

**NATIONAL INSTITUTE FOR HEALTH AND CARE  
EXCELLENCE**

**Final appraisal determination**

**Total hip replacement and resurfacing  
arthroplasty for end-stage arthritis of the  
hip (review of technology appraisal  
guidance 2 and 44)**

**1 Guidance**

- 1.1 Protheses for total hip replacement and resurfacing arthroplasty are recommended as treatment options for people with end-stage arthritis of the hip only if the protheses have rates (or projected rates) of revision of 5% or less at 10 years.

**2 Clinical need and practice**

- 2.1 Arthritis refers to inflammation of a joint, and is a leading cause of pain and disability in the UK. Arthritis can have many causes, the most common of which is osteoarthritis (defined by a loss of cartilage within the joint and related changes in the associated bone). Over 2.8 million people in the UK are estimated to have osteoarthritis. The second most common cause is rheumatoid arthritis (an autoimmune inflammatory disease that affects the synovial lining of joints). Around 400,000 people in the UK have rheumatoid arthritis.
- 2.2 Symptoms of hip arthritis include pain and stiffness that limit daily activities such as walking, climbing stairs and performing household tasks. The diagnosis of arthritis of the hip is usually based on individual patient history and clinical examination

assessing joint pain, deformity and reduced range of movement.

[Osteoarthritis: the care and management of osteoarthritis in adults](#)

(NICE clinical guideline 59) states that clinicians should first offer patients non-surgical treatments including exercise, physical therapy and analgesics, and should consider referring patients for joint replacement surgery if they have ongoing pain, joint stiffness, reduced function and a poor quality of life. People having elective primary surgery to relieve pain and disability caused by arthritis of the hip may receive either a total replacement of the damaged hip (total hip replacement) or a hip resurfacing arthroplasty.

- 2.3 The National Joint Registry 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 private practice, and to monitor the performance of joint replacement prostheses. Since 2009, all NHS patients who are having hip replacement surgery are invited to fill in Patient Reported Outcome Measures (PROMs) questionnaires about their health and quality of life before and after their surgery.
- 2.4 Following publication of [Guidance on the selection of prostheses for primary total hip replacement](#) (NICE technology appraisal guidance 2), the National Health Purchasing and Supply Agency (PASA) was given the task of monitoring adherence to the technology appraisal recommendation. PASA set up a panel of experts known as the Orthopaedic Data Evaluation Panel (ODEP). PASA was subsequently replaced by NHS Supply Chain, which still manages and provides administrative support to ODEP. ODEP provides the NHS with an approved list of prostheses that meet the revision rate standard at 10 years set out in NICE guidance and which are suitable for use in primary hip replacement (see section 3.6). ODEP provides separate ratings for the 2 components

of hip replacement prostheses (that is, stems and cups; see section 3.1). For hip prostheses with less than 10 years of clinical data, there are currently 3 entry standards expressed by ODEP as failure rate: 3% or less at 3 years; 5% or less at 5 years; and 7% or less at 7 years, which are considered to be consistent with the 10-year standard.

### **3 The technologies**

- 3.1 In total hip replacement (THR) surgery, the acetabulum (hip socket) is replaced with either a single-piece cup made from 1 material (polyethylene, ceramic or metal) or a 2-piece (modular) cup made from a metal outer shell and a polyethylene, ceramic or metal liner. The head of the femur (thigh bone) is replaced with either a single-piece metal stem and head, or a modular component consisting of a metal stem (which may consist of more than 1 piece) with a metal, ceramic or ceramicised metal head.
- 3.2 THRs vary in what fixation method is used for each component of the prosthesis. In some THRs, all the components are fixed into position using cement (hereafter referred to as cemented THRs). Other types of THR are designed to be used without cement (hereafter referred to as cementless THRs); instead, they are inserted using press-fit fixation, and natural bone growth over time secures the prosthesis in place. Some prostheses are hybrid, in which the femoral component is cemented into place while the cup is fixed without cement, or reverse hybrid, in which the femoral component is fixed without cement while the cup is cemented into place. THRs may also vary by femoral head size, with a large head defined as being 36 mm or more in diameter.
- 3.3 Hip resurfacing arthroplasty involves removing and replacing the surface of the femoral head with a hollow metal hemisphere, which fits into a metal cup fixed into the acetabulum. All resurfacing

arthroplasty prostheses currently on the market are metal-on-metal (MoM), and can be hybrid or cementless. As with THR prostheses, resurfacing arthroplasty prostheses may also vary by femoral head size.

- 3.4 Patient selection for THR or resurfacing arthroplasty depends on various factors, including but not limited to: patient characteristics (for example a patient's age, activity and underlying hip physiology); the surgeon's choice; and the surgeon's experience of using a particular class of prosthesis.
- 3.5 Adverse events associated with hip replacement surgery (THR or resurfacing arthroplasty) may occur because of complications at the time of surgery, or may occur years afterwards. Complications that may lead to hip replacement revision surgery include prosthesis instability, dislocation, aseptic loosening, osteolysis (bone reabsorption), infection and prosthesis failure.
- 3.6 [Guidance on selection of prostheses for primary total hip replacement](#) (NICE technology appraisal guidance 2) recommends that the best prostheses should have a revision rate of 10% or less at 10 years or, as a minimum, a 3-year revision rate consistent with this. [Guidance on the use of metal on metal hip resurfacing arthroplasty](#) (NICE technology appraisal guidance 44) recommends MoM resurfacing arthroplasty as an option for people with advanced hip disease who would otherwise receive, and are likely to outlive, a conventional primary THR. The guidance recommends that the best prostheses should demonstrate the same revision rates as recommended in NICE technology appraisal guidance 2.
- 3.7 The Medicines and Healthcare products Regulatory Agency (MHRA) monitors the safety of devices used in clinical practice. In June 2010, the MHRA issued an alert on all MoM hip replacement prostheses (both THR and resurfacing arthroplasty) after reports of

soft tissue reactions that may be associated with pain. In June 2012, the MHRA released an updated alert noting that MoM prostheses (THR and resurfacing arthroplasty) may wear at an accelerated rate. The MHRA stated that people with MoM prostheses may develop soft tissue damage caused by wear debris from these prostheses. It advised annual monitoring of the hip using imaging and measurement of metal levels in the blood to determine whether a revision is needed in people with MoM hip replacement prostheses who have symptoms, or who have a certain type of MoM hip replacement, including stemmed MoM THRs with a larger femoral head (36 mm diameter or more) or the recalled DePuy ASR hip replacements (THR and resurfacing arthroplasty).

- 3.8 Over 20 companies manufacture prostheses for hip replacement (THR and resurfacing arthroplasty), and some produce multiple brands of components. The NHS Supply Chain provided the average list price costs for 5 manufacturers of the 5 THR categories (varying by fixation method and bearing surface) identified by the Assessment Group and 3 manufacturers of resurfacing arthroplasty prostheses and associated accessories. The average list prices for THRs across the manufacturers were: £1557 for a cemented polyethylene cup plus a metal head; £3016 for a cementless metal cup with a polyethylene liner plus a metal head (cementless stem); £3869 for a cementless metal cup with a ceramic liner plus a ceramic head (cementless stem); £2650 for hybrid cementless metal cup with a polyethylene liner plus a metal head (cemented stem); and £1996 for cemented polyethylene cup with ceramic head (cemented stem). The average list price for resurfacing arthroplasty prostheses across the manufacturers was £2672. Typically, the price of hip replacement prostheses depends on the volume ordered and locally negotiated discounts, so the

prices paid by the NHS are routinely lower than the list prices listed above.

## **4 Evidence and interpretation**

The Appraisal Committee (section 9) considered evidence from a number of sources (section 10).

### **4.1 Clinical effectiveness**

4.1.1 The Assessment Group conducted a systematic review of randomised controlled trials (RCTs), published systematic reviews, and published registry studies of hip replacement procedures. In addition, the Assessment Group analysed individual patient data from the National Joint Registry (NJR).

#### **Systematic review of randomised controlled trials and published systematic reviews**

4.1.2 The Assessment Group identified 16 RCTs and 8 systematic reviews. It noted that there were a further 20 ongoing clinical trials. Three RCTs and 3 systematic reviews compared resurfacing arthroplasty with total hip replacement (THR); and 13 RCTs and 5 systematic reviews compared different types of THR with each other.

4.1.3 The Assessment Group assessed the risk of bias and methodological quality of the studies (RCTs and systematic reviews), determining whether the evidence could be considered conclusive or non-conclusive based on the precision, consistency and clinical relevance of the effects. The Assessment Group recognised that studies included different measures of patient function and chose, based on previously published research, the following criteria for minimally clinically important differences (MCID): the Harris Hip Score (MCID range: 7–10); the Oxford Hip Score (MCID range: 5–7); the Western Ontario McMaster

Osteoarthritis Index (MCID: 8); and the EQ-5D measure of health-related quality of life (MCID: 0.074). The Assessment Group considered the evidence from an RCT to be conclusive if it showed:

- a statistically significantly different effect between treatments for which the 95% confidence interval included the MCID **or**
- no effect if the MCID was outside the 95% confidence interval for any given outcome.

The Assessment Group considered the evidence from an RCT to be inconclusive if:

- the confidence intervals were wide **or**
- there were missing data **or**
- the effects were inconsistent, if there were 2 separate trials that had assessed the same outcome.

The Assessment Group further considered the evidence from a systematic review to be inconclusive if it:

- did not report pooled results of RCTs (that is, it reported the results narratively) **or**
- used inappropriate methods to pool data **or**
- reported inconsistent summary findings.

### ***Resurfacing arthroplasty compared with total hip replacement***

4.1.4 Of the 3 RCTs comparing the effectiveness of resurfacing arthroplasty with THR, 1 RCT compared metal-on-metal (MoM) resurfacing arthroplasty with large-head MoM THR, 1 RCT compared MoM resurfacing arthroplasty with MoM THR, and 1 RCT compared MoM resurfacing arthroplasty with an unspecified bearing surface of THR. The 3 RCTs randomised a total of 422 patients (ranging from 104 to 192 per study) and the length of follow-up in the trials ranged from 1 to 6 years.

4.1.5 The reported outcomes in the 3 RCTs comparing resurfacing arthroplasty with THR were function (assessed in 3 RCTs), risk of revision (assessed in 1 RCT), infection (assessed in 2 RCTs), aseptic loosening (assessed in 1 RCT), dislocation (assessed in 2 RCTs), deep vein thrombosis (assessed in 2 RCTs) and health-related quality of life (assessed in 2 RCTs; 1 used the EQ-5D and 1 used the SF-36 questionnaire). Five functional measures were used across the 3 RCTs. There was no difference between resurfacing arthroplasty and THR for the Oxford Hip Score, Western Ontario McMaster Osteoarthritis Index score, or the Merle D'Abigine and Postel score. The evidence was inconclusive for the Harris Hip Score and the University of California, Los Angeles activity score. The Assessment Group reported that infection rates differed between patients who had resurfacing arthroplasty and those who had THR. The Assessment Group's meta-analysis of the 2 RCTs that assessed this outcome indicated that, 12-56 months after surgery, patients who had had THR developed more infections than patients who had had resurfacing arthroplasty (pooled odds ratio 7.94, 95% confidence interval [CI] 1.78 to 35.40). All data for the other outcomes (quality of life, revision dislocation, deep vein thrombosis, wound complication, aseptic loosening and mortality) reported in the 3 RCTs were inconclusive.

4.1.6 Of the 3 systematic reviews comparing the effectiveness of resurfacing arthroplasty with THR, 2 synthesised data on function, 2 on risk of revision, 1 on infection, 2 on aseptic loosening, 2 on dislocation and 2 on mortality. The systematic reviews included data from both RCTs and observational studies, including single-arm studies of resurfacing arthroplasty or THR. Two of the systematic reviews assessed resurfacing arthroplasty compared with all types of THR and 1 systematic review compared resurfacing arthroplasty with cementless THR. Two of the systematic reviews included RCTs that the Assessment Group had



critiqued separately. The Assessment Group considered the reported data on function to be inconclusive. The 2 systematic reviews that compared revision rates between resurfacing arthroplasty and THR showed that revision rates were higher after resurfacing arthroplasty (1 estimated a relative risk [RR] of 2.60 [95% CI 1.31 to 5.15] over a 10-year follow-up, 1 estimated an RR of 1.72 [95% CI 1.20 to 2.45] but did not report length of follow-up). Two systematic reviews found that resurfacing arthroplasty was associated with more component loosening than THR (RR 3.00, 95% CI 1.11 to 8.50 and RR 4.96, 95% CI 1.82 to 13.50 respectively). Both of these systematic reviews assessed dislocation rates and 1 found statistically significantly lower dislocation rates associated with resurfacing arthroplasty compared with THR (RR 0.20, 95% CI 0.10 to 0.5). The Assessment Group considered the reported data on all of the other outcomes (mortality, prosthesis failure and infection) to be inconclusive.

### ***Comparison of different types of total hip replacement***

- 4.1.7 The Assessment Group identified 13 RCTs comparing different types of THR with each other, including comparisons of different fixation methods, bearing surfaces, component materials, designs and component sizes. The number of people in each RCT ranged from 100 to 557. The length of follow-up ranged from 3 months to 20 years. Reported outcomes across the RCTs varied and included function, revision, osteolysis (bone reabsorption), aseptic loosening, infection, mortality, femoral fracture, dislocation, deep vein thrombosis, femoral head penetration (prosthesis movement) and quality of life (using SF-12).
- 4.1.8 Four of the RCTs compared THRs with different fixation methods. Of these, 2 compared cemented with cementless cup fixation, 1 compared cemented with cementless femoral stem fixation and 1 compared cemented with cementless cup and femoral stem

fixation. The Assessment Group reported that cemented cups had a lower risk of dislocation compared with cementless cups; its pooled estimate of the odds ratio for the 2 RCTs was 0.34 (95% CI 0.13 to 0.89). The Assessment Group found no other differences between the fixation methods.

- 4.1.9 Six of the RCTs compared THR prostheses with different bearing surfaces, comparing: cross-linked polyethylene with non-cross-linked polyethylene cup liners (2 RCTs); oxinium with cobalt-chromium femoral heads (1 RCT); ceramic-on-ceramic with metal-on-polyethylene femoral head on cup liners (1 RCT); ceramic-on-ceramic with ceramic-on-polyethylene femoral head on cup liners (1 RCT); and steel-on-polyethylene with cobalt-chromium on cross-linked polyethylene and with cobalt-chromium-on-polyethylene femoral head on cup liners. One RCT with 10 years' follow-up, which assessed revision rates, found that THR prostheses with cross-linked polyethylene cup liners had lower revision rates than THRs with non-cross-linked polyethylene cup liners (RR 0.18, 95% CI 0.04 to 0.78). One RCT with 10 years' follow-up found that there was a lower risk of osteolysis with a ceramic-on-ceramic head on cup liner bearing surface than a metal-on-polyethylene femoral head on cup liner bearing surface (RR 0.10, 95% CI 0.02 to 0.32). One RCT with 2 years' follow-up found that steel-on-polyethylene and cobalt-chromium on cross-linked polyethylene femoral head on cup liner bearing surfaces both had a lower rate of femoral head penetration than cobalt-chromium-on-polyethylene or oxinium-on-polyethylene femoral head on cup liner bearing surfaces ( $p < 0.001$ ). There were no other differences reported in the RCTs that assessed THRs with different bearing surfaces.

- 4.1.10 The Assessment Group reported results from 4 other RCTs that compared different types of THR. One RCT compared THRs with

different cup shell designs (porous coated cups compared with arc-deposited hydroxyapatite coated cups). One RCT compared THRs with femoral stems made from cobalt-chromium or titanium. One RCT compared femoral stems with a short metaphyseal fitting with conventional metaphyseal and diaphyseal filling. One RCT compared THRs using a 36-mm femoral head with THRs using a 28-mm femoral head. The Assessment Group reported that the RCT comparing different femoral head sizes found a decreased risk of dislocation associated with 36-mm femoral heads compared with 28-mm femoral heads over a 1-year follow-up (RR 0.17, 95% CI 0.04 to 0.78). No other conclusive differences were reported in these 4 RCTs.

- 4.1.11 The primary focus of the 5 systematic reviews evaluating different types of THR was the comparison of different cup fixation methods (cemented compared with cementless), and the materials used for prosthesis articulation with respect to the postoperative clinical function scores and revision rates. The Assessment Group considered most of the evidence to be inconclusive because the reviews had either reported only a narrative synthesis, or 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 cup liner compared with a non-zirconia head-on-polyethylene cup liner (pooled difference in frequency of revisions over the studies' follow-up periods was 0.02, 95% CI -0.01 to 0.06).

### **Systematic review of registry studies**

- 4.1.12 The Assessment Group reviewed studies based on registries of THR or resurfacing arthroplasty for people with end-stage arthritis of the hip. It identified 30 studies from a number of countries, which

reported different outcomes, had different durations of follow-up, and made different comparisons.

- 4.1.13 The Assessment Group identified 8 registry studies reporting on resurfacing arthroplasty. An analysis of the NJR in England and Wales showed that women had a 30% greater risk of revision with resurfacing than men (hazard ratio [HR] 1.30, 99% CI 1.01 to 1.76). Three of the 4 that compared revision rates between resurfacing arthroplasty and THR found that resurfacing arthroplasty had a higher revision rate than THR. A further analysis of the NJR showed that, although in women resurfacing always had higher revision rates than THR, in men resurfacing arthroplasty prostheses with a larger head size (54 mm) had similar predicted 5-year revision rates to THR prostheses. One study suggested that the risk of revision with resurfacing arthroplasty varied by country, and another study demonstrated lower revision rates in specialist compared with non-specialist centres.
- 4.1.14 The Assessment Group identified 22 registry studies that reported only on THR and that presented analyses of either trends in revision rates or comparisons of revision rates across different types of THR. One study using NJR data from England and Wales (Smith et al. 2012) and 1 using combined data from registries from England, Wales, Australia and New Zealand assessed whether there was an association between femoral head size and revision rates for THR; the studies demonstrated that the relationship was dependent on bearing surface. Both studies showed that the revision rate for MoM THR increased as the femoral head size increased. 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. One study (an analysis of the NJR by McMinn et al. 2012) showed, at a maximum of 8 years' follow-up, a higher mortality rate for patients having cemented

compared with cementless THR (adjusted HR 1.11, 95% CI 1.07 to 1.16).

- 4.1.15 The Assessment Group noted that, of the registries of joint replacement worldwide, the Swedish registry is the oldest. The Assessment Group presented data on revision rates using up to 19 years of follow-up from the Swedish registry for THR and resurfacing arthroplasty grouped together, but noted that these revision rates may include devices and practices no longer in use. The data suggested that revision rates depended on a patient's age at primary surgery. At a maximum of 19 years' follow-up, for people younger than 50 years at primary surgery, 39.8% of women and 37.4% of men had a revision; for people aged between 50 and 59 years, 26.3% of women and 32.8% of men had a revision; for people aged between 60 and 75 years, 12.8% of women and 19.5% of men had a revision; and for people over 75 years, 5.2% of women and 7.9% of men had a revision.

#### **Retrospective cohort analysis of individual patient data from the National Joint Registry**

- 4.1.16 The Assessment Group performed a retrospective cohort analysis of the NJR to estimate revision rates for the different types of prostheses for both populations in the final scope issued by NICE (that is, people for whom both resurfacing arthroplasty and THR were suitable and people for whom only THR was suitable). The Assessment Group obtained individual patient data from the NJR that included data from 2003 to September 2012 and for operations carried out in the NHS and in private practice.
- 4.1.17 The final scope issued by NICE stipulated that different types of hip replacements should be considered separately, if evidence allows. The Assessment Group, advised by its clinical adviser, grouped the types of most commonly used THR into 7 categories. 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. These 5 categories of THR prosthesis accounted for 62% of THRs in the NJR with available data. The categories were:

- category A: cemented polyethylene cup with a metal head (cemented stem)
- category B: cementless hydroxyapatite coated metal cup (with a polyethylene liner) with a metal head
- category C: cementless hydroxyapatite coated metal cup (with a polyethylene liner) with a ceramic head
- category D: cementless hydroxyapatite coated metal cup (with a polyethylene liner) with a metal head (cemented stem)
- category E: cemented polyethylene with a ceramic head (cemented stem).

4.1.18 The Assessment Group addressed the population for whom either resurfacing arthroplasty or THR was suitable. It noted that [NICE technology appraisal guidance 44](#) recommended resurfacing arthroplasty for people who would otherwise receive and outlive a conventional primary THR. This population primarily consisted of people younger than 65 years. The Assessment Group also stated that clinical opinion holds that clinicians offer resurfacing arthroplasty mainly to relatively active younger people, while THR is the usual option for less active older people. The Assessment Group noted that the NJR data did not include data on activity levels. In the absence of data on activity levels, the Assessment Group determined the suitability of resurfacing arthroplasty based on age and sex, and sampled people who had had THR who shared these characteristics. The mean age of this population was 55.8 years and 35% were women.

4.1.19 The Assessment Group addressed the population for whom resurfacing arthroplasty was not suitable. The Assessment Group noted that most people who had THR documented in the NJR were older than 65 years but considered that, because there had been high revision rates after resurfacing arthroplasty, in the future fewer younger people may be considered as candidates for both procedures. As a result, the Assessment Group considered that the population for whom resurfacing arthroplasty was not suitable could be assumed to match the population who had THR documented in the NJR. The mean age of this population was 71.6 years and 64% were women.

***Assessment Group analysis of revision rates of prostheses in the National Joint Registry***

4.1.20 The Assessment Group analysed revision rates using the available data from the NJR (maximum follow-up of 9 years) using Kaplan–Meier estimates. For the population for whom both resurfacing arthroplasty and THR were suitable, the population was matched by age and sex. For the population for whom resurfacing was not suitable, the population was not matched by age and sex and the Kaplan–Meier estimates were not adjusted for these characteristics. The Assessment Group found that, consistent with previous published analyses of the NJR, the revision rate for resurfacing arthroplasty over 9 years of follow-up was about 3 times higher than for all the types of THR prostheses recorded in the NJR. The difference was even larger when comparing resurfacing arthroplasty with THR restricted to the 5 commonly used THR combinations (prosthesis categories A to E, see section 4.1.17). The Assessment Group presented data on revision rates for men and women separately. Revision rates for resurfacing arthroplasty unadjusted for age were higher for women (18% at 9 years) than for men (7% at 9 years). The Assessment Group performed additional analyses in which it excluded data from the

8.8% of people who had the now-recalled DePuy ASR resurfacing prosthesis. Although this lowered the revision rate for resurfacing arthroplasty slightly, the difference between the revision rates for resurfacing arthroplasty and THR remained large.

- 4.1.21 The Assessment Group assessed the time to revision for the 5 categories of THR (A to E) separately. The Assessment Group noted that the revision rate for the cementless prostheses (category C) were higher than for the cemented prostheses (category E and category A). The Assessment Group noted that revision of each category of prosthesis appeared to occur more frequently for men who had any of the prostheses in these 5 categories than for women.
- 4.1.22 To extrapolate revision rates beyond the up-to-9-year data in the NJR, the Assessment Group assessed the fit of various parametric models to the Kaplan–Meier analyses. The Assessment Group noted that, while the bathtub and log-normal models appear to fit the Kaplan–Meier values of revision, after extrapolation these models generated different revision rates. The Assessment Group noted an increasing risk of revision over time with the bathtub model and a decreasing risk of revision over time with the log-normal model. The Assessment Group considered that whether a person underwent revision surgery or not depended both on why the prosthesis had failed and on a person’s suitability for revision surgery. The Assessment Group concluded that, for younger people, the risk of needing a revision would increase over time (because the risk of outliving the prosthesis would increase) and that, for older people, the risks of revision would decrease over time (because the risks of revision surgery might outweigh the benefits). The Assessment Group further concluded that, in active people, prostheses would be more likely to wear out and need revision. The Assessment Group used the bathtub model in its



base case and the log-normal model in its sensitivity analyses of revision rates in people who were over 65 years when they had their THR.

4.1.23 For the population for whom both resurfacing arthroplasty and THR were suitable, the bathtub model predicted revision rates at 10 years of 17.2% and 4.6% for resurfacing arthroplasty and THR respectively, and at 20 years of 48.3% and 12.9% respectively. For the population for whom resurfacing arthroplasty was not suitable, the bathtub model predicted revision rates (unadjusted for age and sex) at 10 years of 2.8% for category A prostheses, 3.9% for category B, 4.6% for category C, 3.0% for category D and 2.1% for category E. The model predicted revision rates at 20 years ranging from 5.2% for category E to 12.3% for category C. The Assessment Group repeated its analysis for the population for whom resurfacing was not suitable, adjusting the bathtub model for age and sex. It found that the relative revision rates across all 5 prosthesis categories were maintained after this adjustment.

4.1.24 For the population for whom both resurfacing arthroplasty and THR were suitable, the Assessment Group predicted revision rates separately for women and men unadjusted for age. In people who had resurfacing arthroplasty, women had higher predicted revision rates at 10-, 20- and 30-year follow-up than men. The estimated 10-year revision rates with resurfacing arthroplasty were 23.1% for women and 12.4% for men.

4.1.25 In the population for whom resurfacing arthroplasty was not suitable, the Assessment Group explored a scenario in which the revision rate in people over 65 years who had THR decreased over time (see section 4.1.22). Using a log-normal distribution and stratifying by sex, the Assessment Group observed lower predicted revision rates compared with the bathtub model. The Assessment Group presented estimates of revision for the mean age in each

category. For men over 65 years, the 10-year modelled revision rates for the 5 THR categories ranged from 1.9% (category E) to 3.9% (category C). For women aged over 65 years, the modelled 10-year revision rates for the 5 THR categories ranged from 1.4% (category E) to 2.8% (category B).

- 4.1.26 The Assessment Group stated that a new rate, setting a standard revision rate for prostheses lower than that of the current standard of less than 10% at 10 years, is appropriate (see section 3.6). The Assessment Group noted that most THR prostheses currently meet this standard, but that most resurfacing arthroplasty prostheses do not.

#### **Manufacturer's clinical-effectiveness evidence**

- 4.1.27 NICE received submissions from 4 manufacturers (DePuy Synthes, JRI, Smith & Nephew and Stryker). The Assessment Group critiqued the submissions and noted that 1 of the 4 manufacturers had performed a systematic review of clinical effectiveness of resurfacing arthroplasty and THR, and that the other 3 manufacturers had provided a narrative review.
- 4.1.28 The manufacturers commented on the difficulties with categorising different types of THR. In particular, 3 manufacturers noted variability in how well different prostheses perform within a category and that some individual manufacturer's brands may have lower revision rates than is typical of their category as a whole. One manufacturer commented that the 7-year revision rates for the 4 most commonly used cementless prostheses range from 2.6% to 4.1%. Another manufacturer noted that data from the NJR showed that its own resurfacing arthroplasty prosthesis, the Birmingham hip resurfacing system, had a revision rate at 7 years that was consistent with the NICE 10% at 10 years standard (it had a revision rate of 5.1%, 95% CI 4.6 to 5.6). Two manufacturers

further stated that categorising by fixation method only may not capture the differences in revision rates that have been seen with different bearing surfaces.

- 4.1.29 Several manufacturers highlighted that the NJR data may not be sufficiently mature to capture changes in risk with different hip prostheses over time. The manufacturers noted that the NJR, the Swedish registry and the Australian registry all showed lower revision rates with cemented prostheses than cementless prostheses in the shorter term after primary surgery, but suggested that this trend may not be maintained if people in the NJR are followed up for longer. The manufacturers highlighted that, after 8 years, the Swedish data showed the risk of revision was higher with cemented than cementless prostheses and, after 6 years, the Australian data showed that cemented THR had a higher revision rate than cementless THR.

## **4.2 Cost effectiveness**

### **Assessment Group's economic model**

- 4.2.1 The Assessment Group developed a Markov model based on the model described by Fitzpatrick et al. (1998), which it adapted to address the decision problem and updated 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 personal social services. The Assessment Group reported results for both a lifetime (80 years) and a 10-year time horizon.
- 4.2.2 Two simulated cohorts entered the model, one of people for whom resurfacing was suitable, reflected by people in England and Wales who underwent resurfacing arthroplasty between 2003 and 2012 (age and sex matched with people who had THR categories A–E; see section 4.1.18); and the other of people for whom resurfacing

arthroplasty was not suitable, represented by people in England and Wales who had THR categories A–E between 2003 and 2012 (see section 4.1.19).

4.2.3 People entered the model at the point of their primary procedure (resurfacing arthroplasty or THR) and moved either to the ‘successful primary’ health state (that is, after successful initial primary surgery) or death. If primary hip replacement failed, people who needed revision moved to the ‘revision total hip replacement state’, received a THR (rather than resurfacing arthroplasty) and stayed in that state for 1 cycle (1 year). If revision was successful, people moved to the ‘successful revision health state’. People in the model could have multiple revisions. The Assessment Group assumed that all sequelae of THR (surgical mortality after primary THR, revision THR or re-revision THR; risk of re-revision) occurred at the beginning of a cycle, and that mortality not related to hip replacement occurred at the end of a cycle.

4.2.4 The transition probability between successful primary surgery and revision THR was based on the revision rates calculated and extrapolated from the NJR data. The Assessment Group based the transition probability between successful revision and further revision THR on the New Zealand Joint Registry (risk of re-revision per procedure 0.0326). The Assessment Group assumed that mortality associated with surgery was 0.5% per procedure (based on the NJR annual report 2012) and used data from the Office for National Statistics on death rates in England and Wales to determine all-cause mortality by age.

4.2.5 To determine the utility associated with each health state, the Assessment Group used the NJR Patient Reported Outcomes Measures (PROMS) database, which reported EQ-5D-3L data post operation by age and sex for the year 2010/2011. The utility values applied in the ‘successful primary’ health state were 0.726 for

people aged between 40 and 50 years; 0.753 for people aged between 50 and 60 years, 0.779 for people aged between 60 and 70 years, 0.764 for people aged between 70 and 80 years, and 0.721 for people aged between 80 and 90 years. The Assessment Group adjusted the utility values for the increasing age of the cohort after every 10 cycles of the model. The Assessment Group assumed that the utility values for people in the 'successful primary health state' were equivalent for people who had resurfacing arthroplasty or THR. The utility value in the 'revision THR health state' was 0.5624 and did not differ by type of THR, age or sex. The Assessment Group assumed that the utility value for a successful revision was the same as for successful primary surgery.

4.2.6 Costs in the model included the costs of the surgery, prostheses, hospitalisation and follow-up. The Assessment Group assumed that the cost of surgery was the same for both THR and resurfacing arthroplasty, 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 £2805.

4.2.7 The Assessment Group obtained the costs of prostheses from the NHS supply chain (see section 3.8). To compare resurfacing arthroplasty with THR for people for whom resurfacing arthroplasty is suitable, the Assessment Group combined the 5 categories of THR prostheses (see section 4.1.17) and generated a weighted average cost based on the frequency of use (from NJR data) of £2571 for THR categories A to E combined. Cemented prostheses needed an additional cost for cement and its preparation (£203.10 for prostheses in which both the stem and cup need cementing and £163.90 for prostheses in which only the stem needs cementing).

- 4.2.8 The Assessment Group derived postoperative hospital costs from Edlin et al. (2012), an RCT that reported the costs of resurfacing arthroplasty and THR over 1 year. The Assessment Group estimated the average cost per day of a hospital stay at £296. People who had resurfacing arthroplasty stayed an average of 5.5 days and people who had a THR stayed an average of 5.7 days, resulting in an overall cost for hospital stays of £1628 for resurfacing arthroplasty and £1687 for THR. Edlin et al. also provided outpatient costs for follow-up after primary THR or resurfacing arthroplasty. The costs over the first 12 months of outpatient care, primary and community care, aids and adaptations provided by the NHS, pain relief and other medications, adjusted for inflation from 2009/2010 to 2011/2012 prices, totalled £501 for resurfacing arthroplasty and £394 for THR. The Assessment Group applied follow-up costs for all consecutive years for the lifetime of the model.
- 4.2.9 The Assessment Group assumed that the costs of revision were the same for THR and resurfacing arthroplasty but depend on the reason for revision (Vanhegan et al. 2012). For example, surgery for infection and peri-prosthetic fracture resulted in longer operating times and lengths of stay than other reasons for revision. Vanhegan et al. reported costs of revision including the costs of the prostheses, materials, theatre, recovery room, inpatient physiotherapy, occupational therapy, pharmacy, radiology and laboratory, with costs based on the NHS 2007/2008 Payment by Results. The Assessment Group adjusted these costs for inflation to 2011/2012 prices by using the projected health service cost index. To derive the 'weighted 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 problems in Vanhegan et al. The Assessment Group applied

the follow-up costs from Edlin et al. (£394; see section 4.2.8) to the successful revision health state.

- 4.2.10 For the population for whom resurfacing arthroplasty and THR were both suitable, the Assessment Group presented deterministic and probabilistic analyses for both a 10-year and a lifetime time horizon. In both the deterministic and probabilistic base case, THR dominated resurfacing arthroplasty (that is, it was less costly and more effective) over both the 10-year and the lifetime time horizons.
- 4.2.11 For the population for whom resurfacing arthroplasty was not suitable, the Assessment Group presented deterministic and probabilistic analyses for both a 10-year and a lifetime time horizon. For a lifetime time horizon, the deterministic incremental analysis showed that THR category E dominated all of the other THR categories. The Assessment Group commented that the difference in quality-adjusted life years (QALYs) was negligible between THR categories A to E (a difference of 0.0064 between the most effective prosthesis category [E] and the least effective prosthesis category [C] in the lifetime deterministic analysis) and that the probabilistic analyses of costs and effectiveness showed that total costs and total QALYs of all categories overlapped.
- 4.2.12 The Assessment Group performed 3 scenario analyses for the population for whom resurfacing arthroplasty and THR were suitable. One scenario analysis tested assumptions used to determine time to revision, and 2 scenarios tested assumptions on the costs of the prostheses. For both the 10-year and lifetime time horizons, all scenario analyses had a minimal effect on incremental costs and QALYs, and the results were consistent with the base case because THR continued to dominate resurfacing arthroplasty.

4.2.13 The Assessment Group performed 7 sensitivity analyses for the population for whom resurfacing arthroplasty was not suitable. Three tested the assumptions used to extrapolate time to revision (including adjusting the analysis for age and sex), 3 tested assumptions on the costs of the prostheses, and 1 tested assumptions on the source of utility values for the successful primary and successful revision health states. The Assessment Group presented results for a 10-year and a lifetime time horizon. For a lifetime time horizon, THR category E continued to dominate all other categories in the following sensitivity analyses: time to revision (bathtub model controlled for age and sex); all 3 cost sensitivity analyses (unadjusted for age and sex with the highest and lowest costs of THR or a 20% discount applied to each prosthesis category); and postoperative utility values (taken from a Swedish cohort study rather than from NJR PROMS data). For the 2 scenarios in which the Assessment Group used the log-normal (rather than the bathtub) model to extrapolate long-term revision rates (1 in which the log-normal model was adjusted for age and sex and 1 in which the model was unadjusted for these characteristics), THR category E was more costly and more effective than category A in the lifetime time horizon (deterministic incremental cost-effectiveness ratio [ICER] £442,830 per QALY gained for the unadjusted model; deterministic ICER £227,031 per QALY gained for the log-normal model adjusted for age and gender). In these log-normal model scenario analyses, THR categories D, B and C continued to be dominated by category E in both the deterministic and probabilistic results.

4.2.14 The Assessment Group performed subgroup analyses for men and women by age for whom resurfacing arthroplasty was suitable. The Assessment Group presented results for each sex stratified by 3 discrete ages, applying a weighting to the modelled revision rates for these subgroups for ages 40, 50 and 60 years. For all ages and



in both men and women, THR dominated resurfacing arthroplasty over both the 10-year and lifetime time horizons.

- 4.2.15 For people for whom resurfacing arthroplasty was not suitable, the Assessment Group presented results for 4 subgroups (men and women under 65 years, and men and women aged over 65 years). For men and women under 65 years, it presented the results for people aged 40, 50 and 60 years separately. For men and women over 65 years, it presented the results for people aged 70 and 80 years separately. For men and women under 65 years, the Assessment Group used the bathtub-modelled revision rates and, for men and women over 65 years, the Assessment Group used the log-normal modelled revision rates. At a lifetime time horizon for men and women aged 70 and 80 years, THR category E was more costly and more effective (QALY difference ranged from 0.0001 and 0.0002) than category A, and dominated categories D, B and C. For women under 65 years, all other categories were dominated by category E. For men aged 40 years, all other 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 dominated categories D, C and B over the lifetime time horizon.

#### **Manufacturer's economic model**

- 4.2.16 Only 1 manufacturer (DePuy Synthes) that made a submission for the appraisal included an economic model.
- 4.2.17 DePuy Synthes developed a transition-state Markov model that had 3 monthly cycles and a lifetime horizon (all patients were assumed to have died by age 100 years). 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 (see section 4.2.3), but the model allowed each patient a maximum of 4 surgical revisions.

- 4.2.18 In the DePuy Synthes model, the populations in the final scope issued by NICE were defined based on the patient characteristics of people in the NJR. The population for whom both resurfacing arthroplasty and THR were suitable was the population in the NJR who had resurfacing arthroplasty. The mean age in this population was 55.3 years and 70.9% were men. The population for whom resurfacing arthroplasty was not suitable was patients in the NJR who had THR. The mean age of this population was 70.4 years and 37.5% were men.
- 4.2.19 For both populations, DePuy Synthes 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 (1 cemented and 1 cementless). DePuy Synthes excluded MoM THR from its analyses, stating that THR using these bearing surfaces are no longer commercially available.
- 4.2.20 DePuy Synthes used individual patient data from the NJR, including data for its own prosthesis brands grouped separately to estimate revision rates for up to 8 years' follow-up. It excluded incomplete entries and those in which osteoarthritis of the hip was not the indication for surgery. DePuy Synthes stated that previous models of revision had fitted different parametric distributions to the periods early and later after surgery, and separately categorised the causes of earlier or later revision. Reasons for early revision included dislocation, mismatch, infection, incorrect sizing and malalignment. Reasons for later revision included fracture of the prosthesis, lysis, pain, acetabular wear, dissociation of the liner, soft tissue reaction and 'other'. DePuy Synthes assessed models that would fit early revisions, late revisions and both combined. It used a Weibull model with a decreasing hazard over time, which it considered realistic for most prosthesis types with the possible exception of cemented prostheses because data from the

Australian registry had shown that the risk of revision with cemented prostheses increases over time.

- 4.2.21 DePuy Synthes based the rate of re-revision (revision subsequent to a first revision) on the New Zealand Joint Registry data (rate 0.0331 per year; 0.0083 per cycle). People stayed in the THR revision/re-revision health state for 1 cycle. The model allowed people to have up to 2 interventions in the same cycle. DePuy Synthes assumed that all people would receive the same type of prosthesis in revision surgery. DePuy Synthes assumed that mortality associated with surgery did not differ by type of prosthesis (0.5% per procedure), and applied an age- and sex-adjusted all-cause mortality rate.
- 4.2.22 The model included the costs of both prostheses and surgery. DePuy Synthes obtained the costs of the prostheses from its own list prices and assumed equal costs for resurfacing arthroplasty and cemented THR. The total prosthesis costs were: cemented THR £1029.00; cementless £2550.50; hybrid £2011.50; and reverse hybrid £1568.00. For the group 'all THR', the manufacturer used a weighted cost (40% cemented, 40% cementless, 17% hybrid, 2% reverse hybrid). DePuy Synthes obtained surgical costs from a micro-costing study that included the costs of anaesthetics, surgical consumables, staff and theatre time. These costs differed across prosthesis type and are academic in confidence. The manufacturer based length of stay on NHS reference costs.
- 4.2.23 DePuy Synthes did not model surgical and post-surgical complications, stating that the average cost and health-related quality of life reflected complications during surgery, and estimates of the risk of revision included complications that occurred after surgery.

- 4.2.24 The manufacturer assumed that the cost of revision was £13,399.42 (which was double the mean cost of the primary procedure). However, unlike the Assessment Group, DePuy Synthes assumed that the cost of revision did not depend on the reason for revision.
- 4.2.25 DePuy Synthes performed a systematic review to identify utility values. For its base case, DePuy Synthes used utility values from Rolfson et al. (2011, Swedish registry). The preoperative utility value was 0.41, and the postoperative utility value was 0.78. It applied a disutility of 0.145 (Briggs et al. 2003) to the postoperative utility value after revision to reflect the lower quality of life associated with a subsequent surgical intervention.
- 4.2.26 In the DePuy Synthes base case for the population for whom both resurfacing arthroplasty and THR were suitable, THR (all types combined) dominated resurfacing arthroplasty. The total incremental cost of resurfacing arthroplasty was £2504.31 for 0.106 fewer QALYS. An incremental analysis calculated using the results for cemented, cementless, hybrid, reverse hybrid and resurfacing prosthesis categories, but excluding DePuy Synthes' own brands (because the costs and QALYs were marked as commercial in confidence and cannot be reported), showed that cemented prostheses dominated both cementless THR and resurfacing arthroplasty. Reverse hybrid prostheses were shown to be extendedly dominated (that is, were dominated by the combination of cemented and hybrid prostheses). The ICER for hybrid prostheses compared with cemented prostheses was £26,636 per QALY gained.
- 4.2.27 For the population for whom resurfacing arthroplasty was not suitable, DePuy Synthes presented an incremental analysis of the results for cemented, reverse hybrid and cementless hybrid prostheses alongside the results for 2 of its own products and all

THR prostheses combined. The results of the incremental analysis for the THR prosthesis categories only showed that cemented prostheses dominated reverse hybrid and cementless prostheses (the results for the manufacturer's own products cannot be reported here because the costs and QALYs are commercial in confidence). The ICER for hybrid prostheses compared with cemented prostheses was £259,667 per QALY gained. The manufacturer noted that the range of QALYs generated by the probabilistic analysis from 10,000 simulations overlapped substantially between the THR prosthesis categories, and concluded that all categories of THR are associated with a similar number of QALYs.

- 4.2.28 DePuy Synthes conducted a number of one-way sensitivity analyses for the population for whom both resurfacing arthroplasty and THR were suitable. It presented the results in net monetary benefit, assuming a maximum acceptable ICER of £20,000 per QALY gained. There was a positive net monetary benefit associated with THR for all parameter values tested. This meant that THR is cost effective compared with resurfacing arthroplasty, given a maximum acceptable ICER of £20,000 per QALY gained. 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. DePuy Synthes also conducted sensitivity analyses for both the population for whom resurfacing arthroplasty and THR were suitable and for the population for whom resurfacing arthroplasty was not suitable, including: using NHS reference costs rather than costs from the micro-costing study; using EQ-5D from the NJR rather than the Swedish registry data; using an exponential rather than a Weibull model to extrapolate revision rate data; and stratifying the population to include people under 70 years or under

55 years. In all scenarios for both populations, the impact on total costs and total QALYs were minimal.

### **4.3 Consideration of the evidence**

- 4.3.1 The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of THR and hip resurfacing arthroplasty for people with end-stage arthritis of the hip for whom non-surgical management has failed. It considered evidence on the nature of surgery for the treatment of pain and disability, and the value placed on the benefits of THR and resurfacing arthroplasty by people needing surgery. It also took into account the effective use of NHS resources.
- 4.3.2 The Committee considered the care pathway for people with end-stage arthritis of the hip and the potential place of THR and resurfacing arthroplasty. The Committee discussed the factors that clinicians take into account when deciding whether to offer a THR or resurfacing arthroplasty to individual patients. The Committee heard from the Assessment Group's clinical adviser that the use of resurfacing prostheses has declined over the past few years, noting the Medicines and Healthcare products Regulatory Agency's alerts to recall some resurfacing prostheses and to monitor patients with metal-on-metal prostheses. The Committee heard that, after any type of hip replacement, some people need revision surgery to replace the primary prosthesis, and that being younger or more active can increase a person's risk of needing revision surgery. The Committee heard that clinicians take into account a person's risk of needing revision surgery when deciding whether to offer resurfacing arthroplasty or THR, and that clinicians in general consider resurfacing arthroplasty more suitable for younger and more active people. The Committee further heard that clinicians may be more likely to offer resurfacing arthroplasty to men than to women because higher revision rates have been observed in

women, which may be associated with women tending to have smaller hips. The Assessment Group's clinical adviser also explained that, because older patients have shorter life expectancies than younger patients, they are less likely to need revision, and that clinicians tend to offer older patients THR. The Committee concluded that both THR and resurfacing arthroplasty are options for treatment of end-stage arthritis of the hip, and that clinicians consider together with patients the factors associated with the risk of revision when choosing the most appropriate procedure.

- 4.3.3 Having considered which type of prosthesis would be appropriate (THR or resurfacing arthroplasty), the Committee then considered the choice of a given prosthesis, noting that prostheses vary in materials and fixation methods. The Committee heard that the operating surgeon generally chooses the type of prosthesis, taking into consideration those that achieve the recommended standard revision rate as provided by the Orthopaedic Data Evaluation Panel (ODEP). The Committee heard that surgeons need specific training for each class of prosthesis (for example, cemented or cementless THR), but that most orthopaedic surgeons in the UK are trained to use both cemented and cementless prostheses. The Committee further heard that an orthopaedic centre's experience and clinical data for individual prostheses further influence the choice of prosthesis. The Committee noted that the NJR contained data for hip replacements carried out in the NHS and in private practice, but heard that the prostheses used in the 2 healthcare sectors were not expected to differ because the same surgeons work in both the NHS and in private practice. The Committee considered whether surgeons offer cemented prostheses and cementless prostheses to different patients, and heard from the manufacturers and the Assessment Group's clinical adviser that there were no specific groups of patients for whom cemented or cementless prostheses

would be specifically indicated. The Assessment Group's clinical adviser explained that a patient's age, sex and activity levels may influence a surgeon's choice of bearing surface for THR. The Committee understood that surgeons tend to choose not only the type but also the brand of hip prosthesis a patient receives, and that this is driven by factors including the surgeon's training, perception of which prostheses perform best, clinical data and experience using different prostheses, among other factors.

4.3.4 The Committee heard from the Assessment Group's clinical adviser that revision surgery is more complex and associated with greater risks than primary THR or resurfacing arthroplasty. It heard from the clinical specialist that patients may need to be referred to a specialist centre for revision surgery. The Committee discussed whether any particular type of THR or resurfacing prosthesis reduced the complexity of subsequent revision surgery, and heard that resurfacing prostheses tended to be easier to replace than THR prostheses, but that the risks associated with surgery to the patients were similar. The Assessment Group's clinical adviser stated that a patient's operative and peri-operative risk depends on why the primary prosthesis failed (for example, infection or fracture) than with the type of prosthesis, or whether it is cemented or cementless. The Committee recognised that revision surgery is more complicated than primary surgery and concluded that the complexity of the revision surgery is primarily determined by why the primary hip replacement failed.

4.3.5 The Committee considered the clinical evidence available for this appraisal. It noted that the Assessment Group presented evidence from RCTs, systematic reviews, published registry studies and its analysis of data from the NJR, and discussed the relevance of each source to its decision making. The Committee noted the Assessment Group's concerns that the RCTs and systematic



reviews it had identified involved small numbers of patients, had relatively short follow-up, reported different outcomes either incompletely or poorly, and were underpowered to detect differences in rates of revision. The Committee accepted that, because of these concerns, it was appropriate that the Assessment Group chose not to meta-analyse the RCTs. The Committee then considered data from registries. It noted that the Assessment Group's retrospective analysis of the NJR provided a record of the revision rates for all types of prostheses used in England and Wales since 2003, and as such provided long-term data generalisable to UK clinical practice. The Committee was aware that, although it is mandatory for NHS organisations to submit data to the NJR, when the registry first started clinicians provided data voluntarily, and that the registry may have missed some procedures that were carried out at the time. The Committee noted that the registry did not provide data on outcomes listed in the scope other than revision, and that it did not provide data on differences in the patient characteristics (for example, activity levels and comorbidities) that might affect both device choice and the risk for revision, causing confounding. The Committee noted the comments received on the appraisal consultation document, stating that there is a problem with an accurate link between the NJR data and Hospital Episode Statistics data, and that data on revision rates from the NJR had not been validated. The Committee concluded that it was appropriate to use both trial and observational data in its decision making, but that uncertainty resulting from the possibility of confounding should be taken into account. The Committee agreed that, although the NJR data had limitations, they are the most comprehensive data reflecting clinical practice in the NHS and therefore the most appropriate for decision making.

4.3.6 The Committee considered the population for whom both procedures are suitable, and the population for whom resurfacing arthroplasty is not suitable. The Committee discussed the Assessment Group's analysis of revision rates of different types of hip replacement in both populations using the NJR data, and whether it had controlled for bias by confounding. The Committee noted that the Assessment Group had controlled for patient age and sex when comparing resurfacing arthroplasty with THR and when comparing different types of THR (in a sensitivity analysis [see section 4.1.23]). The Committee also noted that the Assessment Group's analysis of the NJR in revision after resurfacing arthroplasty compared with THR was consistent with effect measures from RCTs and systematic reviews (see section 4.1.20). The Committee had heard that activity levels influence the choice of whether a person would be offered resurfacing arthroplasty, or which bearing surface of a THR is chosen, and would also affect the rate of wear of a prosthesis (see sections 4.3.2 and 4.3.3), but that the NJR did not contain data on activity. The Committee discussed whether having resurfacing arthroplasty rather than THR would allow people to be more active after their surgery. They heard from the Assessment Group's clinical specialist that observational studies had shown that people were more active after resurfacing arthroplasty than after THR, but were likely to have been more active before resurfacing arthroplasty compared with people who underwent THR, and that 1 RCT showed no difference in activity levels after surgery in people randomised to resurfacing arthroplasty or to THR. The Committee agreed that there was uncertainty around whether the difference in revision rates between THR and resurfacing arthroplasty could just be attributed to risk of failure of the prostheses because it is likely that people who have resurfacing are more active than people who have THR and higher activity may

cause accelerated wear of a prosthesis. The Committee also heard that comorbidities may influence which type of prosthesis a patient receives and whether or not a patient is offered revision surgery. The Committee concluded that the Assessment Group's analysis of revision rates were consistent with published systematic reviews of trials, and controlled for some, but not all, potential confounders, notably activity level and comorbidities, and therefore uncertainty remained surrounding the relative revision rates between different types of prostheses.

- 4.3.7 The Committee considered whether data on revision surgery in the Assessment Group's NJR data set could be considered a proxy for prosthesis failure. The Committee noted that the NJR captured revision rates, but not failure rates of the prostheses, and that some people need revision surgery for pain only (without the prosthesis failing). The Committee further noted that there are people who need a revision because their prosthesis has failed, but who are not fit enough to have surgery or who choose not to have surgery. The Committee appreciated that, in these people, the NJR data on revision rates may underestimate the true failure rate. After consultation on the appraisal consultation document, the Committee further considered revisions that result from prostheses failing and revisions that result from complications during surgery or errors in prosthesis insertion (early revision). The Committee heard from the manufacturers that they expected the proportions of revisions not directly related to device failures to be similar across classes of hip replacement prostheses. The Committee noted a comment received from a manufacturer during consultation stating that early failures associated with dislocation were the fault of the surgeon, but the Committee had no further evidence to support this conclusion. The Committee appreciated that the underlying reason for why a patient needed revision surgery may be difficult to identify and is not routinely recorded in the NJR. In addition, the Committee

was told that there is no system that collects data about the prevalence of people living with a failed prosthesis who are unable to, or chose not to, have revision surgery, and no representative data on the proportion of revisions that are a result of failing prostheses. The Committee accepted that, while revision rates may not fully reflect prosthesis failure, revision was an important outcome both from the patient's perspective and in terms of costs and the resources needed.

- 4.3.8 The Committee considered the approaches to modelling revision rates beyond the maximum 9 years of follow-up in the NJR. It discussed the bathtub model and the log-normal model used by the Assessment Group in its base case, and sensitivity analysis, and the Weibull model used by the manufacturer in its base case. The Committee noted that the bathtub model, which it understood was widely used in manufacturing to describe device failure, assumed that risk of revision would decrease initially and then increase over time, whereas the log-normal and Weibull models assumed an increasing risk of revision over time. The Committee compared the revision rates predicted by all 3 models with data from the Swedish registry, in which people aged between 60 and 75 years who had a hip replacement (resurfacing arthroplasty or THR) were followed up for 19 years. In the population for whom resurfacing was not suitable, the bathtub model predicted longer-term outcomes that fitted the data from the Swedish registry better than the log-normal model. The manufacturer's Weibull model did not fit the Swedish data as well as the Assessment Group's bathtub model did. The Committee noted that there was uncertainty surrounding the generalisability of the Swedish registry data to the UK population, in part because the Swedish registry was initiated earlier than the NJR. The Committee noted that the revision rates in the Swedish registry were higher than the revision rates predicted by the 3 models used to extrapolate data from the NJR. The Committee

concluded that, of the 3 models presented to extrapolate revision rates beyond the 9-year follow-up of the NJR, the Assessment Group's bathtub extrapolation was the most plausible.

- 4.3.9 The Committee examined the economic modelling that had been carried out for the appraisal. The Committee noted that the 2 economic models presented by the Assessment Group and by 1 manufacturer (DePuy Synthes) had similar structures and were based on a model structure that had been used in previous health economic evaluations of hip replacement prostheses. The Committee concluded that the outlined structure of the models adhered to the NICE reference case for economic analysis and was acceptable for the purpose outlined in the scope.
- 4.3.10 The Committee considered the utility values and the source of the health-related quality-of-life data used by the Assessment Group and the manufacturer. The Committee observed that, in both models, the differences in QALYs gained between the types of hip replacement were very small (see section 4.2.11). The Committee discussed how different types of hip replacement surgery would affect a patient's quality of life. The Committee noted that the Assessment Group's utility values came from PROMs in the NJR and were collected postoperatively, but were not specific to individual types of prosthesis. The Committee noted that, in the manufacturer's model, different types of THR and resurfacing arthroplasty were also associated with the same utility value after surgery. The Committee noted that, in the manufacturer's model, a disutility of 0.145 had been applied after a successful revision. This was to reflect that a patient is unlikely to return to the level of health-related quality of life experienced after the primary surgery, whereas the Assessment Group had assumed that utility after a successful revision would be the same as utility after a successful primary hip replacement. The Committee heard from the

Assessment Group's clinical specialist and the manufacturer that, although a successful primary hip replacement would be expected to relieve pain and disability associated with end-stage arthritis of the hip completely, revision surgery was associated with both greater risks and poorer functional outcomes than primary surgery and it was appropriate to apply a disutility value in the post-revision health state, as in the manufacturer's model. The Committee concluded that it was plausible that people who had revision surgery would have a lower quality of life than people who had a successful primary hip replacement. It further concluded that, given the available evidence, it was not possible to determine how use of different types of hip replacement prostheses would affect quality of life.

- 4.3.11 The Committee discussed the costs of the prostheses. It understood that the [Guide to the Methods of Technology Appraisal 2008](#) recommends using public list prices in the reference-case analysis, but noted that the NHS routinely pays a lower price for hip replacement prostheses because of volume-dependent and locally negotiated discounts. The Committee was aware that the Assessment Group obtained an average of sample list prices from the NHS Supply Chain for multiple manufacturers, and that the manufacturer had presented list prices for its own brands. The Committee also noted that the Assessment Group's prices were higher than the manufacturer's (with some exceptions). The Committee concluded that there was considerable uncertainty surrounding the prices of prostheses.
- 4.3.12 The Committee considered the base-case economic analyses presented by the Assessment Group and 1 manufacturer. It noted that they generated broadly similar results, that is, THR dominated resurfacing arthroplasty in both the Assessment Group's and manufacturer's base cases, and that resurfacing arthroplasty

remained dominated in every sensitivity and subgroup analysis. The Committee also noted that, although the categories of THR differed in the Assessment Group's and manufacturer's analyses, cemented prostheses tended to be the least costly and most effective, but with small incremental differences in costs and QALYs compared with other types of THR. The Committee also noted that, in the analyses of cost effectiveness, the Assessment Group and manufacturer used the average revision rate across category, and that the revision rate was the most important key driver of costs and QALYs in the model. The Committee concluded that THR was more effective and less costly than resurfacing arthroplasty in all analyses, but that the small differences between cemented and cementless THR were associated with uncertainty.

- 4.3.13 The Committee discussed the approach of comparing the cost effectiveness of categories of THR and resurfacing arthroplasty by category rather than by individual brands. The Committee was aware that devices can differ only slightly and that, within each category, there are multiple brands. The Committee further noted comments received on the appraisal consultation document that not all categories of THR had been investigated by the Assessment Group. The Committee was aware that the Assessment Group had assessed the 5 most frequently used combinations of bearing surface and fixation method in the NJR and considered this to be appropriate. The Committee considered that the Assessment Group and manufacturer had not taken into account the uncertainty related to revision rates of different brands of prostheses within a category. The Committee noted that the Assessment Group had modelled a revision rate of 17.2% for men and women at 10 years for resurfacing arthroplasty, and 12.4% for men only (in current practice resurfacing is predominantly used in men), and that these revision rates were higher than the current NICE standard of 10% or less at 10 years in [NICE technology appraisal guidance 2](#) and

[NICE technology appraisal guidance 44](#). However, the Committee noted that 1 manufacturer of resurfacing arthroplasty products had provided evidence that its product had a revision rate lower than the NICE standard. In response to the appraisal consultation document, several consultees emphasised that revision rates vary between different brands of prosthesis within a category. The Committee noted again that making recommendations by revision rate allowed individual brands to be assessed separately. The Committee reiterated that it had considered making recommendations for prosthesis by category based on the average revision rate of multiple brands within a category. However, the Committee chose not to make recommendations by category, having concluded that this would disadvantage individual brands of prostheses with low revision rates, and would give an unfair advantage to individual brands with high revision rates within an overall well-performing category.

- 4.3.14 The Committee considered whether it was still appropriate to recommend a revision rate for prostheses of 10% or less at 10 years, as recommended in [NICE technology appraisal guidance 2](#) and [NICE technology appraisal guidance 44](#). The Committee noted that the Assessment Group, having analysed and extrapolated data from the NJR for the population for whom both procedures were suitable, had estimated that the 10-year revision rate for resurfacing arthroplasty was worse (higher) than the standard, and that the 10-year revision rates for THR were much better (lower) than 10% at 10 years. Furthermore, the Committee noted that, in the population for whom resurfacing was not suitable, the highest estimate across the 5 categories of THR was less than 5% at 10 years. The Committee agreed that the current standard was too high for both populations, and was aware that prostheses become more cost effective the lower the revision rates. Therefore, it discussed how a new standard could be determined with the data



available. The Committee considered that, because all of the categories of THR prostheses for both populations had a predicted revision rate of less than 5% at 10 years, the value reflecting the new standard for THRs should be no higher than 5%. Additionally, it considered that, because the predicted revision rate of THR was less than 5% at 10 years in the population for whom both THR and resurfacing arthroplasty were suitable, the revision rate standard for resurfacing arthroplasty should be the same as that for THRs. The Committee noted that, although the average revision rate was predicted to be 5% or less at 10 years, it was likely that within a category of THR some brands would perform poorly and would not meet this standard. The Committee discussed whether the proposed value should be reduced to even less than 5% to provide a more 'aspirational' standard. However, the Committee acknowledged that limitations in the data available (see section 4.3.6) did not allow it to determine the lowest revision rate that current practice could realistically achieve. The Committee concluded that it was appropriate to recommend that a prosthesis (for either resurfacing arthroplasty or THR) should meet a revision rate of 5% or less at 10 years.

- 4.3.15 The Committee was aware that NICE technology appraisal guidance makes recommendations on the most cost-effective use of NHS resources but does not specify how to implement the guidance. It was also aware that the NICE Implementation Programme supports health and social care organisations to maximise the uptake and use of evidence and guidance. The Committee was further aware that ODEP, which is independent of NICE, currently provides the NHS with a list of prostheses that do or do not meet the standard for revision rates outlined in [NICE technology appraisal guidance 2](#) and [NICE technology appraisal guidance 44](#), and that there are initiatives to improve collecting and disseminating information on revision rates. The Committee

discussed whether, given the current support for implementation available to the NHS, it would be possible to implement guidance in which recommendations depended on prostheses meeting a 5% or less revision rate at 10 years, particularly for brands with less than 10 years of data. The Committee was aware that [NICE technology appraisal guidance 2](#) and [NICE technology appraisal guidance 44](#) considered it reasonable to recommend prostheses with a minimum of 3 years of experience, provided the projected revision rate was consistent with the standard recommended at that time; the Committee considered that this remained appropriate. The Committee noted that the ODEP rating system includes 3 entry revision rate benchmarks assuming a linear relationship between the time since primary hip replacement and the proportion of people who would be expected to have a revision. The Committee agreed that, while other appropriate distributions may exist, the analysis of revision rates presented by the Assessment Group for this appraisal had shown it was reasonable to extrapolate using the bathtub function for prostheses with a follow-up of less than 10 years. Furthermore, the bathtub model accounted for a higher rate of early revisions, which may reflect surgical complications or other factors unrelated to the prosthesis (see sections 4.3.7 and 4.3.8). The Committee preferred the Assessment Group's method of extrapolating revision rates to a linear extrapolation, but was content that ODEP needs to determine the methods with which it estimates revision rates based on the quality of the data provided by the manufacturers and the timing of the reporting of revision rates in clinical practice. The Committee concluded that it would be reasonable to recommend prostheses with less than 10 years of data, provided that the revision rate was, in as much as the shorter term follow up data allow, consistent with 5% or less at 10 years and that the recommendation could be implemented within the current support framework provided by ODEP. It also concluded

that prostheses currently with at least 3 years of data, which estimate a higher than 5% revision rate at 10 years when projected, should not continue to be offered to patients.

- 4.3.16 The Committee considered other aspects of how prostheses are currently being rated and noted comments received from ODEP on the appraisal consultation document, in which ODEP clarified that it gave ratings for stem and cup components individually because of the large number of cup and stem components available and their many combinations. The Committee considered whether the revision rate standard of 5% or less at 10 years should apply to each cup and stem component separately. The Committee agreed that total hip replacement or resurfacing arthroplasty can be considered to meet the revision rate standard of 5% or less at 10 years if all components have an ODEP rating consistent with this standard.
- 4.3.17 The Committee considered the cases in which there may be more than 1 prosthesis suitable for a patient that meets the revision rate standard of 5% or less at 10 years. It was aware that current arrangements of generating ODEP ratings do not provide the NHS with the absolute revision rate for an individual prosthesis but only information on whether or not the standard was achieved, and that this was because ODEP receives revision rates from several registries or published papers each with different volumes of implants making a scientifically robust aggregation difficult. The Committee considered that, if more than 1 prosthesis meets the 5% or less revision rate standard, it would prefer to recommend the most cost-effective prostheses (those with the lowest revision rates) but concluded that, without absolute revision rate data for each hip replacement system, this would not be feasible to implement.

4.3.18 The Committee was aware that, because of uncertainties surrounding the costs of prostheses and the discounts available to the NHS, it was not possible to give an estimate of the mean price paid in the NHS for a given prosthesis. The Committee considered that its recommendations should promote maintaining (at least) the level of discount from prostheses' list prices currently offered to the NHS. The Committee discussed whether, if more than 1 prosthesis meets the 5% or less at 10 years revision rate standard, it should recommend the prosthesis with the lowest acquisition costs. The Committee considered comments received during consultation on the appraisal consultation document. It was aware that the cost of THR and resurfacing arthroplasty included both procedure costs and surgical costs. The Committee noted that the Assessment Group had used published literature to determine surgical costs and had assumed that these would be the same for resurfacing arthroplasty and THR. The Committee also noted that 1 manufacturer (DePuy Synthes) had carried out a costing study to estimate time in surgery and consumables, and that the manufacturer stated that procedure costs differed for resurfacing arthroplasty and for THR, and between various types of THR. The Committee heard from the manufacturers that the cost of a prosthesis may be a small proportion of the tariff paid by the NHS for a hip replacement. The Committee noted that the cost of a prosthesis is included in the fixed NHS tariff. The Committee considered the comments received from consultees on the appraisal consultation document, which stated that the benefits of manufacturer support packages had not been taken into account. However, the Committee concluded that tender costs included training in the use of a prosthesis. The Committee concluded that, although the NHS should be mindful of costs, in situations where multiple prostheses with a revision rate of 5% or less at 10 years are suitable for a patient, it could not currently recommend

selecting a prosthesis with the lowest acquisition cost. The Committee further concluded that the recommended standards for revision rate would encourage manufacturers to maintain training programmes to ensure the lowest revision rates possible for their products.

**Summary of Appraisal Committee’s key conclusions**

TAXXX	Appraisal title:	Section
<b>Key conclusion</b>		
	Prostheses for total hip replacement and resurfacing arthroplasty are recommended as treatment options for people with end-stage arthritis of the hip only if the prostheses have rates (or projected rates) of revision of 5% or less at 10 years.	1.1
	Revision rate was the most important key driver of costs and quality-adjusted life years (QALYs) in the model. The Committee was aware that prostheses become more cost effective the lower the revision rates.	4.3.12
	The Committee considered that, because all of the categories of total hip replacement (THR) prostheses had a predicted revision rate of less than 5% at 10 years, the value reflecting the new standard should be no higher than 5%.	4.3.14
	It considered that, because the predicted revision rate of THR was less than 5% at 10 years in the population for whom both THR and resurfacing arthroplasty were suitable, the revision rate standard for resurfacing arthroplasty should be the same as that for THRs.	

<b>Current practice</b>		
Clinical need of patients, including the availability of alternative treatments	Both THR and resurfacing arthroplasty are options for treating end-stage arthritis of the hip, and clinicians consider together with patients the factors associated with the risk of revision when choosing the most appropriate procedure.	4.3.2
	Clinicians may be more likely to offer resurfacing arthroplasty to men than to women because higher revision rates have been observed in women, which may be associated with women tending to have smaller hips.	4.3.2
	The operating surgeon generally chooses the prosthesis, taking into consideration those that achieve the recommended standard revision rate as provided by the Orthopaedic Data Evaluation Panel. The Committee heard that surgeons need specific training for each class of prosthesis (for example, cemented or cementless THR), but that most orthopaedic surgeons in the UK are trained to use both cemented and cementless prostheses. The Committee also heard that an orthopaedic centre's experience and clinical data for individual prostheses further influence choice of prosthesis.	4.3.3
<b>The technology</b>		
Proposed benefits of the technology  How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	A successful primary hip replacement would be expected to completely relieve pain and disability associated with end-stage arthritis of the hip, and hip resurfacing prostheses tend to be easier to replace than THR prostheses, but the risks associated with surgery are similar.	4.3.10
		4.3.4
What is the position of the treatment in the pathway of care for the condition?	The Committee reviewed the data available on the clinical and cost effectiveness of THR and hip resurfacing arthroplasty for people with end-stage arthritis of the hip for whom non-surgical management has failed.	4.3.1

Adverse reactions	<p>Adverse events associated with hip replacement surgery (THR or resurfacing arthroplasty) may occur because of complications at the time of surgery or many years afterwards. Complications that may lead to hip replacement revision surgery include prosthesis instability, dislocation, aseptic loosening, osteolysis (bone reabsorption), infection and prosthesis failure.</p> <p>The Assessment Group's clinical adviser stated that a patient's operative and peri-operative risk after a revision is associated more with why the primary prosthesis failed (for example, infection or fracture) than with the type of prosthesis, or whether it is cemented or cementless.</p>	<p>3.5</p> <p>4.3.4</p>
<b>Evidence for clinical effectiveness</b>		
Availability, nature and quality of evidence	<p>The Assessment Group presented evidence from RCTs, systematic reviews, published registry studies, and its analysis of data from the NJR. The RCTs and systematic reviews involved small numbers of patients, had relatively short follow-up, reported different outcomes either incompletely or poorly, and were underpowered to detect differences in rates of revision.</p> <p>The Assessment Group's retrospective analysis of the NJR provided a record of the revision rates for all types of prostheses used in England and Wales since 2003 and, as such, provided long-term data generalisable to UK clinical practice.</p> <p>The Committee noted comments received during consultation, stating that there is a problem with an accurate link with Hospital Episode Statistics data and that data on revision rates from the NJR have not been validated.</p> <p>The Committee noted that the registry did not provide data on outcomes listed in the scope other than revision, and that it did not provide data on differences in the patient characteristics (for example, activity level and comorbidities) that might affect both device</p>	<p>4.3.5</p> <p>4.3.5</p> <p>4.3.5</p> <p>4.3.5</p>

	<p>choice and the risk for revision, and could therefore cause confounding. The Committee concluded that it was appropriate to use both trial and observational data in its decision making, but that uncertainty resulting from the possibility of confounding should be taken into account. The Committee agreed that, although the NJR data had limitations, they are the most comprehensive data reflecting clinical practice in the NHS and therefore the most appropriate for decision making.</p>	
<p>Relevance to general clinical practice in the NHS</p>	<p>The Assessment Group's retrospective analysis of the NJR provided a record of the revision rates for all types of prostheses used in England and Wales since 2003 and, as such, provided long-term data generalisable to UK clinical practice.</p>	<p>4.3.5</p>
<p>Uncertainties generated by the evidence</p>	<p>The Committee heard that activity levels influence the choice of whether a person would be offered resurfacing arthroplasty, or which bearing surface of a THR is chosen, and would also affect the rate of wear of a prosthesis but that the NJR did not contain data on activity. It agreed that there was uncertainty around whether the difference in revision rates between THR and resurfacing arthroplasty could be attributed to failure of the prostheses because it is likely that people who have resurfacing are more active than people who have THR and higher activity may cause accelerated wear of a prosthesis.</p> <p>The Committee heard that comorbidities may be associated with which type of prosthesis a patient receives and whether or not a patient is offered revision surgery.</p> <p>The Committee concluded that the Assessment Group's analysis of revision rates controlled for some, but not all, potential confounders, notably activity and comorbidities, and that it was consistent with published systematic reviews of trials, but that there remained uncertainty surrounding the relative revision rates between different types of prostheses.</p>	<p>4.3.6</p> <p>4.3.6</p> <p>4.3.6</p>



<p>Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?</p>	<p>Clinicians may be more likely to offer resurfacing arthroplasty to men than to women because higher revision rates have been observed in women, which may be associated with women tending to have smaller hips.</p>	<p>4.3.2</p>
<p>Estimate of the size of the clinical effectiveness including strength of supporting evidence</p>	<p>The Committee noted that the Assessment Group had modelled a revision rate of 17.2% for men and women at 10 years for resurfacing arthroplasty, and 12.4% for men only (in current practice resurfacing is predominantly used in men), and that these revision rates were higher than the current NICE standard of 10% or less at 10 years in <a href="#">NICE technology appraisal guidance 2</a> and <a href="#">NICE technology appraisal guidance 44</a>. All of the categories of THR prostheses for both populations had a predicted revision rate of less than 5% at 10 years.</p>	<p>4.3.13, 4.3.14</p>
<p><b>For reviews (except rapid reviews):</b> How has the new clinical evidence that has emerged since the original appraisal (TA2 and TA44) influenced the current (preliminary) recommendations?</p>	<p>Since the original appraisals <a href="#">NICE technology appraisal guidance 2</a> and <a href="#">NICE technology appraisal guidance 44</a>, data have become available for revision rates of prostheses used in the NHS and private practice in England and Wales and are documented in the National Joint Registry.</p>	<p>4.3.5</p>
<p><b>Evidence for cost effectiveness</b></p>		
<p>Availability and nature of evidence</p>	<p>The Committee considered the base-case economic analyses presented by the Assessment Group and 1 of the manufacturers (DePuy Synthes).</p>	<p>4.3.12</p>
<p>Uncertainties around and plausibility of assumptions and inputs in the economic model</p>	<p>The Committee understood that the <a href="#">Guide to the Methods of Technology Appraisal 2008</a> recommends using publicly available list prices in the reference-case analysis, but noted that the NHS routinely pays a lower price for hip replacement prostheses because of volume-dependent and locally negotiated discounts. The Committee concluded that there was considerable uncertainty surrounding the prices of prostheses.</p>	<p>4.3.11</p>

<p>Incorporation of health-related quality-of-life benefits and utility values</p> <p>Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?</p>	<p>The Committee concluded that it was plausible that people who had revision surgery would have a lower quality of life than people who had a successful primary hip replacement. It further concluded that, given the available evidence, it was not possible to determine how use of different types of hip replacement prostheses would affect quality of life.</p>	<p>4.3.10</p>
<p>Are there specific groups of people for whom the technology is particularly cost effective?</p>	<p>Not applicable</p>	
<p>What are the key drivers of cost effectiveness?</p>	<p>The Committee noted that, in the analyses of cost effectiveness, the Assessment Group and manufacturer used the average revision rate across category, and that the revision rate was the most important key driver of costs and QALYs in the model.</p> <p>Prostheses become more cost effective the lower the revision rates.</p>	<p>4.3.12</p> <p>4.1.14</p>
<p>Most likely cost-effectiveness estimate (given as an ICER)</p>	<p>ICERs were not the relevant parameter in determining the recommendations. This was because the incremental cost-effectiveness ratios (ICERs) were dependent on the predicted average revision rates of the analysed categories of prostheses, the differences in QALYs between categories were small, and individual brands may have different revision rates from the category average.</p>	<p>4.3.10, 4.3.13</p>

<p><b>For reviews (except rapid reviews):</b> How has the new cost-effectiveness evidence that has emerged since the original appraisal (TA2 and TA44) influenced the current (preliminary) recommendations?</p>	<p>The Committee concluded that THR was more effective and less costly than resurfacing arthroplasty in all analyses, but that the small differences between cemented and cementless prostheses were associated with uncertainty.</p>	4.3.12
	<p>The Committee considered making recommendations for particular prostheses categories based on the point estimate reflecting the average revision rate of multiple brands of prostheses within a category. However, it concluded that this would disadvantage individual brands of prostheses with particularly low revision rates and would give an unfair advantage to individual brands with high revision rates within an overall well-performing category.</p>	4.3.13
	<p>The Committee concluded that it was appropriate to recommend that a prosthesis should meet a revision rate of 5% or less at 10 years.</p>	4.3.14
<p><b>Additional factors taken into account</b></p>		
<p>Patient access schemes (PPRS)</p>	<p>Not applicable</p>	
<p>End-of-life considerations</p>	<p>Not applicable</p>	
<p>Equalities considerations and social value judgements</p>	<p>During scoping, consultees said that the rates of total joint surgery in practice may vary in different groups of people. However, no changes were required to the scope because it did not define the population being considered by any of the protected equality characteristics. It was noted by the Committee that <a href="#">NICE technology appraisal guidance 2</a> and <a href="#">NICE technology appraisal guidance 44</a> were published before the current NICE equalities scheme was implemented. No equality issues were raised in the assessment report, the manufacturer's submissions or during the consultation on the assessment report or the Committee's discussions.</p>	n/a

## 5 Implementation

- 5.1 Section 7(6) of the [National Institute for Health and Care Excellence \(Constitution and Functions\) and the Health and Social Care Information Centre \(Functions\) Regulations 2013](#) requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within 3 months of its date of publication.
- 5.2 When NICE recommends a treatment ‘as an option’, the NHS must make sure it is available within the period set out in the paragraph above. This means that, if a patient has end-stage arthritis of the hip and the doctor responsible for their care thinks that total hip replacement or resurfacing arthroplasty is the right treatment, it should be available for use, in line with NICE’s recommendations.
- 5.3 NICE has developed tools [[link to www.nice.org.uk/guidance/TAXXX](#)] to help organisations put this guidance into practice (listed below). [[NICE to amend list as needed at time of publication](#)]
- Slides highlighting key messages for local discussion.
  - Costing template and report to estimate the national and local savings and costs associated with implementation.
  - Implementation advice on how to put the guidance into practice and national initiatives that support this locally.
  - A costing statement explaining the resource impact of this guidance.
  - Audit support for monitoring local practice.
- 5.4 The [Orthopaedic Data Evaluation Panel](#) (ODEP), hosted and facilitated by the NHS Supply Chain, will provide information about revision rates for hip prostheses, enabling commissioners in the

NHS to comply with this guidance. ODEP coordinates, receives and analyses submissions of long-term performance data from prosthesis manufacturers, used in the UK and internationally. Based on the quality and length of follow-up data, prostheses are rated as per the table below.

**Benchmarks developed by ODEP**

<b>Pre-entry</b>	<p><b>Pre-entry A:</b> Product launched under Beyond Compliance.</p> <p><b>Pre-Entry B:</b> Products registered with the National Joint Registry. All primaries and revisions monitored via supplier feedback.</p>
<b>3 years</b>	<p><b>3A rating:</b> 150 with minimum 3 years' follow up with actual revision rates of less than 3%. All deaths, loss to follow-up failures and indications for revisions recorded.</p> <p><b>3B rating:</b> Data for a smaller cohort demonstrating less than 3% revision rates at 3 years, and Patient Time Incidence Rate (PTIR) or Kaplan–Maier survivorship data showing confidence limits on the data.</p>
<b>5 years</b>	<p><b>5A rating:</b> 250 patients (with data from beyond the developing centre submitted) with minimum 5 years follow up with actual revision rates of less than 5%. All deaths, loss to follow up indications for revisions recorded.</p> <p><b>5B rating:</b> Data for a smaller cohort demonstrating less than 5% revision rates at 5 years, and PTIR or Kaplan–Maier survivorship data showing confidence limits on the data.</p>
<b>7 years</b>	<p><b>7A rating:</b> 350 patients (with data from beyond the developing centre submitted) with a minimum of 7 years follow up with actual revision rates of less than 7%. All deaths, loss to follow up, failures and indications for revisions recorded.</p> <p><b>7B rating:</b> Data for a smaller cohort demonstrating 7% at 7 years, and PTIR or Kaplan–Maier survivorship data showing confidence limits on the data.</p>
<b>10 years</b>	<p><b>10A* rating:</b> 500 patients (including 3 centres in cohort and including data from beyond the developing centres) with a minimum of 10 years follow up with actual revision rates of less than 5% at 10 years (that is, demonstrating survivorship of better than</p>

	<p>95%). All deaths, loss to follow up failures and indications for revision included in data.</p> <p><b>10A rating:</b> 500 patients (including 3 centres in cohort and including data from beyond the developing centres) with a minimum of 10 years follow up with a survivorship of better than 90%. All deaths, loss to follow up, failures and indications for revision included in data.</p> <p><b>10B rating:</b> Data for a smaller cohort demonstrating 10% at 10 years, and PTIR or Kaplan–Maier survivorship data showing confidence limits on the data.</p>
--	--

- 5.5 Both strength of data and length of follow up are considered when awarding a rating to individual stem and cup components. While the current A and B rating system is to be retained to support international users of the ratings, hip prostheses will need to track towards no more than a 5% revision rate over 10 years to achieve a 10A\* rating in line with NICE guidance (implants with a 3A rating, followed by a 5A rating and with a less than 5% revision rate at 7 years are considered on track to meet the 10A\* rating). All hip prostheses are expected to progress through benchmarks from pre-entry with, for example, a 5 year submission is expected within 3 years following award of a 3 year benchmark. Failure to follow this progression will result in products being de-registered on the ODEP website. In addition, the option to show products at pre-entry following the [Beyond Compliance Programme](#) has been included with the introduction of a pre-entry ‘A’ benchmark.
- 5.6 When contracts for the purchase of prostheses exist and cannot be changed, the recommendation applies to all new contracts. However, commissioners should explore whether there is flexibility within such contracts to orientate their buying towards prostheses that meet the updated recommended revision rate standard.

## 6 Recommendations for further research

- 6.1 The Committee recommended that research should be carried out to determine the relationship between activity and prosthesis failure.
- 6.2 The Committee recommended the collection of data on prosthesis failure or on the prevalence of people living with a failed hip but for whom revision surgery is not suitable or who choose not to have revision surgery. The Committee further recommended that nomenclature for hip replacement failure needs to be established to allow demarcation of prosthesis-dependent and prosthesis-independent hip replacement failure. Furthermore, patient reported outcome measures collected as part of the National Joint Registry should allow for reporting of hip replacement failure in people who cannot or chose not to have revision surgery.

## 7 Related NICE guidance

Details are correct at the time of publication. Further information is available on the [NICE website](#).

- [Arthroscopic femoro–acetabular surgery for hip impingement syndrome](#). NICE interventional procedure guidance 408 (2011).
- [Minimally invasive total hip replacement](#). NICE interventional procedure guidance 363 (2010).
- [Rheumatoid arthritis: the management of rheumatoid arthritis in adults](#). NICE clinical guideline 79 (2009).
- [Osteoarthritis: the care and management of osteoarthritis in adults](#). NICE clinical guideline 59 (2008).
- [Guidance on the use of metal on metal hip resurfacing arthroplasty](#). NICE technology appraisal guidance 44 (2002).
- [Guidance on the selection of prostheses for primary total hip replacement](#). NICE technology appraisal guidance 2 (2000).

## **8 Review of guidance**

- 8.1 The guidance on this technology will be considered for review in February 2017. The Guidance Executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Amanda Adler

Chair, Appraisal Committee

December 2013



## **9 Appraisal Committee members, guideline representatives and NICE project team**

### **9.1 *Appraisal Committee members***

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

#### **Dr Amanda Adler (Chair)**

Consultant Physician, Addenbrooke's Hospital

#### **Professor Ken Stein (Vice Chair)**

Professor of Public Health, Peninsula Technology Assessment Group (PenTAG), University of Exeter

#### **Professor Keith Abrams**

Professor of Medical Statistics, University of Leicester

#### **Dr Jeff Aronson**

Reader in Clinical Pharmacology, University Department of Primary Health Care, University of Oxford

**Professor John Cairns**

Professor of Health Economics Public Health and Policy, London School of Hygiene and Tropical Medicine

**Matthew Campbell-Hill**

Lay member

**Mark Chapman**

Health Economics and Market Access Manager, Medtronic UK

**Dr Lisa Cooper**

Echocardiographer, Stockport NHS Foundation Trust

**Dr Maria Dyban**

General Practitioner

**Professor Fergus Gleeson**

Consultant Radiologist, Churchill Hospital, Oxford

**Robert Hinchliffe**

HEFCE Clinical Senior Lecturer in Vascular Surgery and Honorary Consultant Vascular Surgeon, St George's Vascular Institute

**Dr Neil Iosson**

General Practitioner

**Anne Joshua**

Associate Director of Pharmacy, NHS Direct

**Dr Rebecca Kearney**

Clinical Lecturer, University of Warwick

**Dr Miriam McCarthy**

Consultant, Public Health, Public Health Agency

**Professor Ruairidh Milne**

Director of Strategy and Development and Director for Public Health Research

National Institute for Health and Care Excellence

at the National Institute for Health Research (NIHR) Evaluation, Trials and Studies Coordinating Centre at the University of Southampton

**Dr Elizabeth Murray**

Reader in Primary Care, University College London

**Dr Peter Norrie**

Principal Lecturer in Nursing, DeMontfort University

**Christopher O'Regan**

Head of Health Technology Assessment & Outcomes Research, Merck Sharp & Dohme

**Professor Stephen Palmer**

Professor of Health Economics, Centre for Health Economics, University of York

**Dr Sanjeev Patel**

Consultant Physician & Senior Lecturer in Rheumatology, St Helier University Hospital

**Dr John Pounsford**

Consultant Physician, Frenchay Hospital, Bristol

**Dr Danielle Preedy**

Lay Member

**Dr John Rodriguez**

Assistant Director of Public Health, NHS Eastern and Coastal Kent

**Alun Roebuck**

Consultant Nurse in Critical and Acute Care, United Lincolnshire NHS Trust

**Cliff Snelling**

Lay Member

**David Thomson**

Lay member

**Dr Nicky Welton**

Senior Lecturer in Biostatistics/Health Technology Assessment, University of Bristol

**Dr Nerys Woolacott**

Senior Research Fellow, Centre for Health Economics, University of York

## **9.2 NICE project team**

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

### **Mary Hughes**

Technical Lead(s)

### **Nicola Hay**

Technical Adviser

### **Jeremy Powell**

Project Manager

## 10 Sources of evidence considered by the Committee

A. The assessment report for this appraisal was prepared by Warwick Evidence:

- Clarke A, Pulikottil-Jacob R, Grove A. et al. Total hip replacement and surface replacement for the treatment of pain and disability resulting from end stage arthritis of the hip (Review of technology appraisal guidance 2 and 44), July 2013

B. The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, assessment report and the appraisal consultation document (ACD). Organisations listed in I, II and III were also invited to make written submissions and have the opportunity to appeal against the final appraisal determination.

### I. Manufacturers/sponsors:

- B Braun
- Biomet
- DePuy-Synthes (Johnson & Johnson)
- JRI
- Smith & Nephew
- Stryker
- Wright Medical
- Zimmer

### II. Professional/specialist and patient/carer groups:

- Arthritis Care
- Association for Perioperative Practice
- British Hip Society
- British Orthopaedic Association
- Chartered Society of Physiotherapy
- Primary Care Rheumatology Society
- Royal College of Nursing
- Royal College of Physicians

### III. Other consultees:

- Department of Health
- Welsh Government

IV. Commentator organisations (without the right of appeal):

- Association of British Healthcare Industries (ABHI)
- Commissioning Support Appraisals Service
- Department of Health, Social Services and Public Safety for Northern Ireland
- EUCOMED
- Healthcare Improvement Scotland
- Medicines and Healthcare products Regulatory Agency
- NHS Confederation
- NHS Supply Chain
- Orthopaedic Research UK

C. The following individual was selected from clinical specialist and patient expert nominations from the consultees and commentators. They participated in the Appraisal Committee discussions and provided evidence to inform the Appraisal Committee's deliberations. They gave their expert personal view on total hip replacement and resurfacing arthroplasty by attending the initial Committee discussion and/or providing written evidence to the Committee. They were also invited to comment on the ACD.

- Mary Drozd, Senior Lecturer, University of Wolverhampton, nominated by the Royal College of Nursing – clinical specialist

E. Representatives from the following manufacturers/sponsors attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Johnson & Johnson
- JRI
- Smith and Nephew
- Stryker