FINAL Protocol NIHR HTA Programme project number 11/118. Updated following final scope (Dec 2012)

Title

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).

1. Research question

To appraise the clinical and cost effectiveness of total hip replacement and surface replacement CE marked interventions for the treatment of pain and disability resulting from arthritis of the hip for which non-surgical management has failed.

2. Name of TAR team and project 'lead'

Produced by:	Warwick Evidence
	Division of Health Sciences
	Medical School, University of Warwick
	Coventry
	CV4 7AL
Lead Author:	Aileen Clarke
Co-authors:	Paul Sutcliffe
	Amy Grove
	Martin Connock
	Ruth Pulikottil-Jacob
	Gaurav Suri
	Alexander Tsertsvadze
	Rachel Court
	Samantha Johnson
	David Metcalf
	Ngianga-Bakwin Kandala
Correspondence to:	Prof Aileen Clarke, Warwick Evidence, Warwick Medical School, University of Warwick, Coventry, CV4 7AL
Tel:	+44 (0) 24761 50063
Email:	Aileen.Clarke@warwick.ac.uk
Date Completed:	22 November 2012

This project is commissioned by the NIHR HTA Programme as project number 11/118. The views expressed in this protocol are those of the authors and not necessarily those of the NIHR

HTA Programme. Any errors are the responsibility of the authors. The authors have no conflicts of interest.

3. Plain English Summary

Arthritis is a group of diseases that affect joints, leading to pain and disability. Osteoarthritis (OA) is the most prevalent form of arthritis, and the most common reason for having a hip replacement (Clinical Guideline No. 59, 2008).

People with end stage arthritic damage to their hip may receive total replacement of the damaged hip. For the purpose of this appraisal, end stage arthritis of the hip is defined as arthritis of the hip for which non-surgical management has failed. One of the most commonly used type of hip replacement is a metal ball on a stem cemented into the femur and a plastic socket cemented into the pelvis. However, some hip replacements do not use cement and have harder bearing surfaces, for example, metal on metal (MoM), ceramic-on-ceramic or ceramic-on-polyethylene and some are hybrid as demonstrated in Table 1

As an alternative to total hip replacement (THR), patients may receive hip resurfacing arthroplasty which involves removing the damaged surfaces of bones inside the hip joint and cementing a metal surface to the reshaped bone. The socket has a metal surface and is fixed into the pelvis without using cement (Vale et al, 2002). Resurfacing conserves more femoral bone and can result in a greater range of movement after surgery. However it requires patients to have relatively strong bones and tends to be used in younger, more active patients (Vale et al, 2002).

Table 1. Hip arthroplasty in NHS England and Wales 2010

Procedure type	**Procedures conducted in 2010 (%)**
Cemented THR	24,806 (36%)
Un-cemented THR	29,630 (43%)
Hybrid THR	11,025 (16%)
Primary resurfacing	2,067 (3%)
Other	1,378 (2%) large head MoM
Total	68,907

Currently artificial hip joints last an average of 10 to 15 years, some considerably longer. Some hip replacements require revision surgery because of loosening of the joint, wear and tear, pain and dislocation. Current National Institute for Health and Clinical Excellence (NICE) guidance says that the best prostheses should demonstrate a ‘benchmark’ revision rate of 10% or less at ten

years or, as a minimum, a three-year revision rate consistent with this benchmark (Technology Appraisal No. 2, 2000).

In 2011, 57,745 hip procedures were carried out in the NHS in England and Wales, with a further 25,138 carried out in independent hospitals. Ninety three per cent of primary hip replacements were for hips that were affected by osteoarthritis (National Joint Registry, 2012) and six per cent were hip resurfacing arthroplasty. The National Joint Registry (NJR) for England and Wales holds information on hip replacement procedures performed in the NHS and the independent sector in England and Wales since 2003.

For the NHS to allocate and deliver its services optimally, relative benefits and costs of THR and hip resurfacing need to be estimated. Moreover, given technical advances in prosthesis design, it would be useful to know which types of THR and resurfacing confer the most benefit and the least harm. Therefore, this report aims to evaluate the clinical and cost effectiveness of THR and hip resurfacing for the treatment of pain and disability in people with arthritis.

4. Decision problem

In people with pain and disability resulting from arthritis of the hip for which non surgical management has failed:

i. who are suitable for both procedures, what is the clinical effectiveness and cost-effectiveness of different types of elective primary total hip replacement compared to primary hip resurfacing arthroplasty?
ii. who are not suitable for hip resurfacing, what is the clinical effectiveness and cost effectiveness of different types of primary total hip replacement compared with each other

Objectives

1) To undertake a systematic review of the clinical and cost-effectiveness for the following:

a. Different types of primary THR compared with surface replacement for people in whom both procedures are suitable;
b. Different types of primary THR compared with each other for people who are not suitable for hip resurfacing

and to

Investigate factors that influence benefits and costs.

2) To develop the cost-effectiveness and cost-utility models published in the 2002 HTA (Technology Appraisal No. 44, 2002) further using updated National Joint Registry data and model inputs where available.

3) To report on findings and make recommendations for future research

Outcomes for both comparisons (a and b) to be considered will include: revision rates, disability, quality of life (QOL), mortality/survival, functional result, pain, bone conservation, radiosteriometric analysis to asses prosthesis movement, adverse treatment (peri-/post-procedural) degradation products, health related quality of life and mortality.

If data are sufficient, the influence of patient and intervention related factors on the magnitude of treatment effects will be explored through subgroup analysis and meta-regression technique. Economic analysis will be undertaken and the cost-effectiveness and/or cost-utility of treatments will be expressed in terms of incremental cost-effectiveness ratios (ICERs). The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from the NHS and Personal Social Services perspectives.

4.1 Background

Disease epidemiology and burden

Arthritis is a group of diseases that involves inflammation of one or more joints, leading to pain and disability. The most common form is OA, others forms are rheumatoid arthritis (RA), psoriatic arthritis, and related autoimmune diseases (Clinical Guideline No. 59, 2008).

OA refers to a clinical syndrome of joint pain accompanied by varying degrees of functional limitation, and reduced quality of life. Structural changes commonly occur without accompanying symptoms. OA is by far the most common form of arthritis and one of the leading causes of pain and disability worldwide (Clinical Guideline No. 59, 2008). RA is an autoimmune disease causing inflammation of joints and is the second most common form of arthritis with approximately 400,000 people affected in the UK (Clinical Guideline No. 59, 2008).

The exact incidence and prevalence of OA is difficult to determine because the extent of the clinical syndrome (joint pain and stiffness) does not always correspond with structural changes (usually defined as abnormal changes in the appearance of joints) on radiographs (Clinical Guideline No. 59, 2008). Estimates suggest that up to ten million people in the UK are affected by joint pain that may be attributed to osteoarthritis (Arthritis Care, 2012).

OA of the hip is the most common reason for having a hip replacement, but it is also undertaken for others forms of arthritis e.g. RA (deVerteuil et al, 2008). In 2005, OA was the primary diagnosis for 94% of THRs in NHS in England and Wales (deVerteuil et al, 2008).

In 2011, 57,745 hip procedures were carried out in the NHS in England and Wales, with a further 25,138 carried out in independent hospitals. Ninety-three per cent of primary hip replacements were for hips that were affected by osteoarthritis (National Joint Registry 2012).

Risk factors for osteoarthritis include: (Clinical Guideline No. 59, 2008)

- Genetic factors - heritability estimates for hand, knee and hip osteoarthritis are high at 40–60%;
- Constitutional factors - ageing, female sex, bone density;
- Biomechanical factors - joint injury, occupational/recreational usage, reduced muscle strength, joint laxity, joint malalignment;
- Environmental factors - overweight and obesity, muscle weakness, occupational or recreational joint trauma.

Impact

OA predominantly affects older people, and often coexists with other conditions associated with aging and overweight or obesity, as well as with common sensory and psychosocial problems (Clinical Guideline No. 59, 2008). Symptoms including pain, stiffness, joint deformity and loss of joint mobility have a substantial impact on every aspect of a person's daily life, and their overall quality of life. Increases in life expectancy are expected to make OA the fourth leading cause of disability by the year 2020 (deVerteuil et al, 2008). Therefore, OA will have considerable impact on health services.

Severity

The severity of end stage arthritis is assessed using established outcome measures. The three most common are the WOMAC score, Harris Hip Score (HHS) and the Oxford Hip Score (OHS). They asses factors including disability, pain, hip function and symptoms which impact upon activities of daily living. The WOMAC score has been validated for measuring clinically important patient-relevant outcomes in osteoarthritis of the hip or knee (Bellamy et al, 1988). The OHS has been developed and validated specifically to assess function and pain for patients undergoing THR surgery (Dawson et al, 1996). The OHS is the most highly evaluated hip-specific measure available (Murray et al, 2007).

THR - Treatment and technology

THR is carried out to relieve the pain and disability caused by arthritis of the hip, which cannot be managed by pain medication and physiotherapy. The damaged hip joint is replaced with an artificial hip prosthesis. Surgery is undertaken either under general or epidural anaesthesia. The surgeon removes the existing hip joint completely. The upper part of the thigh bone (femur) is removed and the natural socket for the head of the femur is hollowed out. A socket is fitted into the hollow in the pelvis. A short, angled metal shaft with a smooth ball on its upper end (to fit into the socket) is placed into the hollow of the thigh bone. The cup and the artificial bone head may be pressed into place or fixed with acrylic cement.

Many variations of the THR operation exist with differences in the design of the implants and their composition (metal, plastic, ceramic), and whether they are inserted with bone cement or not (un-cemented). There are also different combinations of the implants, producing different bearing surfaces (metal or ceramic-on-plastic, metal-on-metal, ceramic-on-ceramic). In 2010, out of the 68,907 primary hip procedures, 36% were cemented total hip replacements (THRs), 43% were un-cemented THRs and 16% were hybrid procedures, 3% were large head metal on metal THRs and 3% were resurfacing arthroplasty (National Joint Registry, 2012) (Table 1).

Surgeons are able to gain guidance from the Orthopaedic Device Evaluation Panel (ODEP) when selecting implants. ODEP is hosted and facilitated by the NHS Supply Chain and coordinates, receives and analyses submissions of long term performance data from manufacturers. ODEP provides the NHS with an approved list of prostheses which meet the benchmarks set out in NICE guidance and which are suitable for use in primary hip replacement.

Hip resurfacing – Treatment and technology

Hip resurfacing arthroplasty is an option for the treatment of arthritis of the hips where the head of the femur is prepared and a large diameter metal cap is fitted, which articulates with a thin walled metal cup implanted into the acetabulum. The main difference between THR and resurfacing is that during resurfacing much less of the bone is removed, as only the joint surfaces are replaced. Hip resurfacing is reserved for use in younger or more active patients with good bone stock, for whom subsequent revision may be easier and more feasible (deVerteuil et al, 2008). In 2010 hip resurfacing arthroplasty accounted for only 3% of all primary hip procedures (National Joint Registry, 2012). Resurfacing surgery can either be carried out as a standard procedure or as a minimally invasive procedure for which additional NICE guidance has been issued (IPG 363 NHS NICE, 2010).

The National Joint Registry (NJR)

The NJR was established in 2002 with combined efforts from the Department of Health and the Welsh government. The NJR's activity mainly relates to collating the data on joint replacement surgeries and to monitoring implant performance (e.g. ankle, hip, knee, shoulder and elbow). The NJR publishes annual reports which include analysis of the data collected from various data collection units. The 2012 report (NJR, 2012) collated data for 1.2 million patients' surgery and provides updated implant survival rates for patients undergoing implant surgery with the use of robust estimation methods (flexible parametric modelling with competing outcomes (Royston & Parmar, 2002)).

4.2 Scoping searches

We undertook web searches on identified manufacturer websites to establish all known devices and to determine their approval status with the ODEP and Conformité Européenne (CE). The scoping searches identified a range of devices and manufacturers, which were discussed with our clinical advisors. Names and manufacturers of all primary total hip replacement and primary resurfacing head and cup manufacturers are displayed in Appendix 1 (Table 2)

5. Review Methods

A systematic review of the evidence for each treatment will be reported according to the general principles recommended in the PRISMA statement (Moher, 2009 a,b). Previous HTA reports and systematic reviews as well as individual primary studies addressing questions relevant to this review will be identified and summarised in the current report.

5.1 Identification of studies

Initial scoping searches were undertaken in Medline in October 2012 to assess the volume and type of literature relating to the assessment question and to inform further development of the search strategies. A search strategy was then developed which focuses the searches to primary THR and resurfacing (see below). All searches will be undertaken in November 2012.

5.1.1 Search strategy for clinical effectiveness and cost effectiveness

An iterative procedure was used to define the scoping searches with input from clinical advisors and previous HTA reports (e.g. Vale et al, 2002 and deVerteuil et al, 2008).

Copies of the draft clinical and cost effectiveness search strategies that are likely to be used in the major databases are provided in Appendix 2. The search strategies were developed for MEDLINE and will be adapted as appropriate for other databases. The strategies have been designed to capture generic terms for arthritis, THR and surface replacement.

Searches (See Appendix 2) will be date-limited from 2002 (the date of the most recent NICE guidance in this area TA 44) to the present day. Clinical searches are restricted to RCT and systematic review evidence, additional searches may be undertaken to capture literature relating to costs, resources use, utilities, cost effectiveness, cost effectiveness models and registries to inform the cost effectiveness analysis.

All bibliographic records identified through the electronic searches will be collected in a managed reference database.

The search strategy will comprise the following main sources:

- Searching of electronic bibliographic databases including trials in progress
- Contact with experts in the field
- Scrutiny of references of included studies
- Screening of manufacturers' websites for relevant publications

These should allow for identification of relevant published and unpublished studies and studies in progress.

Databases will include:

MEDLINE; MEDLINE In-Process & Other Non-Indexed Citations; EMBASE; Cochrane Database (including Cochrane Systematic Reviews, DARE, NHS EED, and HTA databases); Science Citation Index and Conference Proceedings (Web of Science); UKCRN Portfolio Database; and NLM gateway (US Meeting Abstracts and Health Services Research Projects in Progress) and the CEA Registry. The following trial databases will also be searched: CENTRAL; Current Controlled Trials; and ClinicalTrials.gov. Citation searches of included studies will be undertaken using the Web of Science citation search facility. The reference lists of included studies and relevant review articles will also be checked, and the manufacturers' websites will be screened for relevant publications. Grey literature search will be undertaken using Google Scholar and online resources of various regulatory bodies, health services research agencies and professional societies will be consulted via the Internet.

These are likely to include:

- British Hip Society
- British Orthopaedic Association
- Orthopaedic Research UK
- Orthopaedic Device Evaluation Panel (ODEP)
- National Joint Registry (NJR)
- Arthritis Research UK (ARUK)
- Cochrane Musculoskeletal Group
- Arthritis Care

5.1.2 Inclusion of relevant studies

Study design (clinical effectiveness):

- Randomised controlled trials
- Systematic reviews
- Meta-analyses

Study design (economic evaluation)

- Randomised controlled trials
- Observational designs; cohort studies and registry-based studies
- Decision analytic modelling studies

- Systematic reviews
- Meta- analyses

Population:

People with pain or disability resulting from end stage arthritis of the hip for which non-surgical management has failed.

Intervention:

- Elective primary total hip replacement (THR)
- Primary hip resurfacing arthroplasty

Comparator:

- Different types of primary total hip replacement and hip resurfacing arthroplasty will be compared for people in whom both procedures are suitable.

- Different types of primary total hip replacement will be compared with each other for people in whom hip resurfacing arthroplasty is not suitable.

The different types of hip replacement that will be considered separately are dependent on the available evidence, but may include:

- Hip replacements with components made from different materials (metal, ceramic, polyethylene, ceramicised metal)
- Cemented, cementless or hybrid prostheses
- Prostheses with differing femoral head size
- Prostheses with differing revision rates

Record:

Full text articles of completed or in-progress studies (protocols) published in English.

Outcomes:

The effectiveness outcome measures to be considered include: function, pain, bone conservation, revision rates (device failure/revision rates/time to revision), radiosteriometric analysis (to assess

prosthesis movement), radiological result, dislocation rates, health related quality of life and mortality.

Adverse events will include peri- and post-procedural complications (e.g. infection, nerve palsy, dislocation rates, femoral neck fracture, metallosis, muscle weakness) and metal and other degradation products.

Cost-effectiveness outcomes will include mean difference in costs and clinical effectiveness measures or utility measures; incremental cost-effectiveness ratio (ICER), uncertainty measures, the ceiling willingness-to-pay ratios, and probabilities from cost-effectiveness acceptability curves.

5.1.3 Exclusion criteria

- Indications for hip replacement other than end stage arthritis of the hip
- Patients undergoing revision surgery
- Abstract/conference proceedings, letters, and commentaries
- Non-English language publications

5.2 Study selection process

We will collect all retrieved records in a specialised database and duplicate records will be identified and removed. The reviewers will pilot-test a priori screening form based on the predefined study eligibility criteria. Afterwards, two independent reviewers will apply inclusion/exclusion criteria and screen all identified bibliographic records for title/abstract (level I) and then for full text (level II). Any disagreements over eligibility will be resolved through consensus or by a third party reviewer. Reasons for exclusion of full text papers will be documented. The study flow will be documented using a PRISMA diagram (Moher, 2009 a,b).

5.3 Data extraction strategy

The relevant data will be extracted independently from included studies by one reviewer using a data extraction form informed by the NHS Centre for Reviews and Dissemination (CRD) (Khan, 2011). Uncertainty and/or any disagreements will be crosschecked with a second researcher and will be resolved by discussion. In cases when studies fail to report summary statistics (e.g., mean score, standard deviation, standard error), we will attempt to calculate these parameters if individual participant data is provided. If a study reports only a standard error of the mean

response, we will convert it into a standard deviation. The extracted data will be entered into summary evidence tables (see Appendix 3). The extracted information will include:

- Study characteristics (i.e., author's name, country, design, study setting, sample size, funding source, duration of follow-up, and methodological features such as generation of randomization, allocation concealment, blinding, completeness of outcome ascertainment, patient withdrawals/attrition for randomised trials; for observational studies and non-randomised trials, information on potential confounding will be additionally ascertained)
- Patient baseline characteristics (i.e., inclusion/exclusion criteria, number of enrolled/analysed participants, age, race, gender, body mass index, underlying conditions, concomitant conditions, co-interventions, disability, range of movement, activity levels, function, pain intensity, and quality of life, and disease-specific measures such as Oxford Hip Score, Harris Hip Score)
- Experimental treatment characteristics (e.g., type - THR, resurfacing; training/experience of the operator, post-operative rehabilitation staff; method of fixation – cemented, un-cemented, hybrid; bearing surface material – metal-on-metal, ceramic -on-ceramic; metal-on-polyethylene, femoral head size; the name/brand and country of manufacturer; post-operative rehabilitation)
- Outcome characteristics (e.g., definition; timing of measurement; scale of measurement - dichotomous, continuous; measures of association – mean difference, relative risk, odds ratio, hazard ratio). Measures of variability and statistical tests used will also be extracted (standard deviation, 95% CIs, standard error, p-values

For studies of economic evaluation (cost-effectiveness/utility analyses), the reviewers will extract information on utilities, resources use and costs (both direct and indirect) and on incremental cost-effectiveness ratios, statistical analysis (e.g., bootstrap techniques, number of replications, parametric tests, levels of statistical significance), type of economic evaluation (i.e., cost-effectiveness, cost-utility analysis), perspective (e.g., societal, health care payer, patient), study currency and discounting.

If a study fails to report the incremental cost-effectiveness ratios, the reviewers will attempt to calculate ratios if data allows. All costs will be converted to the United Kingdom Pounds (GBP)

using exchange rates applicable to the end (the month of December) of the year for which the cost estimates in each study were reported (www.xe.com).

5.4 Quality assessment strategy

Risk of bias of individual studies will be assessed with respect to any given outcome. Two reviewers will independently assess risk of bias of included studies using published and validated assessment scales and/or checklists (see Appendix 3). Any disagreements between the two reviewers will be resolved by a third reviewer through discussion.

- Randomised Controlled trials (RCTs) - Cochrane Collaboration Risk of Bias tool (Higgins, 2011) which covers the following domains of threat to validity: selection bias (randomization sequence, allocation concealment), performance bias (blinding of participants/personnel), detection bias (blinding of outcome assessors), attrition bias (incomplete outcome data), reporting bias (selective outcome/analysis reporting), and other pre-specified bias (e.g., Funding source, adequacy of statistical methods used, type of analysis [Intention-to-treat/Per protocol], imbalance in the distribution of baseline prognostic factors between the compared treatment groups).
- Observational cohort studies and non-randomised controlled clinical trials (CCTs) for the cost effectiveness study will be assessed using an adapted Cochrane Risk of Bias tool (with randomisation sequence generation and allocation concealment items removed) (Higgins, 2011).
- Methodological quality of included systematic reviews will be assessed using the AMSTAR tool (Shea, et al 2007 a,b, Shea, et al 2009), which covers the following domains: a) research question, b) inclusion/exclusion criteria, c) search strategy (at least two major electronic databases), d) data extraction by independent reviewers, e) assessment of risk of bias by independent reviewers, f) consideration of risk of bias in the analysis, g) exploration of heterogeneity, and h) publication bias.
- Economic evaluation primary studies (cost-effectiveness analysis) will be assessed using the Drummond checklist (Drummond, 1996).

Further details on the methodological quality/risk of bias assessment instruments are presented in Appendix 4. These may be amended following preliminary extraction of included papers.

5.5 Grading overall quality of evidence strategy

The overall quality of evidence for each pre-selected outcome across studies will be assessed using the systematic approach developed by the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) Working Group (http://www.gradeworkinggroup.org).

The GRADE approach (Guyatt 2011) indicates levels of confidence in the observed treatment effect estimate(s), which is categorized as high, moderate, low, or very low. The grading of overall quality of evidence for each key outcome is based on assessments across five domains: a) overall risk of bias (internal validity across studies; study limitations), b) consistency of results (heterogeneity), c) directness of the evidence (applicability of the results; indirect treatment comparisons), d) precision of the results (the width of 95% CI around the estimate), and e) publication/reporting bias (detection of asymmetry in the funnel plot; selective outcome reporting). Examples and explanations of grading process across the five domains are presented in Appendix 5.

The gradable outcomes for the report were selected based on their meaningfulness and importance for decision-making given the objectives of the review outlined above. The proposed outcomes for this process are revision rates, disability, quality of life (QOL), mortality/ survival, functional limitation, pain and adverse events. GRADEpro software (version 3.2 for Windows. Jan Brozek, Andrew Oxman, Holger Schunemann 2008) will be used to generate results for each graded outcome which will be presented in Evidence Profile (EP) and summary of findings (SoF) tables.

5.6 Methods of analysis and synthesis

Study, treatment, population, and outcome characteristics will be summarised and compared qualitatively in text, summary, and evidence tables. The effectiveness of treatments reported in comparative head-to-head studies will be compared for as follows:

a) Different types of primary THR compared with hip resurfacing arthroplasty (for people in whom both procedures are suitable);

b) Different types of primary THR compared with each other (for people who are not suitable for hip resurfacing);

If the evidence allows subgroups based on activity levels will be compared. If the evidence allows different types of hip prostheses will be considered separately such as:

- Hip replacements with components made from different materials (metal, ceramic, polyethylene, ceramicised metal);
- Cemented, cementless or hybrid prostheses;
- Prostheses with differing femoral head size;
- Prostheses with differing revision rates, for example ODEP ratings.

The collection and monitoring of performance data and arrangements for the effective implementation of such recommendations based on long term performance (revision rates, for example ODEP ratings) will be considered.

The clinical diversity of treatment effect of THR and surface replacement will be assessed across a priori specified subgroups defined by activity levels and function as agreed with clinical experts.

If data allow, study results from RCTs will be pooled in a meta-analysis. The decision to pool individual study results will be based on degree of similarity with respect to methodological and clinical characteristics of studies under consideration (e.g., design, population, comparator treatment, and outcome). The estimates of post-treatment mean difference (MD) for continuous outcomes and relative risk (RR) or hazard ratio (HR) for binary outcomes (except for rare events) of individual studies will be pooled using a DerSimonian and Laird random-effects model (DerSimonian & Laird, 1986). The choice of this model is based on the assumption that some residual clinical and methodological diversity will exist across the pooled studies despite the similarities. Where necessary (zero events in one or both arms of a trial) a continuity correction will applied. For binary outcomes with very low event rates < 1%, Peto odds ratios (ORs) will be pooled.

Trials will not be meta-analysed if the mean and standard deviation for the continuous outcome of interest cannot be ascertained. Trials with obvious between-group baseline imbalance in a continuous outcome will not be pooled unless the mean change from baseline and corresponding standard deviation for the compared study groups are reported or can be reliably calculated from p values.

The degree of statistical heterogeneity across pooled studies will be assessed through visual inspection of the Forest plots, Labbe plots, calculation of Cochrans Q and tau squared statistics for between study variance, and the I^2 statistics. If data allow, the potential clinical and methodological sources of statistically significant heterogeneity (Chi-square $p < 0.10$ and/or $I^2 > 50\%$), will be explored through subgroup (age, sex, activity levels, and function) and sensitivity analyses (e.g., Risk of Bias item, ITT vs. per-protocol), respectively.

The extent of publication reporting bias will be examined by visual inspection, funnel plot asymmetry, and linear regression tests (Egger 1997, for continuous outcomes, Harbord 2006, and or Peters 2006, for dichotomous outcomes), if a sufficient number of data points are available.

If there is lack or insufficient evidence of direct treatment comparison from head-to-head studies (different types of THR vs. resurfacing or THR vs. THR) and if time and data permit, we will attempt to conduct adjusted indirect treatment comparison analysis if there is a common treatment comparator across the studies (Bucher et al, 1997).

6. Report methods for synthesising evidence of cost-effectiveness

Economic analysis will be undertaken and the cost-effectiveness and/or cost-utility of treatments will be expressed in terms of incremental cost-effectiveness ratios (ICER). The time horizon for estimating clinical and cost effectiveness will be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. The cost-effectiveness results will be summarized in text and tables. The structure of the economic evaluation will be informed by previous work undertaken deVerteuil et al, (2008).

6.1 Published economic studies

Published cost-effectiveness studies will be reviewed. All papers which present findings on the cost and outcomes of primary THR compared with surface replacement or different types of primary THR compared with each other, will be reviewed in detail, and a narrative review will be undertaken.

6.2 Economic appraisal

Costs and effectiveness of different types of THR compared with each other for those not eligible for surface replacement and different types of primary THR compared with surface replacement for those who are eligible for both procedures will be estimated for patients with pain or disability

resulting from end stage arthritis of the hip and we will also investigate the factors that drive costs and benefits. We will further develop the cost-effectiveness and cost-utility models developed in the 2002 HTA (Technology Appraisal No. 44, 2002) using updated National Joint Registry data and model inputs where available.

Costs will be obtained from systematic review of published literature, National Schedule of Reference Costs, the National Joint Registry, clinical advisors, industry submissions and other ad hoc studies identified through ODEP (Orthopaedic Device Evaluation Panel). Costs to be considered will include NHS resource use before and after THR or hip resurfacing (e.g. theatre cost, prostheses cost, length of hospital stay and follow-up costs). The perspective of the economic analysis will be that of the NHS and personal social service. Data on clinical and quality of life benefits and on revision rates will be sought from the systematic literature review and from the NJR.

Revision rates

Failure and revision rates are of critical importance in estimating cost-effectiveness of THR and resurfacing. The way in which these rates depend on patient characteristics (including age, gender) and on the type of intervention received is also important. This information may be obtained from the literature and/or from the National Joint Registry.

The main objectives of analysis of the NJR include:

1. To report individual patient baseline characteristics and current epidemiology of the interventions;

2. To extract other relevant covariate information of potential relevance for survival modelling of treatment revision rates (including follow up of patients subsequent to surgery);

3. To carry out analysis of revision rates for patient groups and interventions using flexible parametric survival models (Royston and Parmar 2002) with competing outcomes (deaths);

4. To compare the above estimates with those reported in the literature;

5. To undertake a quality assessment of the National Joint Registry data provided using criteria mentioned in Black et al. (2004)

Data permitting, the analysis of resurfacing and THR revision rates will take account of different types of bearing surface (ceramic on ceramic, resurfacing, metal-on-polyethylene and metal-on-metal and head size head size) and differing sub groups.

Revision rates beyond observed data

Flexible parametric models of failure rates of devices derived from the NJR or from the literature will require extrapolation beyond the observation period to extended time horizons in order that all benefits and harms are captured for the economic analysis. The following approaches will be considered: - a) extrapolation using parametric fits from flexible parametric modelling; b) fitting a "bath tub" hazard profile to the modelled data as suggested by Briggs et al (2004); this may done directly using a bathtub equation such as that represented in Collet (2003) or, following Briggs et al (2004), by combining the Weibull fit to early failures with the Weibull fit for late failures

6.3 Industry submission(s) regarding Effectiveness of treatments

We will compare any submitted industry economic model with our own findings. If our conclusion differs, we will identify the key assumptions that lead to the differences, and comment on the different interpretations of the evidence.

7. Expertise in this TAR team

Warwick Evidence is a technology assessment group located within Warwick Medical School. Warwick Evidence brings together experts in clinical and cost effectiveness reviewing, medical statistics, health economics and modelling. The team planned for the work includes: Dr Paul Sutcliffe, Ms Amy Grove, Dr Martin Connock, Dr David Metcalf, Dr Alexander Tsertsvadze and Professor Aileen Clarke who are experienced systematic reviewers; Ms Samantha Johnson and Ms Rachel Court, information specialists; Professor Aileen Clarke, Dr Martin Connock, Ms Ruth Jacobs, Mr Gaurav Suri, and Dr Ngianga-Bakwin Kandala provide modelling and health economic expertise.

8. Competing interests of authors and advisors

None of the authors have any competing interests. The advisors have not declared any competing interests.

9. Timetable/milestones

The project will be undertaken in phases, including: literature search, study selection, data extraction and critical appraisal, evidence synthesis, and NJR data analysis, model building and report writing. A progress report including a draft clinical effectiveness section will be submitted on the 06/02/13, this is conditional upon the rapid approval of the protocol.

The final assessment report including the clinical and cost-effectiveness sections will be submitted on 17/05/12. There will be fortnightly team meetings and correspondence with the clinical advisors will take place every 2-3 weeks via email.

Draft protocol submitted	02/11/2012
Draft protocol finalised	23/11/2012
Progress report including draft clinical effectiveness section	06/02/2013
Final assessment report including clinical and cost-effectiveness sections	17/05/12

10. Team members' contributions

Research team: Warwick Evidence

Lead	Professor Aileen Clarke
Title:	Professor of Public Health and Health Services Research and Director for Warwick Evidence
Address:	Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel:	02476 150189
Email:	Aileen.Clarke@warwick.ac.uk
Contribution:	Co-ordinate review process, protocol development, data analysis, synthesis of findings and report writing

Name	Dr Paul Sutcliffe
Title:	Associate Professor and Deputy Director for Warwick Evidence
Address:	Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel:	02476 574505
Email:	p.a.sutcliffe@warwick.ac.uk

Contribution: Co-ordinate review process, protocol development, assessment for eligibility, quality assessment of trials, data extraction, data entry, data analysis, and report writing

Name: Ms Amy Grove
Title: Research Project Manager
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 515584
Email: A.L.Grove@warwick.ac.uk
Contribution: Co-ordinate review process, protocol development, assessment for eligibility, quality assessment of trials, data extraction, data entry, data analysis, and report writing

Name: Dr Martin Connock
Title: Senior Research Fellow
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 574940
Email: M.Connock@warwick.ac.uk
Contribution: Protocol development, assessment for eligibility, quality assessment of trials, data extraction, data entry, data analysis, and report writing

Name: Ms Ruth Pulikottil-Jacobs
Title: Research Fellow Health Economics
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 151902
Email: R.Jacob@warwick.ac.uk
Contribution: Health economics modeller, assessment for eligibility and data extraction

Name: Mr Gaurav Suri
Title: Research Associate

Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 73163
Email: G.Suri@warwick.ac.uk
Contribution: Operations research modeller, assessment for eligibility and data extraction

Name: Dr Alexander Tsertsvadze
Title: Honorary Senior Research Fellow
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: N/A
Email: a_tsertsvadze@hotmail.com
Contribution: Protocol development, assessment for eligibility, quality assessment of trials, data extraction, data entry, data analysis, and report writing

Name: Mr David Metcalf
Title: Academic Clinical Fellow
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: N/A
Email: d.metcalfe@doctors.org.uk
Contribution: Assessment for eligibility, quality assessment of trials, data extraction and report writing

Name: Ms Rachel Court
Title: Information Specialist
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 522427
Email: R.A.Court@warwick.ac.uk
Contribution: Protocol development, develop search strategy and undertake the electronic literature searches

Name: Ms Samantha Johnson

Title: Information Specialist
Address: Library Services, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 522427
Email: Samantha.A.Johnson@warwick.ac.uk
Contribution: Protocol development, develop search strategy and undertake the electronic literature searches

Name: Dr Ngianga-Bakwin Kandala
Title: Principal Research Fellow Statistician
Address: Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry CV4 7AL
Tel: 02476 150541
Email: N-B.Kandala@warwick.ac.uk
Contribution: Statistician, Health economics modeller

10.1 Methodological advisors

Contribution of methodological advisor: Dr Ewen Cummins will provide Health Economics advice during the MTA. His contribution will include: using previous experience of modelling in this area to advise on multistate models, general evidence synthesis, and statistical issues in health economic modelling, application of statistical methods to various health care settings.

10.2 Clinical and Technical Advisors

Contribution of clinical and technical advisors: to advise on protocol development, help interpret data, provide a methodological, policy and clinical perspective on data and review development of background information and clinical effectiveness and review of report drafts. Clinical advisors include

- Prof Matt Costa – Professor of Trauma and Academic Orthopaedic Surgery at The University of Warwick,
- Prof Ashley Blom – Head of the Orthopaedic Group of the University of Bristol,
- Prof Alister Hart – University College London Chair of Orthopaedic Surgery Consultant Orthopaedic Surgeon Director of Research & Development

Technical advisor:

- Prof Richie Gill – Professor of Healthcare Engineering University of Bath; provides biomechanical advice.

11. References

Arthritis Care (2012) http://www.arthritiscare.org.uk/Home accessed October 2012.

Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833-40.

Black N, Barker M, Payne M (2004). Cross sectional survey of multicentre clinical databases in the United Kingdom. BMJ; 328(7454):1478.

Briggs A, Sculpher M, Dawson J, Fitzpatrick R, Murray D, Malchau H (2004). The use of probabilistic decision models in technology assessment : the case of total hip replacement. Appl Health Econ Health Policy; 3(2):79-89.

Bucher HC, Guyatt GH, Griffith LE Walter SD (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol; 50(6):683-91.

Clinical Guideline No. 59, Feb 2008, 'Osteoarthritis: The care and management of osteoarthritis in adults CG 59. Review in progress, earliest anticipated date of publication Nov 2013.

Collet, D. (2003). Modelling survival data in medical research. Chapman & Hall. London.
Dawson J., Fitzpatrick R., Murray D., Carr A. (1996) Comparison of measures to assess outcomes in total hip replacement surgery. Qual. Hlth. Care;5:81-88.

de Verteuil R, Imamura M, Zhu S, Glazener C, Fraser C, Munro N, Hutchison J, Grant A, Coyle D, Coyle K, Vale L. A systematic review of the clinical effectiveness and cost-effectiveness and economic modelling of minimal incision total hip replacement approaches in the management of arthritic disease of the hip. Health Technol Assess. 2008 Jun;12(26):iii-iv, ix-223.

DerSimonian R and Laird N, (1986) Meta-analysis in clinical trials. Controlled Clin Trials; 7:177-88.

Drummond MF, (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ;313:275.

Egger M, Davey SG, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ; 315(7109):629-634.

Grading of Recommendations Assessment, Development and Evaluation (GRADE) 2012. http://www.gradeworkinggroup.org accessed November 2012.

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. (2011) GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol; 64: 383-94.

Harbord RM, Egger M, Sterne JA (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med; 25(20):3443-3457.

Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. 2011. URL: www.cochrane-handbook.org (accessed 5 November 2011).

Intervention Procedure Guidance No. 112, Feb (2005) 'Minimally invasive two-incision surgery for total hip replacement'. National Institute for Health and Clinical Excellence. Department of Health Publication

Khan KS, Ter Riet G, Galnville J, Snowdon AJ, Kleijnen J (2001). Undertaking Systematic Reviews of Research on Effectiveness. CRD's Guidance for Carrying Out or Commissioning Reviews. 2 nd Edition. (ISBN 1900640201) CRD Report No. 4. York: NHS Centre for Reviews and Dissemination (CRD), University of York. http://www.york.ac.uk/inst/crd/report4.htm

Moher D, Liberati A, Tetzlaff J, Altman DG (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol;62:1006-12. a.

Moher D, Liberati A, Tetzlaff J, Altman DG (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ ;339:b2535. b.

Murray, D. W., Fitzpatrick, R., Rogers, K., Pandit, H., Beard, D. J., Carr, A. J., and Dawson, J. (2007). The use of the Oxford Hip and Knee Scores. J Bone Joint Surg, 89-B (8), 1010-1014

NJR (2012) The NRR Centre Hemel Hempstead. National Joint Registry for England and Wales Annual Report 2012. Available URL: http://www new.njrcentre.org.uk/njrcenre/Default.aspx Accessed 27/02/12.

Oxman, AD (2004) Grading quality of evidence and strength of recommendations. BMJ, 328;1490

Royston, P and M.K.B. Parmar. (2002). Flexible parametric proportional hazards and proportional odds model for censored survival data with application to prognostic modelling and estimation of treatment effects. Stat. Med. 21: 2175-2192.

Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z et al (2007). External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One; 2(12):e1350.

Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et al (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol; 7:10.

Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J et al (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol; 62(10):1013-1020.

Smith AJ, Dieppe P, Howard PW, Blom AW (2012). Failure rates of metal-on-metal hip resurfacings: analysis of data from the National Joint Registry for England and Wales. The Lancet, 380 (9855), 1759 – 1766.

Smith AJ, Dieppe P, Vernon K, Porter M, Blom AW (2012). Failure rates of stemmed metal-on-metal hip replacements: analysis of data from the National Joint Registry of England and Wales. Lancet; 379(9822):1199-1204.

Technology Appraisal No. 2, Apr (2000), 'Hip disease- replacement prostheses'. National Institute for Health and Clinical Excellence. Department of Health Publication

Technology Appraisal No. 44, Jun (2002), 'Hip disease- metal on metal hip resurfacing'. National Institute for Health and Clinical Excellence. Department of Health Publication

The National Joint Registry (2011) The NRR Centre Hemel Hempstead. National Joint Registry for England and Wales Annual Report (2011). Available URL: http://www new.njrcentre.org.uk/njrcenre/Default.aspx Accessed 27/02/12.

Vale, L, Wyness, L, McCormack, K, McKenzie, L, Brazzelli, M , Stearns, S (2002) A systematic reivew of the effectiveness and cost-effectivenss of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease. Health Technology Assessment, Vol 6: No 15.

12. Appendices

Appendix 1.

Manufacturers of all primary total hip replacement and primary resurfacing head and cup manufacturers

Manufacturer Primary total hip replacement
Amplitude
Biomet
B Braun/ Aesculap
Corin
DePuy
Exactech
JRI (Joint Replacement Instrumentation)
Implantcast
Implants International
Lima WG Healthcare
Mathys Orthopaedics
Medacta UK
Othodynamics
Peter Brehm
SERF dedienne santé
Smith & Nephew
Stanmore Implants Worldwide
Stryker
Symbios SA
Waldemar Link
Wright Medical UK
Zimmer
Manufacturer primary resurfacing head and cup manufacturers
Biomet
Corin
Implantcast
Smith & Nephew
Wright Medical UK

Zimmer

Table 1 Names and manufacturers for all Primary total hip replacement and primary resurfacing head and cup manufacturers

Appendix 2.

Clinical effectiveness searches

Medline (1946 to October Week 4 2012) via OVID interface, searched on 05/11/2012

1	exp Arthroplasty, Replacement, Hip/	15246
2	exp Hip Prosthesis/	18304
3	(tha or thr).tw.	23312
4	exp Hip Joint/	20108
5	exp Hip/	8480
6	hip.tw.	79606
7	("femur head*" or "femoral head*" or acetabul*).tw.	20571
8	exp Femur Head/	7700
9	exp Acetabulum/	8243
10	4 or 5 or 6 or 7 or 8 or 9	97057
11	(arthroplast* or replace* or implant* or prosthes*).tw.	514865
12	exp Joint Prosthesis/	33736
13	exp "Prostheses and Implants"/	355910
14	11 or 12 or 13	716289
15	10 and 14	35876
16	(surf* or resurf*).tw.	629176
17	10 and 16	5573
18	1 or 2 or 3 or 15 or 17	61490
19	exp Arthritis, Rheumatoid/ or exp Arthritis/	190095
20	exp Osteoarthritis, Hip/ or exp Osteoarthritis/	39813
21	(arthrit* or osteoarthrit* or osteoarthrosis or "rheumatoid arthrit*").tw.	141102
22	19 or 20 or 21	221909
23	18 and 22	7739

24	meta analysis.pt.	37222
25	randomized controlled trial.pt.	340101
26	(random* or "controlled trial*" or "clinical trial*" or rct).tw.	718263
27	(metaanalys* or "meta analys*" or "meta-analys*").tw.	42924
28	"systematic review*".tw.	34474
29	24 or 25 or 26 or 27 or 28	846326
30	23 and 29	614
31	limit 30 to (english language and yr="2002 -Current")	443

Cost effectiveness searches

Medline (1946 to November Week 2 2012) via OVID interface, searched on 21/11/2012

	Searches	Results
1	exp Arthroplasty, Replacement, Hip/	15452
2	exp Hip Prosthesis/	18500
3	(tha or thr).tw.	23434
4	exp Hip Joint/	20449
5	exp Hip/	8617
6	hip.tw.	80678
7	("femur head*" or "femoral head*" or acetabul*).tw.	20855
8	exp Femur Head/	7859
9	exp Acetabulum/	8395
10	4 or 5 or 6 or 7 or 8 or 9	98321
11	(arthroplast* or replace* or implant* or prosthes*).tw.	517989
12	exp Joint Prosthesis/	34030
13	exp "Prostheses and Implants"/	360271
14	11 or 12 or 13	722394
15	10 and 14	36321
16	(surf* or resurf*).tw.	631946
17	10 and 16	5613
18	1 or 2 or 3 or 15 or 17	62033
19	exp Arthritis, Rheumatoid/ or exp Arthritis/	190844
20	exp Osteoarthritis, Hip/ or exp Osteoarthritis/	40125
21	(arthrit* or osteoarthrit* or osteoarthrosis or "rheumatoid arthrit*").tw.	141771
22	19 or 20 or 21	222856
23	18 and 22	7855
24	*Economics/ or exp *"economics, hospital"/ or *economics, medical/ or *economics, nursing/	27335
25	exp *"Costs and Cost Analysis"/	42087
26	exp *"Cost of Illness"/	6771

27	exp *"Models, Economic"/	3077
28	(cost* or economic*).ti.	96067
29	exp *"Quality of Life"/	46201
30	exp *"Quality-Adjusted Life Years"/	1296
31	(ICER or qaly* or eq5d* or "eq-5d*" or euroqol or "euro-qol" or "quality of well-being" or "quality of wellbeing" or "short-form 36" or "shortform 36" or "36-item short-form" or "36-item short form" or "sf-36" or sf36 or "short-form 12" or "short form 12" or "12-item short-form" or "12-item short form" or "sf12" or "sf-12").ti.	1823
32	("Stanford Health Assessment Questionnaire" or HAQ or "Western Ontario and McMaster University Osteoarthritis Index" or WOMAC or OAKHQOL or JAQQ or PSAQoL).tw.	3220
33	(markov or "time trade off" or "time-trade-off" or standard gamble or utilit* or qol or hrql or hrqol or disutilit* or "net-benefit analysis").ti.	17993
34	(quality adj2 life).ti.	32938
35	(decision adj2 model).ti.	454
36	("resource use" or "resource utili?ation").ti.	1505
37	exp *Health Status/	45793
38	("health state*" or "health status").ti.	7435
39	24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38	247280
40	23 and 39	500
41	limit 40 to (english language and yr="2002 -Current")	348

Registry searches. Medline (1946 to November Week 2 2012) via OVID interface, searched on 21/11/2012

	Searches	Results
1	exp Arthroplasty, Replacement, Hip/	15452
2	exp Hip Prosthesis/	18500
3	(tha or thr).tw.	23434
4	exp Hip Joint/	20449
5	exp Hip/	8617
6	hip.tw.	80678
7	("femur head*" or "femoral head*" or acetabul*).tw.	20855
8	exp Femur Head/	7859
9	exp Acetabulum/	8395
10	4 or 5 or 6 or 7 or 8 or 9	98321
11	(arthroplast* or replace* or implant* or prosthes*).tw.	517989
12	exp Joint Prosthesis/	34030
13	exp "Prostheses and Implants"/	360271
14	11 or 12 or 13	722394
15	10 and 14	36321
16	(surf* or resurf*).tw.	631946
17	10 and 16	5613
18	1 or 2 or 3 or 15 or 17	62033
19	exp Arthritis, Rheumatoid/ or exp Arthritis/	190844
20	exp Osteoarthritis, Hip/ or exp Osteoarthritis/	40125
21	(arthrit* or osteoarthrit* or osteoarthrosis or "rheumatoid arthrit*").tw.	141771
22	19 or 20 or 21	222856
23	18 and 22	7855
24	exp Registries/	50193
25	(registry or registries).tw.	48804
26	(register or registers).tw.	34468
27	Databases as Topic/	7949
28	Databases, Factual/	37575
29	24 or 25 or 26 or 27 or 28	145461
30	23 and 29	244
31	limit 30 to (english language and yr="2002 -Current")	208

Appendix 3. Data extraction forms

Data extraction form for primary studies

Name of the reviewer:

Study details
Study ID (Ref man):
First author surname:
Year of publication:
Country:
Study design:
Study setting:
Number of centres:
Duration of study:
Follow up period:
Funding:
Aim of the study:
Participants
Total number of participants:
Sample attrition/drop out:
Inclusion criteria:
Exclusion criteria:
Characteristics of participants:
Mean age:
Mean sex:
Race:
Diagnosis:
Intervention
Indication for treatment:
Type of device used:
Any comparison:
Duration of treatment:
Other interventions used:

CE approval: Yes/No;
Outcomes
Primary outcomes:
Secondary outcomes:
Method of assessing outcomes:
Timing of assessment:
Study end point:
Survival analysis: Yes/No
Mortality: Yes/No
Physiological data: Yes/No
Adverse event: Yes/No
Health related quality of life: Yes/No; which measures used?
Length of follow up:

Number of participants	**Intervention**	**Comparator, if present**
Screened		
Randomised/Included		
Excluded		
Missing participants		
Withdrawals		
Patient's baseline characteristics	**Intervention**	**Comparator, if present**
Age, years		
Sex		
BSA, m^2		
Weight, kg, BMI		
Survival data	**Intervention**	**Comparator, if present**
Actuarial survival		
Overall survival		
Kaplan-Meier estimates		
Physiological data	**Intervention**	**Comparator, if present**
Adverse events	**Intervention**	**Comparator, if present**

Cause of death		
≤12 months		
≥12 months		
Quality of life	**Intervention**	**Comparator, if present**

Authors conclusion
Reviewer's conclusion

Data extraction form for economic studies.

Name of the reviewer:

Study intervention (clearly defined?)
Objective (clearly defined?)
Design
Analytical framework (type of model):
Patient population:
Comparator (clearly defined?)
Analytic horizon:
Perspective:
Setting:
Clinical measures:
Effectiveness measures:
Economic measures:
Methods
Health care system:
Model description:
Data sources (efficacy, resource use, costs, appropriately measured, all costs included?:
Data collection (primary data collection, if appropriate):

Probabilities:
Healthcare use:
Sensitivity analysis (allowance made for uncertainty):
Discounting (costs/benefits?):
Results (incremental analysis of costs and consequences?)
Conclusion: Assessment:

Authors conclusion
Reviewer's conclusion

Data extraction form for systematic reviews

Name of the reviewer:

Study details
Study ID (Ref man):
First author surname:
Year of publication:
Country:
Funding:
Aim of the study:
Methods
Databases searched:
Last date of search:
Inclusion criteria:
Participants:
Interventions:
Comparators:
Outcome measures:
Types of studies included:
Quality assessment criteria used:
Application of methods:
Methods of analysis:
1. narrative, 2. meta-analysis, 3. indirect comparison, 4. others
Results
Quantity and quality of included studies:
Treatment effect:
Economic evaluation:
Conclusions:
Implications of the review:
Methodological comments
Search strategy:
Participants:
Inclusion/exclusion criteria:

Quality assessment of studies: Method of synthesis:
General comment
Generalisability: Funding:
Authors conclusion
Reviewer's conclusion

Appendix 4 Methodological quality and risk of bias assessment instruments

Quality assessment forms

Cochrane Collaboration's tool for assessing risk of bias for a randomized controlled trial

(adapted from Higgins et al. 2011) (BMJ 2011;343:d5928 doi: 10.1136/bmj.d5928)

Bias domain	Source of bias	Support for judgment*	Authors' judgment**
Selection bias	Random sequence generation		
	Allocation concealment		
Performance bias	Blinding of participants and Personnel [for each outcome or group of subjective/objective outcomes of interest]		
Detection bias	Blinding of outcome assessors [for each outcome or group of subjective/objective outcomes of interest]		
Attrition bias	Incomplete outcome data [for each outcome of interest]		
Reporting bias	Selective reporting of the outcome [for each outcome of interest], subgroups, or analysis		
Other bias	Funding source, adequacy of statistical methods used, type of analysis [ITT/PP] baseline imbalance in important characteristics		

* Statement, description or quote supporting the judgment
** Low risk of bias, high risk of bias, or unclear risk of bias

ITT=intention to treat
PP=per protocol

Summary risk of bias assessment for each outcome within and across randomized controlled trials
(adapted from Higgins et al. 2011) (BMJ 2011;343:d5928 doi: 10.1136/bmj.d5928)

Study ID	Random sequence generation	Allocation concealment	Blinding of participants and Personnel	Blinding of outcome assessors	Incomplete outcome data	Selective reporting	Other bias [ζ]	Summary risk of bias (within trial)[*]
Summary risk of bias (across trials)[**]								

[ζ] Funding source, adequacy of statistical methods used, type of analysis (ITT/PP), baseline imbalance in important characteristics

[*] Low risk of bias (low risk of bias for all key domains), high risk of bias (high risk of bias for one or more key domains), or unclear risk of bias (low or unclear risk of bias for all key domains)

[**] Low risk of bias (most information is from trials at low risk of bias), high risk of bias (the proportion of information from trials at high risk of bias is sufficient to affect the interpretation of results), or unclear risk of bias (most information is from trials at low or unclear risk of bias)

ITT=intention to treat
PP=per protocol

Quality assessment criteria for systematic reviews: The AMSTAR tool for assessing methodological quality of systematic reviews

(Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS One* 2007; 2(12):e1350,
Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007; 7:10.
Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009; 62(10):1013-1020)

1. Was an 'a priori' design provided?
2. Was there duplicate study selection and data extraction?
3. Was a comprehensive literature search performed?
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
5. Was a list of studies (included and excluded) provided?
6. Were the characteristics of the included studies provided?
7. Was the scientific quality of the included studies assessed and documented?
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
9. Were the methods used to combine the findings of studies appropriate?
10. Was the likelihood of publication bias assessed?
11. Was the conflict of interest stated?

Rating (by criteria fulfilled, i.e. 'yes' response): 9 to 11 high quality, 5 to 8 medium quality, 0 to 4 low quality.

Quality assessment criteria for economic studies: Drummond checklist (Drummond, 1996 Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ;313:275)

Item		Yes	No	Not clear	Not appropriate
Study design					
1.	The research question is stated.	¨	¨	¨	
2.	The economic importance of the research question is stated.	¨	¨	¨	
3.	The viewpoint(s) of the analysis are clearly stated and justified.	¨	¨	¨	
4.	The rationale for choosing alternative programmes or interventions compared is stated.	¨	¨	¨	
5.	The alternatives being compared are clearly described.	¨	¨	¨	
6.	The form of economic evaluation used is stated.	¨	¨	¨	
7.	The choice of form of economic evaluation is justified in relation to the questions addressed.	¨	¨	¨	
Data collection					
8.	The source(s) of effectiveness estimates used are stated.	¨	¨	¨	
9.	Details of the design and results of effectiveness study are given (if based on a single study).	¨	¨	¨	¨
10.	Details of the methods of synthesis or meta-analysis of estimates are given (if based on a synthesis of a number of effectiveness studies).	¨	¨	¨	¨
11.	The primary outcome measure(s) for the economic evaluation are clearly stated.	¨	¨	¨	
12.	Methods to value benefits are stated.	¨	¨	¨	¨
13.	Details of the subjects from whom valuations were obtained were given.	¨	¨	¨	¨
14.	Productivity changes (if included) are reported separately.	¨	¨	¨	¨
15.	The relevance of productivity changes to the study question is discussed.	¨	¨	¨	¨
16.	Quantities of resource use are reported separately from their unit costs.	¨	¨	¨	
17.	Methods for the estimation of quantities and unit costs are	¨	¨	¨	

	described.				
18.	Currency and price data are recorded.	¨	¨	¨	
19.	Details of currency of price adjustments for inflation or currency conversion are given.	¨	¨	¨	
20.	Details of any model used are given.	¨	¨	¨	¨
21.	The choice of model used and the key parameters on which it is based are justified.	¨	¨	¨	¨
Analysis and interpretation of results					
22.	Time horizon of costs and benefits is stated.	¨	¨	¨	¨
23.	The discount rate(s) is stated.	¨	¨	¨	¨
24.	The choice of discount rate(s) is justified.	¨	¨	¨	¨
25.	An explanation is given if costs and benefits are not discounted.	¨	¨	¨	¨
26.	Details of statistical tests and confidence intervals are given for stochastic data.	¨	¨	¨	¨
27.	The approach to sensitivity analysis is given.	¨	¨	¨	¨
28.	The choice of variables for sensitivity analysis is justified.	¨	¨	¨	¨
29.	The ranges over which the variables are varied are justified.	¨	¨	¨	¨
30.	Relevant alternatives are compared.	¨	¨	¨	¨
31.	Incremental analysis is reported.	¨	¨	¨	¨
32.	Major outcomes are presented in a disaggregated as well as aggregated form.	¨	¨	¨	
33.	The answer to the study question is given.	¨	¨	¨	
34.	Conclusions follow from the data reported.	¨	¨	¨	
35.	Conclusions are accompanied by the appropriate caveats.	¨	¨	¨	

Appendix 5. Explanations of grading process across the five domains

Definitions of grades of evidence (adapted from Oxman BMJ 2004; 328:1490-4)

Overall grade of evidence	Interpretation
High	Further research is unlikely to change our confidence in the effect estimate
Moderate	Further research is likely to have an important impact on our confidence in the effect estimate and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the effect estimate and is likely to change the estimate
Very low	Any estimate of effect is very uncertain

Grading quality of evidence across the five domains

Domain of assessment	Assessment target	Examples of downgrading the overall rating
Overall risk of bias	Internal validity across studies	Overall risk of bias is moderate (-1) or high (-2)
Consistency	Heterogeneity	Inconsistent direction of estimates that cannot be explained (-1)
Directness	Applicability of the results; indirect treatment comparisons	The study populations are highly selected or only a subgroup effect estimate is known limiting applicability (-2); only a surrogate (i.e., laboratory measure) of a clinical important outcome was reported (-1); absence of evidence on head-to-head comparison of treatments (-1)
Precision	The width of 95% CI around the estimate	Non-significant effect with wide CIs equally compatible with benefits and harms (-1)
Publication or outcome reporting bias	Assessing the funnel plot for asymmetry; all or most studies included in the review are funded by industry; selective reporting of outcomes	Visual inspection and regression-based tests reveal asymmetry in the funnel plot (-1); graded outcome not reported in one or more studies (-1)

CI=confidence interval

GRADE evidence profile (adapted from Guyatt J Clin Epidemiol 2011; 64: 383-94)

Type of outcome [follow-up timing]	# Of studies [design]	Overall risk of bias	Consistency	Directness	Precision	Publication or outcome reporting bias	# Patients with outcome n/N		Effect estimate [95% CI]	Quality (GRADE)
							Treatment arm	Comparator arm		
Outcome 1										
Outcome 2										
Outcome 3										
Outcome 4										
Outcome 5										
GRADE= Grading of Recommendations, Assessment, Development, and Evaluation; RCT=randomized controlled trial; CI=confidence interval										

Summary of findings (adapted from Guyatt J Clin Epidemiol 2011; 64: 383-94)

Type of outcome [follow-up timing]	Estimated risks [95% CI]		Effect estimate [95% CI]	# Patients [# studies]	Quality (GRADE)	Comments
	Control risk*	Intervention risk**				
	Comparator	Intervention				
Outcome 1						
Outcome 2						
Outcome 3						
Outcome 4						
Outcome 5						
GRADE= Grading of Recommendations, Assessment, Development, and Evaluation; CI=confidence interval						

* The median control group risk across studies ** Based on the control risk in the comparison group and the relative effect of the intervention (and its 95% CI)