

# 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): systematic review and economic evaluation

Aileen Clarke,<sup>1\*</sup> Ruth Pulikottil-Jacob,<sup>1</sup>  
Amy Grove,<sup>1</sup> Karoline Freeman,<sup>1</sup> Hema Mistry,<sup>1</sup>  
Alexander Tsertsvadze,<sup>1</sup> Martin Connock,<sup>1</sup>  
Rachel Court,<sup>1</sup> Ngianga-Bakwin Kandala,<sup>1</sup>  
Matthew Costa,<sup>1</sup> Gaurav Suri,<sup>1</sup> David Metcalfe,<sup>2</sup>  
Michael Crowther,<sup>3</sup> Sarah Morrow,<sup>4</sup>  
Samantha Johnson<sup>1</sup> and Paul Sutcliffe<sup>1</sup>

<sup>1</sup>Warwick Evidence, Warwick Medical School, University of Warwick, Coventry, UK

<sup>2</sup>Warwick Orthopaedics, University Hospitals Coventry and Warwickshire, Coventry, UK

<sup>3</sup>Department of Health Sciences, University of Leicester, Leicester, UK

<sup>4</sup>Oxford Medical School, University of Oxford, Oxford, UK

\*Corresponding author

**Declared competing interests of authors:** Aileen Clarke is a member of the NIHR HTA Editorial Board and the Warwick Medical School receive payment for this work.

Published January 2015

DOI: 10.3310/hta19100

## Scientific summary

### Total hip replacement and surface replacement

Health Technology Assessment 2015; Vol. 19: No. 10

DOI: 10.3310/hta19100

NIHR Journals Library [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

# Scientific summary

## Background

Osteoarthritis (OA) is a leading cause of pain and disability in the UK. The incidence rates of OA of the hip in men and women aged 70–79 years are estimated to be 430 and 600 per 100,000 person-years, respectively. The economic impact of arthritis is vast, because of both direct costs to the health-care system, community and social services and indirect costs to patients, carers and families. Patients who do not respond to non-surgical measures are referred for elective surgical interventions, most commonly total hip replacement (THR) or resurfacing arthroplasty (RS).

Total hip replacement involves replacement of the hip joint with an artificial hip prosthesis consisting of a cup (with or without liner), a femoral stem and head. There are different types of THR including different types of articulation surface (metal, ceramic, polyethylene, ceramicised metal); implant component fixation (cemented, cementless, hybrid, reverse hybrid); and implant component size. RS involves replacement of the femoral head surface with a metal covering. The resurfacing component articulates with a hollow metal cup located in the acetabulum. Revision is undertaken when implants fail because of infection or loosening. Previous National Institute for Health and Care Excellence guidance indicated that the benchmark for selection of prostheses for THR should be a revision rate of  $\leq 10\%$  at 10 years.

## Decision problem and objectives

The main objective was to undertake a clinical effectiveness and cost-effectiveness analysis of different types of THR and RS for the treatment of pain and disability in people with end-stage arthritis of the hip.

Specific aims were to compare the clinical effectiveness and cost-effectiveness of different types of primary THR:

- (a) with RS for people in whom both procedures are suitable
- (b) with each other for those not suitable for hip RS.

## Systematic reviews

### Methods

Searches were undertaken of clinical effectiveness, cost-effectiveness and registry studies in December 2012. For clinical effectiveness, studies were limited to 2008 onwards and a sample size of  $\geq 100$  participants. Electronic databases searched included MEDLINE, EMBASE, The Cochrane Library and Current Controlled Trials. Reference lists, and manufacturer and professional organisation websites were screened. Full-text English-language reports of randomised controlled trials (RCTs), systematic reviews and meta-analyses were included.

Two independent reviewers screened all records and extracted data, with disagreements resolved through consensus. Methodological quality was assessed using the Cochrane Collaboration's risk of bias and the Assessment of Multiple Systematic Reviews (AMSTAR) tools. Estimates of post-treatment mean difference (MD) for continuous outcomes and risk ratios (RRs) for binary outcomes of individual studies were pooled using a random-effects model. Dichotomous outcomes were pooled as RRs using a Mantel–Haenszel fixed-effects model or as odds ratios (ORs) using the Peto fixed-effects model. Heterogeneity was

determined through Cochran's  $Q$  and the  $I^2$  statistics. Overall quality of evidence was assessed using GRADE (Grading of Recommendations, Assessment, Development and Evaluation).

This report contains reference to confidential information provided as part of the NICE appraisal process. This information has been removed from the report and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

## Systematic review results

### Clinical effectiveness: resurfacing arthroplasty compared with total hip replacement

A total of 2469 records were screened, of which 37 were included, representing 16 RCTs and eight systematic reviews. Mean age ranged from 45 to 72 years with a maximum follow-up of 20 years. Mean post-THR Harris Hip Score (HHS) and Western Ontario and McMaster University Osteoarthritis Index (WOMAC) and Short Form questionnaire-12 items (SF-12) scores measured at different follow-up periods did not differ between THR groups, including between cross-linked polyethylene and traditional polyethylene cup liners [HHS pooled MD 2.29, 95% confidence interval (CI) -0.88 to 5.45].

There was a reduced risk of implant dislocation with the use of a cemented cup compared with a cementless cup (high-grade evidence; pooled OR 0.34, 95% CI 0.13 to 0.89) or larger femoral head size (36 mm vs. 28 mm). Femoral head penetration rates were reduced for cross-linked compared with conventional polyethylene cup liners (low-grade evidence). Recipients of ceramic-on-ceramic articulations (vs. metal-on-polyethylene) experienced a reduced risk of osteolysis. Evidence from two RCTs indicated an increased infection risk for THR compared with RS (pooled OR 7.94, 95% CI 1.78 to 35.40).

The eight systematic reviews identified (five on different types of THR and three on THR vs. RS) were inconclusive.

## Cost-effectiveness

### Methods

We drew on systematic review results and the National Joint Registry for England and Wales (NJR). Using a series of cross-tabulations, we identified the top four most commonly used mutually exclusive categories of THR (> 25,000) as reported in the NJR and on clinical advice included a further mutually exclusive fifth category. We built a Markov multistate model to investigate both RS and THR using the observed time to revision. Cycle length was 1 year and we adopted 10-year and lifetime horizons.

Analysis was conducted from the perspectives of the NHS and personal and social services. All costs are in UK pounds at 2011/12 prices. Health outcomes were measured in quality-adjusted life-years (QALYs). Results are expressed as incremental cost per QALY gained. An annual discount rate of 3.5% was applied to both costs and outcomes. We ran the model deterministically and probabilistically with 1000 iterations, calculated cost-effectiveness acceptability curves (CEACs) and undertook sensitivity analyses.

### Resurfacing arthroplasty compared with total hip replacement

We propensity matched RS patients drawing age–sex matched pairs from the data set of all identified categories of THR combined, to identify patients who underwent THR but who were also eligible for RS. We used NHS reference costs for RS and THR for follow-up and revision. We drew age- and sex-adjusted utility values from the patient-reported outcome measures (PROMs) data set for both THR and RS.

For the comparison of RS with THR we undertook sensitivity analyses stratified by sex and controlled for age. We assessed estimates of cost-effectiveness for men and women aged 40, 50 and 60 years using lifetime revision rates. We constructed CEACs comparing RS with THR overall, in separate age groups and at different levels of willingness to pay (WTP).

## Total hip replacement compared with total hip replacement

We compared the five categories of THR with each other and in sensitivity analyses investigated patients aged > 65 years who are less eligible for RS. In the base case we used costs supplied by the manufacturers for each of the components of THR; we used alternative costs in sensitivity analyses. We used age- and sex-adjusted utility values from the PROMs data set for before and after hip replacement and for revision.

We undertook various sensitivity analyses and analysis of cost drivers. These included different age and sex categories, stratifying by age (> 65 years or < 65 years), different methods of extrapolation of revision rates, varying prosthesis costs (using NHS list prices) and discount rates. We constructed CEACs comparing different types of THR overall and in separate age groups at different levels of WTP.

### Cost-effectiveness results

#### Resurfacing arthroplasty compared with total hip replacement

Using the NJR we found a total of 31,222 people who had undergone RS and 386,556 who had undergone a THR. In total, 3% of those undergoing THR and 11% undergoing RS had a revision by 9 years. The revision rate for all RS was always higher than that for THR. The prostheses cost £2672 and £2571 for RS and THR, respectively.

For all analyses, the mean cost for RS was higher than that for THR and the mean QALYs were lower. The incremental cost-effectiveness ratio for RS was dominated by THR; that is, THR was cheaper and more effective than RS (for a lifetime horizon in the base-case analysis, the incremental cost of RS was £11,490 and the incremental QALYs were -0.0879).

Similar results were obtained for the deterministic and probabilistic analysis of RS compared with THR and when analysed separately in sensitivity analyses for men and women by age group (40, 50 and 60 years). For all age and sex groups RS remained clearly dominated by THR. CEACs showed that, for all patients, THR was almost 100% cost-effective at any WTP level.

#### Total hip replacement compared with total hip replacement

We identified five categories of commonly used types of THR from the NJR:

- category A: metal head (cemented stem) on cemented polyethylene cup (CeMoP) (125,285 patients)
- category B: metal head (cementless stem) on cementless hydroxyapatite-coated metal cup (polyethylene liner) (CeLMoP) (37,874 patients)
- category C: ceramic head (cementless stem) on cementless hydroxyapatite-coated metal cup (ceramic liner) (CeLCoC) (34,754 patients)
- category D: hybrid metal head (cemented stem) on cementless hydroxyapatite-coated metal cup (polyethylene liner) (HyMoP) (28,471 patients)
- category E: ceramic head (cemented stem) on cemented polyethylene cup (CeCoP) (12,075 patients).

There were age and sex differences between recipients of different types of THR and variations in revision rates (category A: 2.5%; B: 3.2%; C: 3.5%; D: 2.5%; E: 1.6% at 9 years). For all interventions, revision rates at 9 years were substantially less than the 10% benchmark. Costs of the different prostheses were as follows: category A: £1557.38; B: £3015.60; C: £3868.80; D: £2649.78; and E: £1995.98.

In the base-case analysis, for all age and sex groups combined and using a bathtub model (indicating an increasing likelihood of need for revision with time) and a lifetime horizon, the mean cost for category E (CeCoP) was slightly lower and the mean QALYs for category E were slightly higher than for all other THR categories in both deterministic and probabilistic analyses. Category E dominated the other four categories.

For example, in the deterministic analysis, compared with category E, category A (CeMoP) cost £278 more (£14,801 vs. £14,523) and generated 0.0022 fewer QALYs (14.7887 vs. 14.7909). The probabilistic results were very similar. The CEACs demonstrated that, over a lifetime horizon, category E was 97.2% likely to be cost-effective compared with 2.8% for category A at a WTP of £20,000 per QALY. For patients aged > 65 years, category A was more likely to be cost-effective in all groups (category A: 100% probability of being cost-effective; categories B–E: < 1% probability of being cost-effective). When examining the lifetime horizon for all age groups, category E was more clinically effective (except for men aged 80 years for whom QALYs generated by categories A and E were the same).

Sensitivity analyses using a log-normal model (indicating a decreasing risk of revision over time) for extrapolation beyond the observed data for revision rates resulted in category A being cheaper over a lifetime horizon for all age–sex groups combined. Although category E was more clinically effective than the other four categories, category A was 100% cost-effective at a WTP threshold of £20,000 per QALY. Additional sensitivity analysis using an age- and sex-adjusted log-normal model also demonstrated that over a lifetime horizon category A was 100% cost-effective at a WTP of £20,000 per QALY.

Varying the main inputs by 30% in the base-case analysis for all age–sex groups, and comparing category A with category E, demonstrated that the main drivers of difference were the costs of the components, the discount rate and the modelled revision rates.

## Strengths and limitations

We reanalysed comprehensive national audit data to calculate outcomes and used the PROMs data set as a source for utility data coupled with costs from the rigorous literature review, NHS reference costs and manufacturers' costs. We did not find any relevant RCTs comparing RS and THR or different types of THR to allow us to model differences in revision rates relevant to a lifetime horizon. As NJR data are non-randomised and may be subject to selection bias, we worked to reduce confounding by propensity matching and by undertaking extensive analyses by age and sex.

In comparing RS with THR, our clinical advisors suggested that the selection of patients for RS may be based on activity levels (levels of physical fitness, athleticism); however, the only characteristics reliably collected at patient level in the NJR were age and sex. This means that we were unable to identify other characteristics or subpopulations for whom RS might be more beneficial. However, age and sex are likely to act as a proxy for physicality and it is of interest that revision rates for RS were higher in every age and sex group that we examined, including in the youngest category of men.

We identified five categories of the most commonly used combinations of THR components. To our knowledge this is the first time that different types of THR have been investigated in this comparative way. It has the advantage of more precisely reflecting current practice.

Revision rates are one of the main factors affecting the cost-effectiveness of the different categories. We had preselected category E on the recommendation of our clinical advisors before assessing revision rates. We undertook extensive modelling of revision rates to find the best methods for extrapolation beyond the observed data for all THR categories. We found that category E had lower revision rates overall and generally across age–sex groups. This pertained across different methods for extrapolation, suggesting that the relative cost-effectiveness of category E is a robust finding.

## Conclusions

### *Systematic reviews*

Total hip replacement is a common operation and is clearly beneficial. Improvements post surgery were reported in the literature for functional/clinical and quality-of-life measures regardless of the type of surgery. Much of the evidence was inconclusive because of poor reporting, missing data, inconsistent results and uncertainty in treatment effect estimates. Evidence on the relative benefits of RS compared with THR or of different types of THR was largely lacking. Certain types of THR appeared to confer some benefit including larger femoral head sizes, use of a cemented cup, use of a cross-linked polyethylene cup liner and a ceramic-on-ceramic articulation as opposed to a metal-on-polyethylene articulation.

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

Compared with THR, revision rates for RS were higher, mean costs were higher and mean QALYs were lower; RS was therefore dominated by THR. Very similar results were obtained for deterministic and probabilistic analyses and for all age and sex groups. THR was almost 100% cost-effective at any WTP level.

### *Total hip replacement compared with total hip replacement*

Revision rates for all types of THR were low. Costs of prostheses varied depending on composition. There were small but clear differences between categories in both costs and effectiveness as measured by QALYs, and when age and sex groups were factored in. Category A was more cost-effective for older age groups for whom revision rates are lower. However, across all age–sex groups combined, mean costs were slightly lower and mean QALYs were slightly higher for category E, than for all other THR categories in both deterministic and probabilistic analyses; category E therefore dominated the other four categories.

## Recommendations for research

1. Consideration should be given to setting up RCTs with long-term follow-up.
2. We were not able to link PROMs data with NJR data or with costs in our analysis; however, the NJR will embed these utility data from 2013.
3. We would welcome work to validate our new findings on the relative cost-effectiveness of different combinations of prosthesis components for THR.

## Study registration

This study is registered as PROSPERO CRD42013003924.

## Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

*Health Technology Assessment* is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [nhredit@southampton.ac.uk](mailto:nhredit@southampton.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hta](http://www.journalslibrary.nihr.ac.uk/hta). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

## Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

## This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 11/118/01. The protocol was agreed in November 2012. The assessment report began editorial review in July 2013 and was accepted for publication in January 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

**© Queen's Printer and Controller of HMSO 2015. This work was produced by Clarke *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.**

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## **Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library**

**Professor Tom Walley** Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

### **NIHR Journals Library Editors**

**Professor Ken Stein** Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

**Professor Andree Le May** Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

**Dr Martin Ashton-Key** Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

**Professor Matthias Beck** Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

**Professor Aileen Clarke** Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Peter Davidson** Director of NETSCC, HTA, UK

**Ms Tara Lamont** Scientific Advisor, NETSCC, UK

**Professor Elaine McColl** Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Professor of Health Sciences Research, Faculty of Education, University of Winchester, UK

**Professor John Powell** Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professor of Child Health Research, UCL Institute of Child Health, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:  
[www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [nihredit@southampton.ac.uk](mailto:nihredit@southampton.ac.uk)