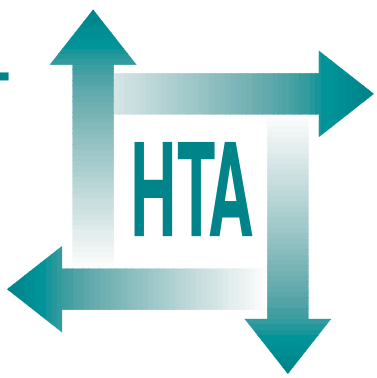
Health Technology Assessment 1998; Vol. 2: No. 20

Review

Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses

R Fitzpatrick
E Shortall
M Sculpher
D Murray
R Morris
M Lodge
J Dawson
A Carr
A Britton
A Briggs

Health Technology Assessment
NHS R&D HTA Programme



Standing Group on Health Technology

Chair: Professor Sir Miles Irving,
Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Dr Sheila Adam,
Department of Health

Professor Martin Buxton,
Professor of Economics, Brunel University †

Professor Angela Coulter,
Director, King's Fund, London

Professor Anthony Culyer,
Deputy Vice-Chancellor, University of York

Dr Peter Doyle,
Executive Director, Zeneca Ltd,
ACOST Committee on Medical Research & Health

Professor John Farndon,
Professor of Surgery, University of Bristol †

Professor Charles Florey,
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee †

Professor John Gabbay,
Director, Wessex Institute for Health Research & Development †

Professor Howard Glennester,
Professor of Social Science & Administration, London School of Economics & Political Science

Professor Sir John Grimley Evans,
Department of Geriatric Medicine, Radcliffe Infirmary, Oxford †

Dr Tony Hope,
The Medical School, University of Oxford †

Mr John H James,
Chief Executive, Kensington, Chelsea & Westminster Health Authority

Professor Richard Lilford,
Regional Director, R&D, West Midlands †

Professor Michael Maisey,
Professor of Radiological Sciences, UMDS, London

Dr Jeremy Metters,
Deputy Chief Medical Officer, Department of Health †

Mrs Gloria Oates,
Chief Executive, Oldham NHS Trust

Dr George Poste,
Chief Science & Technology Officer, SmithKline Beecham †

Professor Michael Rawlins,
Wolfson Unit of Clinical Pharmacology, University of Newcastle-upon-Tyne

Professor Martin Roland,
Professor of General Practice, University of Manchester

Mr Hugh Ross,
Chief Executive, The United Bristol Healthcare NHS Trust †

Professor Ian Russell,
Department of Health Sciences & Clinical Evaluation, University of York

Professor Trevor Sheldon,
Director, NHS Centre for Reviews & Dissemination, University of York †

Professor Mike Smith,
Director, The Research School of Medicine, University of Leeds †

Dr Charles Swan,
Consultant Gastroenterologist, North Staffordshire Royal Infirmary

Dr John Tripp,
Department of Child Health, Royal Devon & Exeter Healthcare NHS Trust †

Professor Tom Walley,
Department of Pharmacological Therapeutics, University of Liverpool †

Dr Julie Woodin,
Chief Executive, Nottingham Health Authority †

† Current members

HTA Commissioning Board

Chair: Professor Charles Florey, Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee †

Professor Ian Russell,
Department of Health Sciences & Clinical Evaluation, University of York *

Dr Doug Altman,
Director of ICRF/NHS Centre for Statistics in Medicine, Oxford †

Mr Peter Bower,
Independent Health Advisor, Newcastle-upon-Tyne †

Ms Christine Clark,
Honorary Research Pharmacist, Hope Hospital, Salford †

Professor David Cohen,
Professor of Health Economics, University of Glamorgan

Mr Barrie Dowdeswell,
Chief Executive, Royal Victoria Infirmary, Newcastle-upon-Tyne

Professor Martin Eccles,
Professor of Clinical Effectiveness, University of Newcastle-upon-Tyne †

Dr Mike Gill,
Director of Public Health and Health Policy, Brent & Harrow Health Authority †

Dr Jenny Hewison,
Senior Lecturer, Department of Psychology, University of Leeds †

Dr Michael Horlington,
Head of Corporate Licensing, Smith & Nephew Group Research Centre

Professor Sir Miles Irving
(Programme Director), Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Professor Alison Kitson,
Director, Royal College of Nursing Institute †

Professor Martin Knapp,
Director, Personal Social Services Research Unit, London School of Economics & Political Science

Dr Donna Lamping,
Senior Lecturer, Department of Public Health, London School of Hygiene & Tropical Medicine †

Professor Theresa Marteau,
Director, Psychology & Genetics Research Group, UMDS, London

Professor Alan Maynard,
Professor of Economics, University of York †

Professor Sally McIntyre,
MRC Medical Sociology Unit, Glasgow

Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield †

Professor Gillian Parker,
Nuffield Professor of Community Care, University of Leicester †

Dr Tim Peters,
Reader in Medical Statistics, Department of Social Medicine, University of Bristol †

Professor David Sackett,
Centre for Evidence Based Medicine, Oxford

Professor Martin Severs,
Professor in Elderly Health Care, Portsmouth University †

Dr David Spiegelhalter,
MRC Biostatistics Unit, Institute of Public Health, Cambridge

Dr Ala Szczepura,
Director, Centre for Health Services Studies, University of Warwick †

Professor Graham Watt,
Department of General Practice, Woodside Health Centre, Glasgow †

Professor David Williams,
Department of Clinical Engineering, University of Liverpool

Dr Mark Williams,
Public Health Physician, Bristol

Dr Jeremy Wyatt,
Senior Fellow, Health and Public Policy, School of Public Policy, University College, London †

* Previous Chair
† Current members

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses

R Fitzpatrick[1]
E Shortall[1]
M Sculpher[2]
D Murray[3]
R Morris[4]
M Lodge[1]
J Dawson[1]
A Carr[3]
A Britton[5]
A Briggs[1]

[1] Institute of Health Sciences, University of Oxford
[2] Centre for Health Economics, University of York
[3] Nuffield Orthopaedic Centre, Oxford
[4] Royal Free Hospital School of Medicine, London
[5] London School of Hygiene and Tropical Medicine, London

Published December 1998

This report should be referenced as follows:

Fitzpatrick R, Shortall E, Sculpher M, Murray D, Morris R, Lodge M, *et al.* Primary total hip replacement surgery: a systematic review of outcomes and modelling of cost-effectiveness associated with different prostheses. *Health Technol Assess* 1998;**2**(20).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Acute Sector Panel and funded as project number 93/11/08.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will, in England, be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

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.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Editorial Assistant: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.soton.ac.uk/~hta

Contents

List of abbreviations

CI	confidence interval
HA	hydroxyapatite
HRQL	health-related quality of life
NOC	Nuffield Orthopaedic Centre
OA	osteoarthritis*
QALY	quality adjusted life-year
RA	rheumatoid arthritis*
RCT	randomised controlled trial
THR	total hip replacement

* Used only in tables

Executive summary

Objectives

- To identify the literature on primary total hip replacement (THR) surgery that is relevant to the question of whether prostheses differ in their medium to longer term outcomes, and to synthesise this evidence.
- To use evidence regarding both costs and outcomes of primary THR to model how much more effective newer prostheses must be to justify higher costs.

Methods

Data sources

- Electronic searches of MEDLINE and EMBASE (1980–1995).
- Hand-searches (1980–1995) of the 11 journals with the highest yield of relevant articles in the electronic searches.

Study selection

- Randomised controlled trials (RCTs) of any kind that compared prostheses for primary THR.
- Observational cohort studies that included concurrent controls.
- Observational studies of single prostheses with at least 5 years of follow-up and reporting outcomes in terms of revision rate or semi-standardised clinical assessment.

Data extraction and synthesis

It was not possible to carry out meta-analysis of the evidence from RCTs because each trial compared a unique pair of prostheses. A more informal form of meta-analysis was performed in which all data (randomised and observational) were combined for any prosthesis for which at least five independent studies reporting revision surgical rates were obtained. The meta-analysis was termed 'informal' because of the impossibility of controlling for numerous biases in the data and the poor quality of reporting of much of the evidence. Revision rates for eligible prostheses were calculated, adjusted for person-years at risk. Data were also combined for meta-analysis for other outcomes (i.e. hip scores, global ratings of success, and proportion of patients pain-free). However, studies lacked evidence of patient-based outcomes, and clinicians' views of outcome required substantial modification of diverse clinical ratings to produce a standardised score.

Costs and benefits of primary THR were assessed using Markov modelling, and calculation of costs per quality adjusted life-year, with sensitivity analysis of the results. Outcomes data were taken from a prospective study of a series of patients followed up for 14 years after THR. Costs were estimated from cost-generating events for THR and unit costs from a single centre (Nuffield Orthopaedic Centre, Oxford).

Results

Eleven RCTs were found that compared outcomes of prostheses. The trials followed up patients for short time periods (mean, 3.9 years) and had quite small sample sizes (mean, 168 patients). A significant difference between prostheses in terms of revision rate was observed in only one RCT.

When results of all reports that included a revision rate were combined, ten prostheses met the criterion set for a meta-analysis that at least five independent studies should be available for a prosthesis to be included. Adjusted THR revision rates (revision rate per 100 person-years at risk) were calculated for each of the ten prostheses to take account of different lengths of observation. The most favourable adjusted revision rates were found for the Exeter, Lubinus and Charnley prostheses. Intermediate results were found for the Müller, McKee-Farrar and Stanmore prostheses. The least favourable adjusted revision rates were observed for the Ring, Harris-Galante, PCA and Charnley-Müller prostheses.

Economic modelling indicated that to be cost-effective the following improvements in THR outcome and revision rates would be needed.

- For a new prosthesis costing three times more than the standard Charnley (i.e. typical cost of a new cementless prosthesis): ≥ 35–44% improvement in patients aged 50–70 years; ≥ 21–27% improvement in patients aged < 50 years.
- For a new prosthesis costing 1.5 times more than the standard Charnley (i.e. typical cost of a new

cemented prosthesis): 9–12% improvement in patients aged 50–70 years; 6–7% improvement in patients aged < 50 years.

From the available evidence, the extent of the improvement required of new and more expensive prostheses is particularly implausible for older patients. However the new cheaper prostheses may be cost-effective because the improvements required are more likely to be achievable.

Conclusions

There is a striking paucity of clear and relevant evidence on which to make well-informed choices about prostheses for primary THR. Although basic scientific innovation continues in relation to THR, the knowledge base to inform selection of prostheses is unlikely to improve in the foreseeable future.

Of prostheses commonly used in the NHS by far the greatest volume of evidence is available for the Charnley and on the basis of that evidence the Charnley appears to perform relatively well. However, the Charnley design has changed, and it is not clear how much of the evidence is relevant to the current design.

Of other prostheses currently used in the NHS, positive evidence (but no data from RCTs) was found in support of the Exeter prosthesis, and some positive evidence was found for the Stanmore (for example, evidence that it performed as well as the Charnley in an RCT). Positive evidence for the Lubinus IP (less widely used in the NHS) was also found. The quality of the evidence for other prostheses was either poor or non-existent. No substantial evidence could be found for cementless prostheses in terms of independent observation of results from five or more studies.

None of the analyses used in this review, such as meta-analysis of evidence, could overcome the fundamental weaknesses of the available evidence. The poor quality of evidence overall does not provide a basis clearly and authoritatively to identify prostheses that could be – or should not be – recommended for use by the NHS. However, it is clear that the more expensive the prosthesis, the more difficult it is to provide justification for its selection on the basis of the current evidence. On the basis of the economic analysis it seems that the use of the more expensive (i.e. cementless) prostheses is hard to justify on current evidence.

Recommendations for future research

As a substantial proportion of the evidence on outcomes of THR comes from healthcare systems quite different from the NHS (i.e. the Swedish and Norwegian national registers) it is recommended that the case for a UK register should be evaluated.

Least biased assessments would be from RCTs, but to detect the small but important differences that may exist between prostheses such trials must be more adequately designed and powered than those carried out previously, and should involve multicentre participation and long-term follow-up. Economic modelling in this review indicates that such trials might identify differences in cost-effectiveness between cemented prostheses.

Patient-based outcomes provide relevant and feasible methods to conduct large multicentre studies. To obtain unbiased assessments of outcome, the focus should be on outcomes of concern to patients, particularly pain and function, and not solely on revision surgery.

Chapter 1

Background and study objectives

Elective total hip replacement (THR) is a biomechanical solution to severe arthritis of the hip, involving removal of the damaged hip joint and its replacement with an artificial prosthesis. It is normally expected to achieve relief of pain and substantial improvement in mobility and physical function. THR has become the most common major orthopaedic surgical procedure and in the UK about 40,000 hip replacement operations are performed annually.

The artificial joint usually involves three elements: (1) a metal ball that replaces the original femoral head and which sits on (2) a metal stem which is inserted into the femur, and (3) a plastic cup which is inserted into the acetabulum. These three elements, referred to as the head, femoral and acetabular components, function much as a hip joint. Together they are referred to as a prosthesis. Prostheses are made by many different private UK and overseas companies, each of which manufactures its own brand or variant. Each manufacturer's prosthesis differs from those of competitors in details of design, materials and cost.

In the early hip replacement operations carried out in the 1950s, prostheses were implanted and fixed in place directly in the bone without use of cement. A major breakthrough in hip replacement surgery which resulted in decreased loosening was achieved by Charnley, who from 1962 onwards used acrylic bone cement to fix into place a metal femoral component and plastic acetabular component. Initially Charnley used Teflon for the acetabular component but because of catastrophic wear he later used polyethylene instead. Metal and polyethylene provide an interface with low friction and wear. Charnley's design has been considered very successful and is still the most widely used in the UK. Improvements in cement and in techniques of applying cement are thought to have further improved the durability of this form of prosthesis since the pioneering work of Charnley.

From the 1970s onwards a number of other prostheses were developed that were primarily intended to eliminate the need for bone cement. Press-fit prostheses were designed to fix in place by close fit alone. Threaded acetabular components which were intended to be screwed into bone were developed. Porous-surfaced prostheses were developed: in these the parts of the prosthesis adjacent to bone have small beads or mesh, and the aim is to encourage bone ingrowth into pores of the prosthesis surface to produce firm fixation. Another method intended to achieve cementless fixation, which was introduced in the 1980s, is the coating of the prosthesis with biological products such as hydroxyapatite (HA) that are intended to induce bone to grow to fit closely around the prosthesis. Many prostheses are now hybrid in that the femoral component is fixed with cement and the acetabular component is cementless. Other developments have included the use of ceramic rather than metal femoral heads to reduce wear. There is also a return to metal-on-metal bearing combinations, with prostheses made of various materials based on stainless steel, titanium or cobalt chrome. A wide variety of shapes for the femoral component have been developed. More recently both acetabular and femoral components have been made modular by many manufacturers.

As a result of this proliferation of technical variants of the original Charnley form of prosthesis there are now over 60 different prostheses from which to choose in the UK and new designs are continually being introduced.[1,2] A survey of current use of prostheses by orthopaedic surgeons, carried out as part of this report, indicates that although the Charnley prosthesis has a dominant position in terms of use by surgeons, many surgeons report using other designs (*Table 1*; see also appendix 1). Although reference is made to specific prostheses such as the Charnley it is important to recognise that designs are continually being changed.

THR is a successful operation for the majority of patients, and with most prostheses it has been estimated that only 10% of patients will require the THR to be revised within 10 years of the primary surgery.[3] Indeed some follow-up studies of the Charnley prosthesis at 20 years after surgery indicate 89% survival.[4] Apart from complications such as deep infection or dislocation, the main reasons for failure of THR are problems that occur after several years, primarily aseptic loosening of the prosthesis, due to resorption of bone supporting the prosthesis, fragmentation of bone cement or wear in the polyethylene part of the prosthesis.

TABLE 1 *Use of prostheses by UK orthopaedic surgeons, 1996*

Prosthesis	Surgeons reporting use as:			
	Femoral component % (n)		Acetabular component % (n)	
Charnley	52.6	(159)	55.6	(168)
Exeter	15.9	(48)	8.3	(25)
Müller	11.6	(35)	15.2	(46)
Furlong (cemented)	8.6	(26)	6.9	(21)
Furlong (cementless)	8.3	(25)	1.3	(4)
Ultima	6.9	(21)	5.6	(17)
Stanmore	6.3	(19)	6.3	(19)
Howse	5.0	(15)	3.6	(11)
CPT	4.3	(13)	3.9	(12)
ABG	4.0	(12)	3.0	(9)
CLS	3.6	(11)	3.6	(11)
C-Fit	2.6	(8)	1.0	(3)
Freeman	2.0	(6)	2.0	(6)
Harris-Galante	2.0	(6)	3.6	(11)
Cenator	1.7	(5)	2.0	(6)
Trilogy	0	(0)	2.6	(8)

Source: appendix 1.

For many designs of prosthesis, results in the first 5 years are good but failure rates increase markedly by 10 years after surgery.[5–7] In particular, the incidence of loosening or failure in the acetabular component is very low in the first 5–7 years after surgery but then rises rapidly.[8] It is for this reason that longer term follow-up is absolutely essential in this field of surgery.

A substantial amount of orthopaedic surgical time is now spent on revising primary hip replacement surgery. Approximately 13% of hip surgery involves revision rather than primary replacement.[9] Revision surgery is more expensive and time-consuming surgical work and is usually less successful than primary surgery. The need to revise surgery does not give a complete picture of the full extent of poor outcomes of primary surgery because patients may suffer pain and disability without having revision surgery.

A variety of factors may be implicated in poor outcomes of THR. Patients' characteristics, such as younger age, heavier weight, and higher level of physical activity are associated with poorer outcomes.[10,11] Evidence that surgical expertise is another factor comes from the observation that surgeons in training have higher revision rates for the THRs that they have performed.[12] Lower volume surgery of units performing the procedure also appears to be associated with poorer outcome, possibly because of lower levels of experience and expertise.[13] The current review is focused on the role of prostheses in influencing outcomes. Because of the relatively unregulated way in which prostheses are introduced into clinical practice, it is essential to estimate to what extent the proliferation of prostheses contributes to differing outcomes of THR surgery. It is commonly observed that orthopaedic surgery generally and hip replacement surgery specifically has an inadequate evidence base from well-designed trials with which to make choices between different surgical techniques and prostheses.[14–16]

The price to the NHS of a prosthesis ranges from £250 to £2000.[2] At present, it has been argued, there is insufficient systematic evidence of the outcomes of THR using different prostheses to enable surgeons and purchasers to make sound choices between prostheses which differ very substantially in terms of costs.[1,17,18] As will be argued later, if different prostheses are associated with different outcomes, particularly the need for revision surgery, a more complete analysis might indicate even greater variation in total costs of different prostheses.[19,20]

As the present review will demonstrate, there is a substantial volume of published evidence regarding outcomes of THR. A systematic review makes possible an efficient synthesis of such evidence. A study was therefore carried out in the form of a systematic review of available evidence regarding costs and outcomes of prostheses used in THR surgery. The study had two objectives:

(1) to carry out a systematic literature review of available evidence regarding medium to longer term outcomes of surgery involving prostheses used in primary THR in the NHS and to assess the extent to which such outcomes do vary by prosthesis
(2) to perform an economic model of the total costs (including costs of revision surgery) in relation to benefits associated with different prostheses to consider how much more effective newer and more expensive prostheses need to be to justify higher costs.

Chapter 2

Review of outcomes associated with different prostheses for THR: methods

Study question for systematic review

A systematic review was undertaken to examine whether different prostheses for THR were associated with different medium to longer terms outcomes. By medium to longer term is meant 5 years or later after surgery – an arbitrary but useful cut-off point based on consistent evidence that adverse outcomes such as revision surgery are rare for shorter periods of follow-up. Outcomes were defined as either the occurrence of revision surgery or standardised assessment of patients' pain and function. Studies assessing outcome solely in terms of radiological evidence of loosening were not included for two reasons: (1) studies are less likely to use standardised measurement of radiological evidence, making meta-analysis difficult, and (2) relationships between radiologically detected loosening and outcomes of importance to patients (pain or the need to have further surgery) are unclear.

Inclusion criteria

For inclusion in the structured review studies had to be concerned with primary THR surgery and to have been published during the period 1980–1995. The following types of study were to be included:

- randomised trials of prostheses for THR regardless of length of follow-up
- comparative observational studies of prostheses with concurrent controls regardless of length of follow-up
- observational studies of single prostheses with at least 5 years follow-up and using outcome measures as described below.

The cut-off point of 5 years was selected because of general agreement in the literature that, with existing clinical measures, rates of poor outcome such as revision surgery are very low for shorter periods (less than 1% per year following surgery).[8] A review of survival analyses of joint replacements confirmed that a cut-off point of 5 years follow-up was conservative and would omit studies with only small numbers of adverse outcomes.[3] For inclusion, observational studies also had to include outcomes assessed either in relation to need for revision surgery, or as a clinical assessment expressed in global terms (success or failure) or in terms of pain and function.

Exclusion criteria

Studies were to be excluded if they focused on issues that were not considered to be relevant to the review (congenital hip problems, hip fracture, hemiarthroplasty, outcomes of revision surgery). Also excluded were single case studies, and observational studies not using outcome measures as defined above.

Search strategy

The search for studies was to be conducted by handsearching selected journals and by electronic searches of the MEDLINE and EMBASE databases. As many as half of relevant trials may be missed if only electronic searches are used.[21] Journals to be handsearched were selected by identifying those with the highest rate of relevant publications. This was done by taking 3 years of the MEDLINE search database (1985, 1990, 1994) and identifying the 11 journals that appeared with the highest frequency (appendices 2 and 3). A similar strategy was repeated with the EMBASE search database (appendices 4 and 5), and it was found that the seven journals identified as having the highest yield of papers had already been selected via MEDLINE. Three further journals identified via EMBASE were not included for handsearching because of the very low total yield of relevant publications. Articles were included provided that either the article itself or an abstract was available in English.

The 11 journals listed in appendix 3 were therefore searched by three experienced handsearchers for the years 1980 to 1995. The searchers were instructed to identify and photocopy all reports of THR that met the criteria defined above.

Electronic searches of MEDLINE and EMBASE were also conducted for the years 1980 to 1995 based on the strategies shown in appendices 2 and 4. Photocopies of relevant reports were obtained.

Planned analyses

The analysis of studies found from the searches was governed by principles of a hierarchy of evidence.[22] Following these principles, evidence from well-designed randomised controlled trials (RCTs) was considered to be stronger. Evidence from observational studies was considered less strong. Observational studies were considered to provide stronger evidence if they had concurrent controls, that is they compared the performance of more than one type of prosthesis in THRs carried out by the same surgical unit(s) over an identical period.[22,23]

Application of these principles resulted in three planned analyses.

Analysis 1

Evidence from RCTs was to be inspected for patterns of effectiveness of prostheses. After this evidence had been examined and found to provide no strong or consistent evidence in favour of any prosthesis (see page 7), largely because of inadequate periods of time to follow up patients in RCTs, it was decided to conduct further analyses on the body of evidence that did contain longer lengths of follow-up, even though it was unlikely to contain strong evidence in terms of quality of study design.

Analysis 2

All observational cohort studies with concurrent controls were to be reviewed for any consistent pattern of evidence of outcomes associated with different prostheses. Although less convincing than evidence from well-conducted RCTs, this analysis would provide moderately convincing evidence of differences in outcomes associated with different prostheses. The results of this analysis are described in chapter 3 (page 7).

Analysis 3

A third analysis was to be undertaken that took account of all series* of observations of patients undergoing THR. Since evidence from the first two analyses provided few clear patterns of superior performance of particular prostheses (see chapter 3), this third analysis attempted to take account of all evidence available for prostheses regardless of the quality of study design, provided that the inclusion criteria for review were met (see above). It was recognised that such analyses might lend spurious precision to available evidence. However meta-analysis of observational evidence has been used to throw light on other questions in orthopaedic surgery. Moreover if no such analysis was undertaken by our group, there might always be a lingering question: for example, when considering the case for the need for large RCTs in this field, whether available evidence could not be better exploited if examined more fully. The methods used for the meta-analysis are described below.

Meta-analysis of all data combined

Because the data from RCTs had produced so little clear evidence of the longer term performance of prostheses, it was decided to combine data from as many sources as possible into a form of meta-analysis to obtain a more informal estimate of the possible extent of differences in outcomes between prostheses. The methods used for this analysis were adapted from a recent meta-analysis of prostheses in knee replacement surgery.[24]

Principles applied to the meta-analysis

The following rules and principles were decided in advance to guide the meta-analysis.

1. All reports obtained from the initial search were eligible to be included in the meta-analysis.
2. Studies were excluded if they provided no prosthesis-specific data on the primary outcome of surgical revision rates.
3. Duplicate reports of outcomes of prostheses were eliminated.
4. The meta-analysis would only examine evidence for prostheses for which at least five independent studies were found.
5. Where available, clinical assessments of outcome would be standardised so that a larger sample of studies could be included in an analysis of outcomes other than revision surgery.
6. A standard set of possible confounding variables would be extracted for each prosthesis series (length of follow-up, age of sample, proportions of patients in study samples who were female, proportions of patients with a diagnosis of rheumatoid arthritis or osteoarthritis).

* A series is a report of one sample of patients receiving a particular prosthesis, whether in the context of a RCT, comparative observational study or single observational study.

In relation to principle 2 above, the meta-analysis was used to calculate two rates of revision surgery, both of which were based on the total numbers of revisions as a percentage of all individuals at risk. The first (unadjusted) rate was simply a calculation of the overall rate of revision associated with a prosthesis which was derived from the number of revisions performed overall from all of the eligible studies of that prosthesis and expressed as a percentage of the total number of patients in the eligible studies. The second (adjusted) revision rate was considered more appropriate and informative. The adjusted revision rate was based on the same principles as the unadjusted rate, except that the revision rate for individual studies of a prosthesis was adjusted for the average length of follow-up of patients in the sample. When all studies for a prosthesis were combined, the adjusted revision rate was therefore the overall rate of revision surgery associated with a prosthesis in eligible studies per hundred person-years at risk. This is a more appropriate expression of available evidence for a prosthesis because it allows for the fact that as patients are followed up for longer periods of time, the number of cases of revision surgery tends to increase. Confidence intervals (95% CIs) were calculated for the adjusted revision rate. Although the adjusted revision rate is more informative than the unadjusted rate it is still based on the assumption that the annual revision rate for each prosthesis is constant.

In relation to principle 4, it was decided that there were sufficient numbers of methodological weaknesses in observational studies to require extreme caution in drawing inferences from single series. These problems were compounded when observational studies that derived from different periods, levels of surgical expertise, patient characteristics, and quality of data were combined to compare series in a meta-analysis. For this reason the principle was adopted that a prosthesis would only be included in the meta-analysis if at least five independent sets of observations from different data sets were available.

In relation to principle number 5 above, a preliminary scan of studies suggested that in addition to the rate of revision surgery for each study it would be possible to examine three other clinical outcomes, when reported. These outcomes were as follows.

- Many studies reported the surgeon's global rating of the outcome of the surgery ('excellent', 'good', 'fair', 'poor'). It was decided to record the percentage of patients with an outcome rated 'excellent' or 'good' combined reported at the last follow-up assessment of each prosthesis series.
- The second clinical outcome measure reported in THR studies and extracted for this meta-analysis was the percentage of patients who were considered by the surgeon or investigator at follow-up to be pain-free.
- The third outcome measure extracted from studies for meta-analysis was the mean hip score (reported at baseline, at follow-up, or as a change in score between baseline and follow-up). This measure required a simple procedure to standardise the variety of hip scores in use to a 0–100 range. The most commonly used clinical outcome measure used in THR research is the Harris hip score,[25] which provides a range from 0 to 100. Some other scores such as the Mayo hip score[26] are also standardised to this range. The Charnley score[27] provides separate scores on a scale 0–6 for each of three dimensions: pain, mobility and walking. Although these three dimensions are not intended to be combined, the Charnley score can be converted to a 0–100 range by calculating a percentage score for each dimension and calculating an overall mean for the three dimensions. Although it was recognised that the different scoring systems are not comparable and that they may produce different results when applied to a set of patients,[28] a standardised score might detect gross differences in outcomes in an exploratory meta-analysis. In this way, mean scores for each prosthesis series could be calculated, when recorded, both before surgery and at the last follow-up assessment. For some studies it was therefore also possible to calculate a change score between pre-surgical and follow-up assessment.

The meta-analysis was conducted on those reports of prosthesis series that were included after application of the pre-defined principles outlined above. In the analysis each separate report of a series of patients was treated as a single unit or case. For the purposes of this meta-analysis such series partly derive from the observational studies of single prostheses that constitute the subject matter of most of the papers found for this literature review (see chapter 3). The RCTs and comparative observational studies that were found were also included. Each prosthesis in the RCTs and comparative studies was entered as a separate series.

Estimates of outcomes

From the absolute numbers of events reported in each series, simple summing from across the

relevant studies provided overall estimates for each prosthesis of the revision rate, the percentage of patients with an outcome assessed as 'excellent' or 'good' at last follow-up, and the percentage of patients assessed as pain-free at last follow-up. Thus, for example, the revision rate for a prosthesis derived from all available studies is the absolute number of revisions divided by the total number of patients from all relevant studies. For the variables for which studies provided mean data rather than absolute numbers (i.e. the mean pre-surgical and follow-up Harris hip score, or equivalent, for a sample), to derive an overall Harris hip score for each prosthesis from all available studies a sample size-adjusted mean score was calculated. In this calculation the mean Harris hip score for each study was multiplied by the number of patients in that study. The products of the calculations were added and the result was divided by the total number of patients for all studies of the prosthesis in question.

Statistical analysis

To test whether there were statistically significant differences between prostheses in the clinical outcomes described above, and whether any such differences were independent of possible confounding effects of the characteristics of study samples, analysis of variance with covariates was performed with each study entered as a case. The dependent variables of these four analyses were, for each study:

- the rate of revision surgery adjusted for length of follow-up
- the mean Harris (or equivalent) follow-up score
- the percentage of patients with an outcome rated as 'excellent' or 'good'
- the percentage of patients who were pain-free.

Main effects were examined for prosthesis type and whether or not the prosthesis was fixed with cement. The following four variables were included as covariates:

- mean age of patients in the sample
- the percentage of female patients in the sample
- the percentage of patients who had rheumatoid arthritis as a main diagnosis
- the percentage of patients who had osteoarthritis as a main diagnosis.

The duration of follow-up was also used as a covariate in analyses of the dependent variables other than the adjusted revision rate (which included the influence of duration of follow-up as a transformed variable). Inevitably this analysis can only take account of covariates at the level of study, rather than on the basis of individual patients.

Quality assessment of studies

Quality assessments of studies should be carried out because more accurate estimates of the outcomes of treatments can be derived from more rigorous evidence. Assessments are usually made to exclude less rigorous studies, to weight studies in meta-analysis, or to perform sensitivity analyses of assumptions or results of meta-analysis. As part of the current review a study was undertaken to assess the quality of studies of medium to longer term outcomes (appendix 6). The criteria that resulted from this assessment were not applied to every study identified as a result of the literature search because it became clear from an early stage that few strong inferences of clear advantages between prostheses were going to emerge from the evidence. Since hardly any of the evidence provided any support for particular prostheses, stratifying evidence or excluding evidence could not make any difference to our conclusions. The methodological review of orthopaedic research (appendix 6) nevertheless provides important evidence of the limitations of current evidence for selecting between prostheses.

Chapter 3

Review of outcomes associated with different prostheses for THR: results

A total of 191 reports were found that fulfilled the agreed criteria. In these 191 papers, 11 RCTs were reported[12,29–38] and the remaining 180 papers reported observational studies. There is some duplication in papers reporting studies.

Evidence from RCTs

From a methodological perspective, the 11 RCTs had quite short periods of follow-up: the mean length of follow-up was 3.9 years (range of follow-up, 1 year to 6.5 years). This is considerably short of the period at which adverse events following THR are likely to occur at higher frequency. All of the trials involve small sample sizes: the mean sample size overall was 168 patients (range, 28–413 patients). Such trials are unlikely to have sufficient power to detect relatively uncommon adverse outcomes. Only one trial used independent measures of outcome as assessed directly by patients.[38]

The RCTs are summarised in *Table 2*. Significant differences between the prostheses investigated were reported in six of the trials, and in only one was a significant difference reported in terms of revision rate: Reigstad and colleagues[29] found a significantly higher revision rate for THR with the ICLH prosthesis than with the Müller prosthesis. Two trials observed statistically significant differences between prostheses on the basis of clinical hip scores. Jacobsson and colleagues[32] found significantly better Harris hip scores for the cementless rigid PCA prosthesis compared with the isoelastic Butel prosthesis. Søballe and colleagues[34] found that the uncemented Biometric prosthesis coated with HA was associated with more favourable Harris hip scores after 1 year than the same device coated with titanium alloy.

Three trials produced significant results in terms of radiographic evidence of migration or loosening. Two of the three trials indicated advantages of cemented over cementless prostheses.[31,33] The third trial[35] provided evidence of the advantages of HA compared with other methods of coating.

The advantages of combining RCTs to increase sample size for meta-analysis could not be explored as no two trials addressed the same question. In an informal sense, Søballe and colleagues[34] and Karrholm and colleagues[35] both provided consistent evidence of the value of HA coating for cementless prostheses. In a similarly informal sense, four out of 11 trials included the Charnley prosthesis as one arm of the trial and in none did it perform significantly worse than other prostheses.

Another way of reviewing the evidence from RCTs is to concentrate on the trials that followed up patients for 5 years or more. Three of the four trials that reported results for 5 years included the Charnley prosthesis as one of the prostheses in the comparison. In these three trials no difference in outcome was observed when the Charnley prosthesis was compared with the Spectron prosthesis,[37] or when it was compared with the Stanmore prosthesis.[12] In the third trial, the Charnley prosthesis was found to have a better survival rate than the cementless HP Garches prosthesis.[30]

Overall, the evidence from RCTs in THR provides no clear evidence of the relative advantages of prostheses. For this reason, the next most robust source of evidence was also examined – observational studies with comparative data on prostheses.

Evidence from comparative observational studies

The second set of data that was examined for trends in outcomes following THR with different prostheses comprised 21 papers[5,6,39–57] reporting 18 comparative studies that fulfilled the initial inclusion criteria set, namely that (1) comparative data were available for more than one prosthesis, (2) evidence was based on concurrent rather than historical comparisons, and (3) evidence was available on outcomes with each prosthesis in terms of the need for revision surgery, clinical assessment (such as a global scale or a rating of pain and/or physical function), or patient-based outcome. Although these 18 studies did not involve random allocation, they were considered of some value

because the performances of different prostheses can be broadly compared because surgery and follow-up for patients in each prosthesis group were broadly contemporaneous. These studies are summarised in *Table 3*.

In terms of the relevance of the evidence to the questions of this review, the 18 comparative studies were more valuable than the RCTs for THR in that all except three studies[45,50,52] followed up patients for at least 5 years. These three shorter comparative studies are included in *Table 3*. In terms of methodological criteria however the 18 comparative studies varied considerably. A small number of studies had the advantage of a large sample size, particularly the studies based on the Swedish[6] and Norwegian[52,53] THR registers. The Swedish study in particular, by providing data on outcomes on over 92,000 patients undergoing THR, has by far the largest sample size of any study of prostheses. The Swedish and

TABLE 2 *RCTs of prostheses for THR*

Study	Prostheses compared	No. of patients (hips) per prosthesis	Duration of follow-up	Main results*
Reigstad *et al.*, 1986[29]	Müller vs. ICLH (double cup)	155 149	4 years	Revision: Müller 0% vs. ICLH 8.7% ($p < 0.001$).
Wykman *et al.*, 1991[30]	Charnley cemented vs. HP Garches	75 75	5 years	Survival at 5 years: Charnley 88%, HP Garches 82%. Median Harris hip score: Charnley 95.3, HP Garches 88.7.
Godsiff *et al.*, 1992[31]	Ring UPM femoral cemented vs. Ring cementless	30 28	2 years	No differences in % pain free but 96% cemented and 62% cementless walked with no aid ($p < 0.05$).
Jacobsson *et al.*, 1993[32]	Two cementless femoral: isoelastic Butel vs. rigid PCA	 28 28	4 years	Failure (revision or loosening) 43% in Butel and 11% in PCA. Harris hip score good/excellent in 50% Butel, in 82% PCA ($p < 0.001$).
Reigstad *et al.*, 1993[33]	Cemented Landos Titane vs. uncemented Zweymuller/ Endler	60 60	5 years	No differences in bone loss. More radio-opaque double line in cementless vs. cemented ($p < 0.001$) but no difference in loosening.
Søballe *et al.*, 1993[34]	Biometric cementless with femoral component Ti-coated vs. HA-coated	 13 (hips) 15 (hips)	1 year	Migration: Ti-coated 3.9 units vs. HA-coated 1.7 units ($p < 0.02$). Harris hip score: Ti-coated 87 vs. HA-coated 98 ($p < 0.01$).
Karrholm *et al.*, 1994[35]	TiFit straight femoral stem; fixation by (a) cement, (b) HA coating or (c) porous coating	(a) 20 (hips) (b) 23 (hips) (c) 21 (hips)	2 years	No clinical differences but subsidence less for HA than for cement or porous ($p < 0.002$).
Onsten *et al.*, 1994[36]	All polyethelyne Charnley socket vs. Cementless porous Harris-Galante	 21 (hips) 21 (hips)	2.25 years	0.2-mm migration for both.
Garellick *et al.*, 1995[37]	Charnley vs. Spectron	206 (hips) 204 (hips)	5.6 years	No differences in revision rate: Charnley 2% vs. Spectron 1%.
Marston *et al.*, 1996[12]	Charnley cemented vs. Stanmore cemented	200 213	6.5 years	No differences in revision rate (4%) or radiological migration.
Mulliken *et al.*, 1996[38]	Mallory cemented vs. Mallory cementless	76 71	4.8 years	No revision cases in either. Harris hip score 93 in both. No differences in quality of life.

* *Differences not significant unless otherwise stated;* p *values are given when stated by authors; otherwise* p *values were unreported or reported as non-significant.*

TABLE 3 *Comparative observational studies*

Study	Prostheses compared	Sample size (n)	Mean duration of follow-up	Methodological strengths of study	Main results
Sudman *et al.,* 1983[39]	Charnley, Christiansen	203	6.5 years	–	Failure: Charnley 4%, Christiansen 31%.
Djerf & Wahlstrom, 1986[40]	McKee-Farrar, Charnley	177	5 years	–	No differences.
Updated: Jacobsson *et al.,* 1990[41]			11.5 years	Survival analysis	Survival: McKee-Farrar 82%, Charnley 90%.
Ritter & Campbell, 1987[5]	Charnley, Müller, T-28	746	10 years	Survival analysis and adjustment by Cox regression; single surgeon.	Survival: Charnley 91%, T-28 88%, Müller 80%, ($p < 0.05$).
Updated: Ritter, 1995[42]			Follow-up, 1–22 years		Survival: Charnley and T-28 continued to show better survival rate than Müller ($p < 0.001$).
Agins *et al.,* 1988[43]	Charnley, Charnley-Müller, T-28	122	11.2 years	Survival analysis and adjustment by multiple regression.	Survival: Charnley better survival rate (83%) than Charnley-Müller (64%) ($p < 0.05$).
Carlsson *et al.,* 1988[44]	Brunswick, Charnley	321	9 years	–	Survival: Brunswick 90.4%, Charnley 100%.
Collis, 1988[45]	10 prostheses (most common were Iowa, TR-28, Charnley)	1436	3.9 years	Single surgeon.	Revision rate: Charnley lower (3.3%) than Müller (14.3%).
Wilson-MacDonald & Morscher, 1989[46]	Müller: straight, curved or 130°	545	Unclear (between 5 and 10 years)	–	Radiological loosening greater with 130° than other types of Müller ($p < 0.05$).
Ritter *et al.,* 1990[47]	Miami MOSC, metal-backed or non-metal-backed	238	7.6 years	Single surgeon, survival analysis (Kaplan-Meier).	Revision rate: metal-backed 6%, non-metal-backed 2% ($p < 0.005$).
Timperley *et al.,* 1992[48]	Ring: cemented and cementless	526	9 years	Single surgeon, survival analysis.	Survival: cemented 97%, cementless 96%.
Malchau *et al.,* 1993[6]	All prostheses commonly used in Sweden	92,675	Approx. 10 years	Survival analysis (Kaplan-Meier), use of CIs.	Survival: Charnley, Lubinus, CAD do better than Exeter matt, Müller curved, which do better than Christiansen.
Müller *et al.,* 1993[49]	Cementless: Zweymuller vs. Endler and RM (both poly-ethylene cup)	1025	Approx. 10 years	Survival analysis.	Failure rate: Zweymuller (1.3%) lower than both polyethylene cups (8%).
Huracek & Spirig, 1994[50]	Mecron cementless, with and without HA coating	80	4 years	Surgery by one surgeon and outcomes assessed independently.	Pain: less with HA-coated ($p < 0.05$). No other differences in clinical outcomes.
Johnsson *et al.,* 1994[51]	Charnley, Brunswick, Lubinus snap fit	799	10 years	Survival analysis and Cox regression adjustment with CIs, also no loss to follow-up.	Revision rate: Charnley 5.5%, better than Brunswick 17.5% ($p < 0.001$), or Lubinus 17.6% ($p < 0.01$)

continued

TABLE 3 contd *Comparative observational studies*

Study	Prostheses compared	Sample size (n)	Mean duration of follow-up	Methodological strengths of study	Main results
Espehaug *et al.*, 1995[52]	10 most commonly used cemented prostheses in Norway	12,179	3.2 years	Survival analysis with Cox regression adjustment.	Failure rate: Spectron/ITH better than Charnley ($p < 0.05$); Müller type and Elite worse than Charnley ($p < 0.01$).
Havelin *et al.*, 1995[53]	11 most commonly used cementless prostheses in Norway	4352	Between 3 and 4 years	Survival analysis (Kaplan-Meier) with Cox proportional hazards adjustment.	Revision rate: HA-coated and hemispheric cups revised less often than either metal-backed screw ($p < 0.01$) or all polyethylene ($p < 0.001$).
Hwang and Park, 1995[54]	Cementless: PCA, AML, HGP	270	5 years	Single surgeon.	No significant clinical differences between prostheses.
Jacobsson *et al.*, 1995[55]	Cemented: ITH, Lubinus SP 2	142	5 years	Multiple regression analysis of differences.	Radiolucency: ITH greater ($p < 0.02$).
Ranawat *et al.*, 1995[56]	Charnley, Triad	226	9 years	–	Failure rate (radiographic loosening): Charnley 10%, Triad 1.5%.
Weidenhiem *et al.*, 1995[57]	Cemented Exeter CPT, cementless porous-coated PCA	151	6.5 years	Single surgeon.	Revision rate: Exeter 3.5%, PCA 34.8%.

Norwegian registers also attempted to represent the full range of clinical practice in those countries since all THRs were expected to be reported. Most other comparative studies represent the clinical practice of one, or a small number of, surgeons. The Swedish register is also of particular value because there are 10 years of observation for many of the prostheses in the register. The results of the Norwegian register are particularly valuable because the investigators have used Cox proportional hazards models to adjust for the effects of type of cement, use of antibiotic prophylactic, type of operating theatre, and the age and gender of the patients. Such analyses have not been carried out in the majority of studies reported in this review. Unfortunately the Norwegian register began later than the Swedish register and so prostheses have been observed for a much shorter period. At the other extreme to the two national registers are comparative studies with very small sample sizes and in which no adjustment is made for possible confounding effects.[39,44]

In the Swedish register study[6] it was found that the Charnley, Lubinus and CAD prostheses had the lowest revision rates, that the Exeter matt surface and the Müller (curved stem) prostheses had intermediate rates, and that the Christiansen prosthesis had the highest revision rate. Some more recently developed prostheses, the Spectron, Lubinus SP and Scan Hip, had very low revision rates but had not been in use for as long an observation period. Results from the Norwegian register are reported separately for cemented[52] and cementless[53] prostheses. With regard to cemented prostheses, after adjusting for possible confounding, the Spectron/ITH prosthesis had the lowest revision rate. The Müller (curved stem) had the highest revision rate. Next highest was an Elite/Charnley combination. The rate for the basic Charnley prosthesis was intermediate. Amongst uncemented prostheses, all polyethylene-backed prostheses (such as Endler) had by far the highest revision rates, and metal-backed screw prostheses (such as Ti-Fit) also had high revision rates. In comparison, porous-coated and HA-coated cups had low revision rates.

Overall, the two largest studies therefore provide evidence of some prostheses that performed significantly worse, in particular the Christiansen prosthesis and, to a lesser extent, the Müller (curved stem) prosthesis. The Charnley was by far the most commonly used prosthesis in both countries, accounting for 32% of all prostheses in Sweden in the period 1978–1990 and for 55% of all cemented prostheses in Norway during the period 1987–1993. The results with the Charnley prosthesis were favourable in both countries.

The other 16 comparative observational studies are much smaller in sample size: the combined total

sample (n = 7007) from all other comparative studies is smaller than the Norwegian register, and considerably smaller than the Swedish register. The smaller comparative observational studies compared a diverse range of prostheses for different lengths of time. It is striking however that, of the eight studies that include the Charnley as one of the prostheses observed, all except one small study which used radiographic loosening as the measure of failure[56] show the Charnley performing more favourably than comparison prostheses. No specific pairs of prostheses are compared sufficiently frequently to warrant combining studies at this point.

Observational data for single prostheses

The 159 remaining papers which were identified from the search and fulfilled inclusion criteria for this report contain observational data of single prostheses. The studies reported in these articles were included in the meta-analysis, the results of which are described below.

Results of meta-analysis

The primary merit of the meta-analysis is that it makes fullest use of the large volume of data and therefore of the large total sample sizes for the more commonly reported prostheses. However it is important to stress that this form of meta-analysis provides only an informal and approximate estimate of likely differences between prostheses. The meta-analysis required that studies were combined from different periods over the study period of 1980–1995, and from surgical centres with different levels of experience and expertise. The analysis combined series of patients with differing disease severity, co-morbidity, age and activity levels. There is therefore very substantial scope for bias in such evidence. Moreover the standard of reporting of studies was sometimes so poor that it was not possible to make any systematic adjustment for quality of measurement, degree of observer bias in assessment of outcome, quality of follow-up and degree of representativeness of reported cases. This is therefore a fairly tentative form of meta-analysis which might more accurately be considered an approximation of possible degrees of differential outcomes of prostheses.

Of the 191 reports initially found for the review, 97 were omitted either because they were duplicate reports or because they reported results on a prosthesis that failed our criterion requiring at least five independent reports. The informal meta-analysis was therefore based on 94 separate reports and papers. Eleven papers report on prostheses in the context of RCTs.[12,29–38] Twenty-one papers report on prostheses in the context of comparative observational studies.[5,6,39–57] Sixty-two papers report on prostheses in the form of a report of a single prosthesis.[58–119] These 95 reports produce evidence on 118 separate series of observations on a prosthesis.

There were ten prostheses for which the literature search found at least five independent studies of outcomes assessed in terms of rate of revision surgery: Charnley, Müller, PCA, Ring, McKee-Farrar, Harris-Galante, Stanmore, Charnley-Müller, Lubinus, Exeter. The number of patient series (hereafter referred to as 'studies') found for each prosthesis varied considerably. Overall, 44% of all studies reported on patients who had received a Charnley prosthesis *(Table 4)*. No other prosthesis has been as frequently examined. In terms of the numbers of patients observed, after Charnley, the Lubinus prosthesis has been next most frequently monitored, largely because of the Swedish and Norwegian registers *(Table 4)*. Four prostheses (Charnley, McKee-Farrar, Stanmore, Charnley-Müller) have been monitored in studies that when combined have a mean length of observation of at least 10 years.

The characteristics of patients differ for different prostheses *(Table 5)*. The mean age of patients is lower and the proportion of female recipients is lower for the PCA and Harris-Galante prostheses than for other prostheses.

Analysis of variance showed significant differences between prostheses for the adjusted revision rate ($F = 3.13$, df 9, $p < 0.01$), the Harris hip score ($F = 5.42$, df 4, $p < 0.01$), and for the percentage of patients pain-free ($F = 3.52$, df 4, $p < 0.05$). Differences were not significant for the percentage of patients with outcome rated excellent or good at last follow-up.

The meta-analysis identified three prostheses with an overall unadjusted revision rate greater than 10%: Ring, Charnley-Müller and McKee-Farrar *(Table 5)*. The adjusted revision rate takes into consideration the fact that prostheses have been available for differing lengths of time and that the longer the length of follow-up of a prosthesis in a study the greater the likelihood that revision surgery will occur. The adjusted revision rate identified four prostheses with revision rates

greater than 1 per 100 patient-years of follow-up: Ring, Harris-Galante, PCA, and Charnley-Müller. A fifth prosthesis, McKee-Farrar, shows an adjusted revision rate very close to 1 per 100 patient-years of follow-up.

Three prostheses – Exeter, Lubinus, and Charnley – had very favourable results in the meta-analysis, whether expressed in terms of crude unadjusted rates or rates adjusted for person-years at risk. Their adjusted revision rate was less than 0.5% per annum.

TABLE 4 *Characteristics of studies included in meta-analysis of outcomes of prostheses*

Prosthesis	No. of studies	Total no. of patients	Mean study sample size (n)	Mean length of follow-up in years (range)
Charnley	52	39,249	755	11.0 (2–20)
Müller	16	3349	209	9.1 (5–14)
PCA	10	1246	125	5.6 (5–7)
Ring	9	1817	202	7.5 (2–21)
McKee-Farrar	6	3577	596	10.0 (5–14)
Harris-Galante	5	346	69	4.0 (2–6)
Stanmore	5	2474	495	11.0 (8–15)
Charnley-Müller	5	3296	659	11.0 (6–15)
Lubinus	5	15,707	3141	8.8 (5–12)
Exeter	5	6314	1263	6.6 (5–10)
Total	**118**	**77,375**	**656**	**9.4**

TABLE 5 *Patients' characteristics and outcomes of studies in meta-analysis of prostheses*

Prosthesis	Patients' characteristics*				Outcomes*					
	Age (years)	Female (%)	RA (%)	OA (%)	Unadjusted revision rate (%)	Adjusted revision rate (per 100 person-years at risk)†	Follow-up mean hip score (%)	Mean hip score change (%)	Rated 'excellent' or 'good' (%)	Pain-free (%)
No. of studies‡...	–	–	–	–	118	118	38	21	25	27
Charnley	66.1 (33–78)	65.9 (31–90)	13.2 (0–100)	69.6 (0–100)	4.7 (0–18)	0.37 (± 0.02)	85.6 (59–95)	39.3	83.5 (51–97)	84.1 (46–100)
Müller	66.8 (49–76)	63.1 (43–83)	5.1 (0–16)	82.7 (42–100)	7.4 (0–27)	0.68 (± 0.10)	83.8 (79–88)	45.8	81.5 (63–92)	–
PCA	57.6 (47–67)	47.7 (32–53)	10.2 (0–23)	40.1 (0–100)	7.5 (1–38)	1.31 (± 0.29)	89.6 (85–94)	40.3	85.9 (84–90)	73.7 (73–74)
Ring	63.8 (63–67)	60.1 (53–73)	7.9 (0–8)	61.1 (0–100)	23.1 (0–26)	2.04 (± 0.19)	93.0 (78–98)	–	68.7 (38–98)	43.2 (34–49)
McKee-Farrar	63.4 (59–75)	62.7 (55–75)	7.7 (5–11)	46.5 (0–87)	13.2 (4–23)	0.98 (± 0.08)	73.9 (73–76)	–	54.3 (49–62)	81.6 (55–75)
Harris-Galante	55.0 (49–69)	41.8 (39–44)	2.3 (0–5)	53.8 (0–100)	8.4 (0–10)	1.40 (± 0.58)	90.1 (83–93)	35.5	–	–
Stanmore	66.0 (63–81)	64.1 (51–81)	2.9 (0–6)	86.2 (80–100)	7.3 (6–22)	0.62 (± 0.09)	–	–	–	72.8 (43–87)
Charnley-Müller	61.1 (59–65)	65.7 (61–71)	18.8 (5–100)	55.6 (0–68)	15.5 (3–45)	1.10 (± 0.10)	–	–	–	–
Lubinus	65.4 (61–73)	64.9 (47–72)	11.2 (0–19)	83.3 (71–100)	3.2 (0–17)	0.27 (± 0.03)	–	–	–	–
Exeter	70.0 (63–71)	65.3 (48–67)	10.1 (6–18)	83.2 (73–85)	1.3 (1–7)	0.18 (± 0.04)	–	–	–	–

RA = rheumatoid arthritis; OA = osteoarthritis

* *Values are means with range in brackets unless otherwise stated.*

† *95% CI in brackets.*

‡ *No. of studies on which outcomes are based.*

Chapter 4

Discussion of evidence for medium to longer term outcomes for prostheses

Quality of evidence

To maximise the yield of relevant studies this review was based on literature identified by an electronic search as well as by handsearching of selected journals. It is possible that relevant evidence was under-sampled because of the linguistic limitations of the study. Items were included in the review provided that an article or abstract was in English. Although this strategy limited access to non-English language sources, English language abstracts enabled us to include European and other non-English language journals in both electronic and handsearching. However since the focus of this review was upon prostheses used in the NHS, it was thought unlikely that important high-quality evidence on prostheses in use in the UK would be omitted. So remarkable are the rare high-quality trials in orthopaedics it is unlikely (although impossible to rule out completely) that a high-quality trial of prostheses currently in use in the NHS would pass unnoticed in discussion in English language abstracts or would not be publicised by manufacturers or other interested parties.

There are reasons to think that the current review is based on most of the available relevant evidence for prostheses in use in the NHS. An informal assessment of the thoroughness of the current literature review can be made by comparing the yield of RCTs obtained by our search strategy with other reviews with similar aims reported in 1995 and 1996. Cowley[120] conducted a systematic review of prostheses, the literature review for which yielded eight RCTs. Yahiro and colleagues[121] carried out a systematic review with a slightly narrower focus on the acetabular component and found no RCTs. The review on THR published by the NHS Centre for Reviews and Dissemination[122] refers to five RCTs. A much earlier search, reported in 1988,[14] found only one RCT of THR. The fact that our search yielded 11 RCTs, a somewhat higher number than any previous review, provides informal but encouraging evidence that we have obtained a substantial proportion of the relevant published evidence. Moreover, a survey inviting manufacturers and distributors of prostheses in the UK to make available any evidence they had regarding the performance of prostheses identified no RCTs and no other important evidence that was not found by our search strategy.[2]

This review provides clear evidence for the commonly expressed view that there is a paucity of high-quality research evidence, particularly RCTs, for orthopaedic surgery in general, and for THR specifically. Our review provides further confirmation of the lack of RCTs for THR that has been commented on previously.[14–16,120–122] A similarly low yield of only four RCTs was found in a systematic review of outcomes of total knee-replacement surgery.[24]

Detailed discussion of the reasons for the lack of RCTs in THR surgery are beyond the scope of this review. Surgeons generally have difficulties with randomisation between surgical options when they may have greater skill and experience in one option and tend to have preconceived views of the advantages of options.[14,123] There are other barriers that may discourage surgeons from performing RCTs: for example, the perception that multi-surgeon and multicentre trials are more difficult to conduct, analyse and interpret in surgery.[124] It is clear that a particular problem in THR surgery is that outcomes of importance do not begin to emerge until 5 or even 10 years after intervention. This may make the funding, logistics and motivation to participate in RCTs problematic.[125] At present investigators undertaking a long-term RCT also have to contend with the recognition that novel prostheses and related changes will continue to be introduced before trial results are reported with the consequent risk that results will be perceived as redundant.[15]

Examination of the results from RCTs found in our literature search provided no very clear indications of the relative advantages of different prostheses because of the small number of trials, their small sample sizes, the short periods of follow-up and because no two trials examined the same comparative question and so no meta-analysis of RCTs could be performed.

Overall we conclude that there is little if any high-quality evidence available to inform decisions about

choice of prostheses for primary THR by orthopaedic surgeons in the NHS.

It has been argued that orthopaedic surgery is one of the areas of health care in which well-designed observational studies are needed to examine effectiveness because of the many difficulties of conducting RCTs.[126] Because of the weakness of the evidence from RCTs, in our review we examined the evidence from comparative observational studies with concurrent controls. When the comparative observational studies are examined as a group, the evidence from the Swedish register is by far the most important for the purposes of this review because it has a variety of methodological advantages, especially sample size, length of observation and generalisability.[6] The Norwegian data and other comparative studies provided some additional support for some of the trends observed from the Swedish data.

There remain concerns about potential biases from reliance on observational sources. Whereas Swedish observational data show that after 9 years the Spectron prosthesis was performing considerably better than the Charnley prosthesis,[127] the evidence from the Swedish RCT of the Spectron and Charnley prostheses[37] shows no difference at 5.5 years. It is possible to speculate that newer prostheses such as the Spectron are taken up by enthusiastic and interested surgeons who also have superior technique. It is difficult to estimate possible effects of the surgeon and other potentially biasing factors on the observational evidence.

Finally we performed a meta-analysis on a broader range of studies that included observational studies of single prostheses – by far the most common form of evidence in this field. We recognised beforehand that this would provide a weaker form of evidence than that from RCTs or observational studies with concurrent comparisons. Nevertheless the informal meta-analysis of the ten most commonly studied prostheses was consistent with the notion that THR outcomes differ with different prostheses, and that such differences remain statistically significant when some possible confounders have been controlled for to some extent.

Conclusions

There is overwhelming evidence that overall THR is a successful form of surgery. The majority of patients can expect to enjoy improved function and reduced pain for many years after THR and will not need revision of their surgery. The present review is concerned with a more specific question: the contribution to medium to longer term outcomes of the large array of available prostheses. On that more specific issue, the review primarily underlines how little is known with confidence on the basis of clear evidence.

Nevertheless, a consistent trend of good performance of the Charnley prosthesis can be observed across a range of data sets, study designs and analyses. For the Charnley prosthesis, outcomes have been studied in a larger number of patients for longer periods of observation than is the case with any other prosthesis. There remain difficulties in the interpretation of this evidence because the design has not remained the same over the history of the Charnley prosthesis.

In comparison with this reasonably confident statement that can be made about the Charnley, conclusions for all other prostheses commonly used in the UK are much more tentative because they are based on much weaker evidence. The Exeter prosthesis had particularly favourable revision rates in the meta-analysis. It was introduced quite early in the history of THR, first being used in 1970. Its distinctive feature is the lack of any collar on the femoral stem, which is polished and tapered. In the data from the Swedish register, the Exeter polished performed well compared with the Charnley, whereas the Exeter matt surface performed less well. The Exeter has been included in no RCTs that could be found. In the Norwegian register it was the second most commonly used prosthesis after the Charnley,[52] and in the Swedish register it was the third most commonly used prosthesis after the Charnley and the Lubinus.[6] In our survey of NHS orthopaedic surgeons it was found to be second only to the Charnley in frequency of use.

The meta-analysis indicated very favourable results for the Lubinus prosthesis. Studies of the Swedish and Norwegian registers indicated that the Lubinus SP performed particularly well. However as Malchau and colleagues[6] comment, because this device was introduced in the 1980s it has had a shorter period of observation. It is used by less than 1% of the surgeons in our NHS survey.

By contrast some prostheses – Ring, Harris-Galante, PCA, and Charnley-Müller – appear to have less favourable results. The highest adjusted revision rate overall in the meta-analysis was found for the Ring. This device also had poor results in terms of the percentage of patients at last follow-up who were reported as pain-free.

The Ring uncemented prosthesis was one of the early attempts to improve upon the Charnley cemented, its main advantage being ease of revision if needed.[102] Like many prostheses it has undergone many modifications: initially it was an all-metal prosthesis but subsequently a metal and polyethylene version was developed. In a small RCT significant improvement after 2 years was reported when the Ring was fixed into place with cement compared with cementless fixation.[31] However it was also reported that the trial was discontinued on the grounds that it was unethical after the investigators noted an unacceptable number of femoral stem breakages. The authors of the paper also report discontinuing the use of the prosthesis altogether. The Ring prosthesis was reported as being used by less than 1% of orthopaedic surgeons in our survey.

The Harris-Galante is a cementless prosthesis, fixed by its porous coating. It has tended to be used in younger patients. Methods of analysis in this review do not allow fully to adjust for younger age which may independently contribute to poorer outcomes. The Harris-Galante prosthesis is somewhat more commonly used in the UK than the Ring prosthesis; 3% of NHS orthopaedic surgeons reported using the acetabular component. Some of this use may be in the form of hybrid combinations of Harris-Galante with other components.

The PCA is also fixed in position without cement. Like the Harris-Galante, it is used on somewhat younger patients. In the meta-analysis the revision rate for the PCA was intermediate between favourable and less favourable prostheses. However it is noticeable that the length of follow-up for this prosthesis is somewhat shorter than for most other prostheses. Owen and colleagues[7] noted a marked deterioration in survival of this prosthesis at 6–7 years after surgery. Although the PCA is distributed in the UK, less than 1% of orthopaedic surgeons used it.

From the pooled evidence of the meta-analysis the Charnley-Müller also appeared to have a very high revision rate. The use of this prosthesis was not reported by any of the orthopaedic surgeons in our survey, and it was not reported as being distributed in the UK in the survey by Murray and colleagues.[2] The McKee-Farrar prosthesis also had a high overall revision rate and is used by less than 1% of orthopaedic surgeons.

A few prostheses, most notably the Christiansen, have performed particularly poorly as evidenced by the Swedish register.[6] However that prosthesis is no longer in use. The same source provides some suggestion of less consistently satisfactory performance of the Müller (curved stem) prosthesis.

Chapter 5

Modelling of cost-effectiveness of THR: methods

Introduction

As has already been indicated, the costs set by manufacturers for prostheses for THR vary considerably.[2] It has been estimated that costs of prostheses may represent up to 40% of the total costs of this form of surgery.[120] In this second part of our report, an attempt is made to model costs in relation to effectiveness of prostheses to consider to what extent higher costs of some prostheses may be justified by greater potential overall benefits.

Despite the dearth of data on the effectiveness of the newer prostheses, decisions need to be taken not only on whether their use can be justified in the context of routine practice but also on whether the NHS should devote additional research and development resources to resolving the uncertainties surrounding their use. Economic evaluation based on decision analytic modelling can generate valuable information to assist in making these sorts of decision.[128] Such economic evaluation can:

- identify which uncertainties really matter in terms of their impact on the cost-effectiveness of different prostheses
- establish the potential for newer implants to offer cost-effective gains over prostheses with a long track record
- indicate the optimum research methods for investigating crucial uncertainties
- identify the optimum design of future trial-based evaluation.

There is widespread confusion about the meaning of the term 'cost-effective'.[129] In the context of THR, since revision operations are markers for significant resource expenditure and adverse health outcome for patients (since a failure of a hip replacement results in pain and loss of mobility), a new prosthesis will be cost-effective if the total cost of the lifetime care package is less than for the standard prosthesis, or if the additional costs of the lifetime care package are worthwhile in terms of the additional benefits to patients of avoiding the suffering associated with a hip failure and consequent revision. Daellenbach and colleagues[130] and Gillespie and colleagues,[131] using the perspective of the New Zealand and Australian healthcare systems, respectively, explored what reduction in revision rate was necessary with a new prosthesis to generate the same overall expected health-service costs as established implants. However, this work did not allow for the fact that higher revision rates may have a significant impact on expected benefits to patients. Even if the new implants generate a higher overall cost of treatment than the established ones, they may still be considered cost-effective if their additional effectiveness is manifested in terms of avoided pain and disability, and if this is valued sufficiently by patients or by society.

Our analysis explored the potential for newer prostheses to represent a cost-effective use of NHS resources in primary THR. The long-term costs and benefits of standard primary THR were estimated using a Markov model based on one of the largest prospective studies of long-term THR survival in the UK.[132] These estimates were then used to represent the baseline against which the potential costs and benefits of new implants could be compared. The central question of the analysis was: given their higher acquisition price, how much more effective do the newer prostheses need to be in order to be considered cost-effective?

The Markov model

Firstly, a detailed description of the construction of the Markov model is given. This is followed by a description of how this model was used to analyse the potential cost-effectiveness of new prostheses.

The analysis was based on a simple Markov process, which is a form of decision analytic model used widely in health-services research,[133] and in economic evaluation in particular.[134] A Markov model involves dividing a patient's possible prognoses into a series of health states. The probabilities defining transition between each of these states are specified over a particular time frame (a 'cycle') such as a month or a year. With the aid of a computer, the model is run over a large number of cycles to see how a typical patient would move between states. Different probabilities are defined for each form of management under evaluation, and the costs and benefits of the comparators are estimated on the basis of the

length of time a hypothetical cohort of patients spends in each state.

The Markov model was used to predict the prognosis of patients who have undergone primary THR. Following the operation, patients were assumed to enter one of four distinct states in the Markov model.

- **Successful primary**: if patients survive the initial THR they move to this state.
- **Revision THR**: patients move to this state if their hip replacement fails (e.g. due to infection or loosening) and they then require revision surgery. As some patients require more than one revision operation, it is possible for a patient to move into this state more than once. Patients only remain in this state for one cycle.
- **Successful revision**: if patients survive revision surgery they progress to this state.
- **Death**: patients can die and enter this state at any point in the model. Patients can enter this state due to death related to surgery or due to the underlying risk of death.

A diagrammatic representation of the model is shown in *Figure 1*. Following the primary THR operation patients entered the model in the 'Successful primary' state if they survived the operation or the 'Death' state if they did not. The cycle length used in the model was 1 year. The model was run over a period of 60 years to estimate the lifetime costs and benefits of THR (this ensured that over 98% of patients in the youngest cohort analysed had died).

Three sets of parameters were used in the model: transition probabilities (which define the rate of transfer between the four states at each cycle); costs (incurred by patients in each state); and utilities (which define a value for the health-related quality of life [HRQL] experienced by patients in each state and which, together with life expectancy, can be used to define patient benefits). The sources of these data are described below.

Transition probabilities

As this review has already shown, there is a dearth of appropriate and useful long-term follow-up data on the effectiveness of hip prostheses. However, for the model presented in this paper to be of use, estimates of how patients progress through the four states were required. The starting point of the analysis was to estimate probabilities in order to model the long-term prognosis of patients undergoing primary THR using established prostheses, and the probabilities described below all relate to this baseline model.

A key probability within the model is the risk of failure of the prosthesis and consequent need for revision surgery. To estimate this risk of prosthesis failure, we used one of the few UK sources of data on the long-term outcomes of THR. The study was a prospective comparison of 1190 patients (mean age, 68 years; 71% female) who received either a Charnley prosthesis (n = 208) or a Stanmore prosthesis (n = 982) between 1973 and 1987.[132] Together, the Charnley and Stanmore prostheses represent up to 40% of the market in the UK,[2] and so this study represents a valuable source of data on THR revision rates over a period of 14 years for a large number of patients undergoing THR with

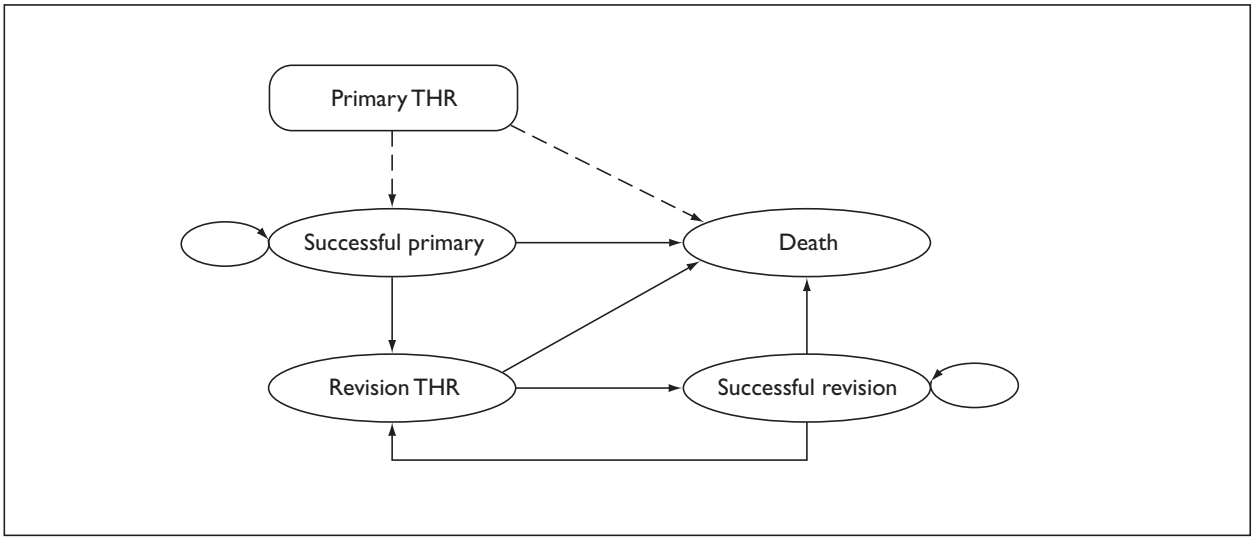


FIGURE 1 *The Markov model*

established prostheses in the NHS. For the model, the risk of prosthesis failure was estimated on the basis of the observed survival at 14 years in this data set, both for the whole patient group and for a series of patient sub-groups based on age and gender. After the risk of failure had been estimated on the basis of 14-years survival, the model was extrapolated to 60 years to estimate the lifetime costs and benefits of THR.

The predicted prosthesis survival curves from the model were compared with actual curves taken from the Swedish National Hip Arthroplasty Register,[127] one of the largest available sources of long-term follow-up of patients undergoing THR.

In addition to the probability of prosthesis failure following primary THR, a number of other transition probabilities were required for the baseline model of prognosis. The risk of death was considered both for risk of death associated with surgery and for underlying mortality risk. The former was assumed to be 1% for both primary and revision THR;[135] the latter was assumed to vary with age and was based on annual mortality rates taken from standard life tables for men and women.[136] The risk of re-revision was assumed to be 4% and constant over time. Although this risk level is slightly higher than the 2% rate used by Pynsent and colleagues in their study,[20] it is consistent with many of the point estimates of revision survival quoted by those authors.

Resource costs

The cost analysis was carried out from the perspective of the NHS, with a focus on the cost of primary and revision THR operations. To cost these two types of operation, estimates of resources used during the procedures were required. Estimates of length of stay in hospital and time in theatre were based on information provided by the Nuffield Orthopaedic Centre (NOC) NHS Trust in Oxford derived from costs for patients treated at that centre. Estimates of the numbers of out-patient attendances and X-rays were based on clinical opinion. The unit costs of each item of resource were based on data from the NOC at 1996–1997 prices, except for the prices of the prostheses and cement (which were based on manufacturers' list prices including VAT) and clinician and nurse salary rates (which were taken from the relevant review bodies' reports[137,138]). The cost of the established implant was based on the acquisition price of a standard Charnley prosthesis, although the price of a Stanmore prosthesis is very similar.[2]

The cost of a revision procedure included the acquisition cost of a long-stem Charnley prosthesis. *Table 6* details the resource use and cost estimates used in the analysis. Implant costs include both femoral and acetebular components.

It is standard practice in economic evaluation to discount costs that occur in future years,[139] and an annual rate of 6% per annum was used.[140]

Patient benefits

There are two components to the benefits patients experience from health care – improvements in life expectancy and in HRQL. Therefore, any measure of benefit in economic evaluation should ideally incorporate both of these elements. The measure of benefit used in the model described here was the quality adjusted life-year (QALY) which has been used widely in economic evaluation.[141] In addition to incorporating the impact of an intervention on length and HRQL, the QALY is a generic measure of benefit, which means that, in principle, it can be used in any area of health care. This assists in the process of resource allocation between programmes and disease areas. To estimate the number of QALYs generated by an intervention, a measure of life expectancy is weighted on an HRQL scale running between 0 (equivalent to death) and 1 (good health).

In the current analysis QALY estimates were generated as follows. Within the Markov model, there is a probability of death from any cause over time and, as the model was run for 60 years, it was straightforward to estimate the life expectancy of a given patient over that period. The quality adjustment of these life-years was undertaken on the basis of the degree of severity of pain patients would be likely to experience in given states in the model. On the basis of data from the UK study,[132] it was assumed that after a successful operation 80% of patients would experience no pain and 20% would have mild pain. On the basis of the same data,[132] it was assumed that for patients whose replacement hips fail, 15% would have experienced severe pain and 85% would have experienced moderate pain in the year before that in which revision surgery was carried out. It was assumed that in the year of a revision operation, all patients would experience a level of pain midway between severe and moderate.

The values, on the 0–1 scale, for the calculation of QALYs were taken from the paper by Laupacis and colleagues.[142] In that study, 188 patients undergoing THR were interviewed and presented with

TABLE 6 *Cost estimate for primary and revision THR*

Resource	No. of units		Unit cost (£)	Total cost (£)	
	Primary	Revision		Primary	Revision
Theatre overheads	134 (min)*	195 (min)*	4.89*	655	954
Theatre staff					
Consultant surgeon	134 (min)*	195 (min)*	0.496†	66	97
Consultant anaesthetist	134 (min)*	195 (min)*	0.496†	66	97
Registrar	134 (min)*	195 (min)*	0.229†	31	45
Grade F nurse	134 (min)*	195 (min)*	0.179	24	35
Grade F nurse	134 (min)*	195 (min)*	0.179	24	35
Grade E nurse	134 (min)*	195 (min)*	0.153	21	30
X-rays	6 (number)‡	6 (number)‡	22.41*	134	134
Stay in hospital	12 (days)*	14 (days)*	201.00	2412	2814
Out-patient visits	3 (number)‡	3 (number)‡	84.00*	252	252
Cost of procedure excluding consumables	–	–	–	3685	4493
Charnley prostheses	1 (number)	1 (number)	–	306¶	676¶
Cement	2 (packets)‡	4 (packets)‡	30.55§	61	122
Cost of procedure including consumables	–	–	–	4052	5291

Sources:
* *Nuffield Orthopaedic Centre NHS Trust.*
† *Review body for nursing staff, midwives, health visitors and professions allied to medicine (1996)[137]; Review body on doctors' and dentists' pay (1996)[138].*
‡ *Clinical opinion.*
¶ *De Puy (Leeds). Primary = standard Charnley; revision = long-stem Charnley.*
§ *CMW Laboratories (Exeter). 40 g packet of Gentamicin cement.*

descriptions of mild, moderate and severe osteoarthritis which focused on six areas of HRQL: pain and stiffness on exertion, use of walking aids, use of analgesics, pain at night, ability to do housework, and social life. The patients were asked to value each of these descriptions using an instrument known as the time trade-off.[143] Given the focus on pain in these descriptions of HRQL, these values were taken as being reasonable estimates of the value patients attach to severe, moderate and mild pain. Although this data set came from Canada, and so it was recognised that there could be a problem in applying it generally to a UK population, the study provided the best available data from which to derive values for the calculation of QALYs in the model. The health state values used to quality-adjust life expectancy on the basis of levels of pain were 0.69 for mild, 0.38 for moderate and 0.19 for severe. It was assumed that patients experiencing no pain had an HRQL valued at 1. In the same way as costs, benefits occurring in future years were discounted at a rate of 6% per annum.

Analysis

The starting point of the analysis was to model the prognosis of the typical patient following primary THR with an established prosthesis, and the estimates of transition probabilities, costs and benefits described above were used for that purpose. This model became a baseline against which to assess the potential impact of newer prostheses. The central question then addressed was: given that newer prostheses invariably have higher acquisition prices than established implants, how much more effective do they need to be to considered cost-effective?

Clearly, if the new prosthesis reduces health-service costs overall as well as having a lower revision rate,

it would be considered cost-effective. It is more difficult to be precise about when the new implant would be considered cost-effective if it has a higher overall health-service cost and a lower revision rate which generates more QALY benefit to patients. The question is whether the incremental cost per additional QALY is considered worth paying for, and this depends on whether a purchaser has additional resources and whether resources can be freed up elsewhere in the system.

Work undertaken in Canada suggested that a new technology would probably be considered cost-effective if it cost less than Can$20,000 per additional QALY, but that it was very unlikely to be considered cost-effective if it cost more than Can$100,000 per additional QALY.[144] No such tentative guidelines have been presented in the UK. However, the critical ratios are likely to be lower given that less money *per capita* is spent on health care in the UK.

To assess what revision rate new implants have to achieve to be considered cost-effective, a two-way threshold analysis was undertaken. In this, the baseline revision probability was varied by a factor x, which was assumed to be constant for each cycle in the model. Then the acquisition cost of the new prosthesis took on the value which achieved one of three critical thresholds:

- an expected overall health-service cost which was the same as that of the conventional prosthesis
- higher expected cost and higher expected QALYs with the new implant, with each additional QALY having an incremental cost of £6500
- higher expected cost and higher expected QALYs with the new implant, with each additional QALY having an incremental cost of £10,000.

This threshold analysis was then presented graphically in order to show, for a given percentage increase in the acquisition cost of a prosthesis (including cement), the percentage reduction in revision risk the new implant would have to achieve over the established prosthesis if it was to be considered cost-effective in terms of the three thresholds above.

Chapter 6

Modelling of cost-effectiveness of THR: results

Baseline revision risk functions

Using the data from the prospective UK study,[132] three approaches were taken to estimate the failure risk (or hazard): the first assumed that the hazard was constant, whereas the second and third assumed a hazard increasing in a linear or quadratic fashion, respectively. The linear increasing hazard is estimated by

$$h = bt$$

where h is the hazard, t is time in years (i.e. model cycle number) and b is the coefficient of increase. Similarly, a quadratic hazard function was estimated by

$$h = ct^2$$

where c is the coefficient of increase.

Figure 2 shows the estimated revision-free survival derived from each of these three approaches in comparison with the original revision-free survival data from which they were estimated. The figure clearly shows that the linear increasing hazard produces a revision-free survival curve which is more consistent with the observed revision-free survival and hence this assumption has been used in the analysis. *Figure 2* illustrates the predicted revision-free survival curves for hazards estimated on the basis of the full data set, but the linear hazard also produced the best fit for the sub-group analyses by age and gender. The revision-free survival for the sub-groups predicted on the basis of a linear increasing hazard is shown, together with the original data points, in *Figures 3* and *4*.

Clearly, a good model will predict the data from which it was constructed and *Figures 2, 3* and *4* show that the model does indeed quite closely predict the revision-free survival observed in the UK prospective study. However, models should also be validated by comparing their predictions with data not used to construct the model. The predicted revision-free survivals for each cohort after 14 years

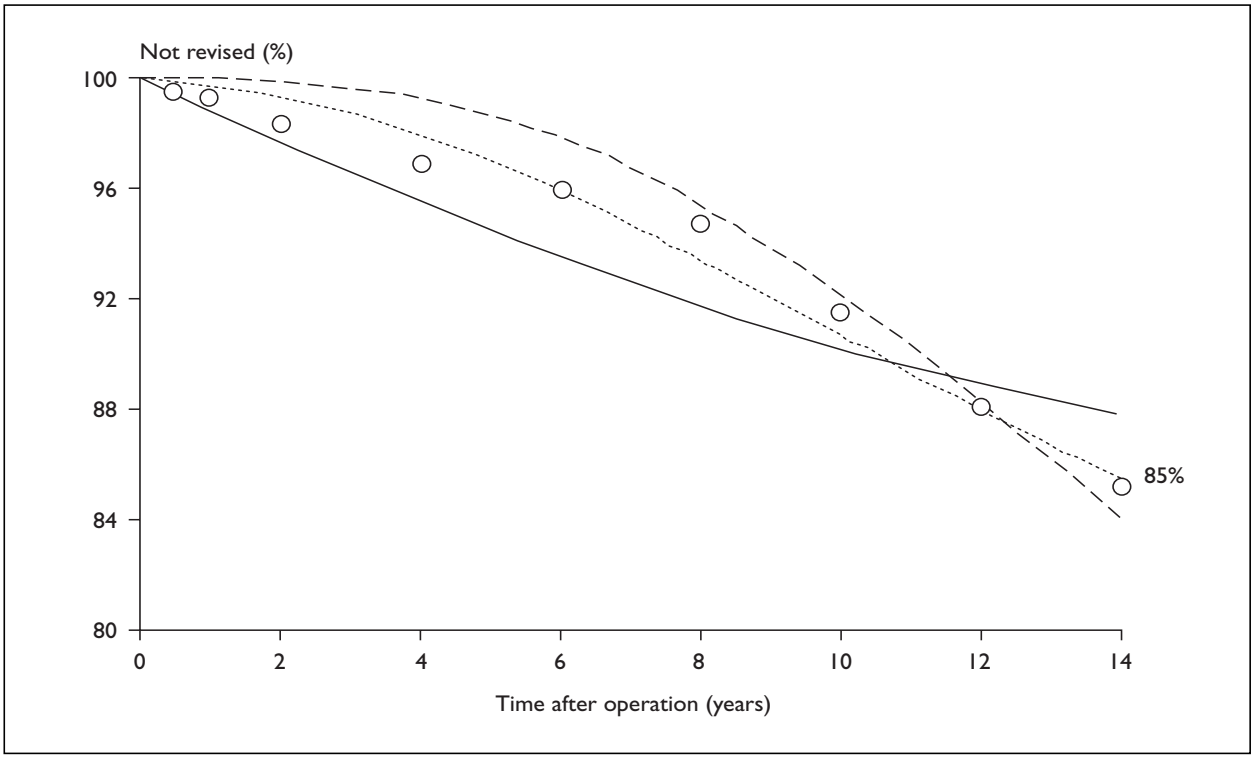
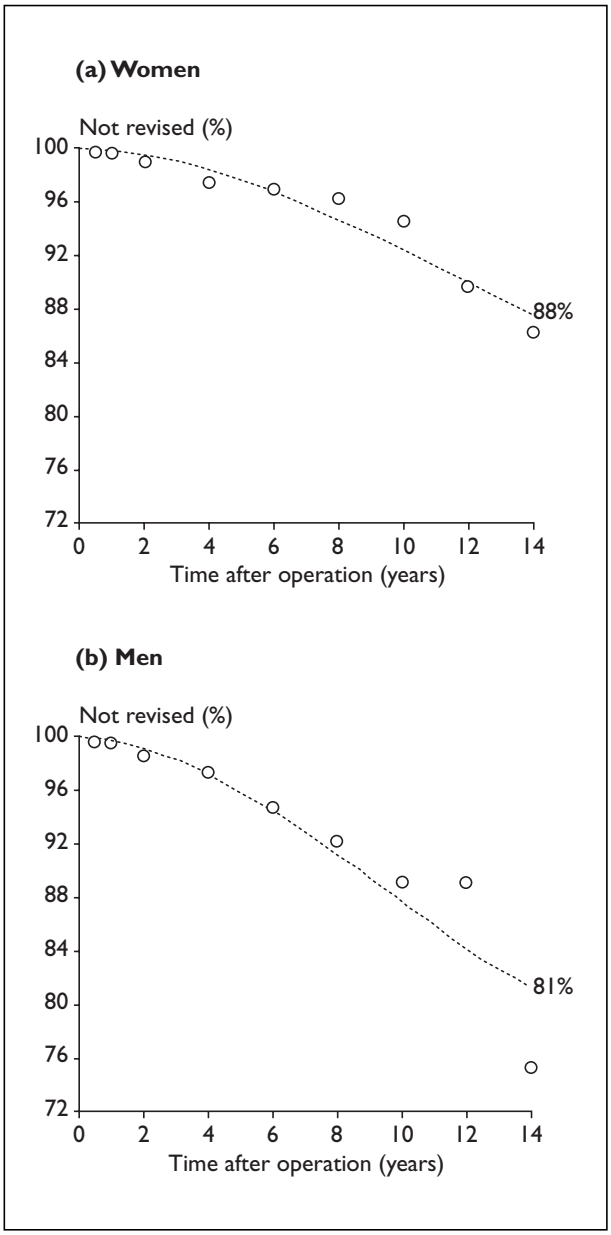
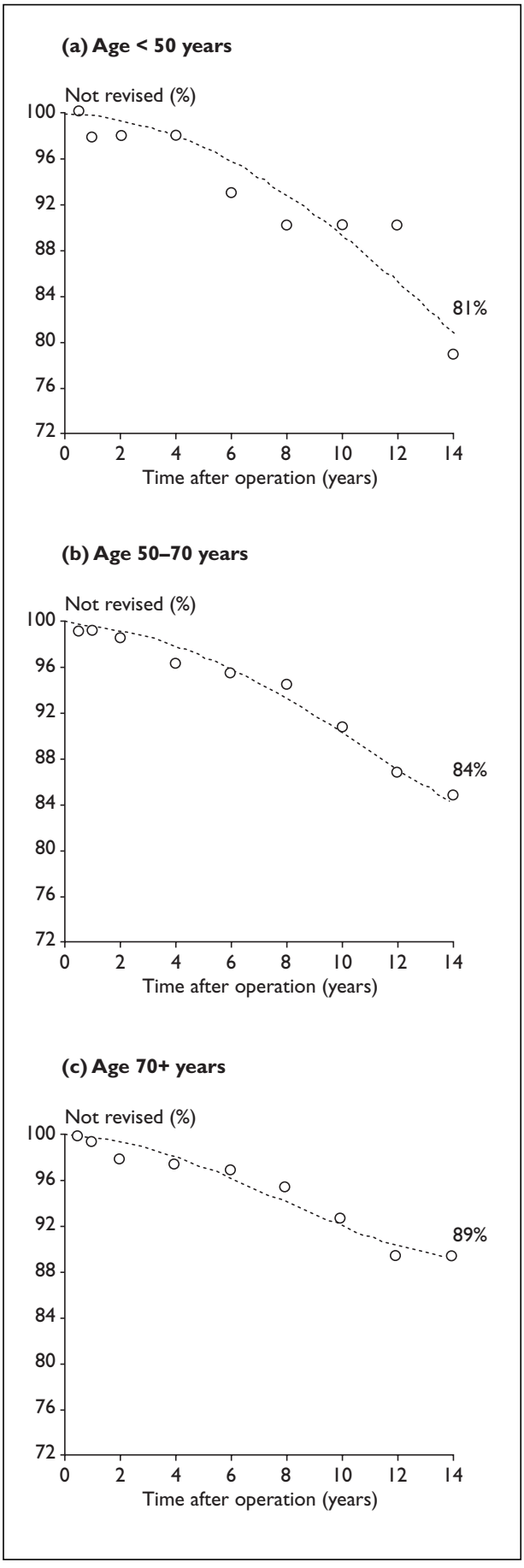


FIGURE 2 *Three assumptions for estimating revision risk (○, observed survival; ——, constant hazard; ....., linear hazard; – – –, quadratic hazard)*



ABOVE: FIGURE 3 *Predicted versus observed revision-free survival for (a) women and (b) men (○, observed survival; ....., predicted survival)*



RIGHT: FIGURE 4 *Predicted versus observed revision-free survival for three different age groups (○, observed survival; ....., predicted survival)*

shown in *Figures 3* and *4* have therefore been compared with 14-year revision rates for patients undergoing THR as recorded in the Swedish registry. The Swedish registry records 89% of women and 85% of men as being free of revision at 14 years, in comparison with 88% of women and 81% of men predicted by the model. Comparisons between the model data and the Swedish registry data are less straightforward for age sub-groups

because different categories were used to describe the Swedish data and the Swedish data are split by diagnosis. Data at 14 years are available for patients with osteoarthritis in the Swedish registry: 84% of patients aged < 55 years, 81% of patients aged 55–64 years, 86% of patients aged 65–74 years, and 93% of those aged > 75 years had not had a revision at 14 years. The model predictions of revision-free survival are 81% for patients aged < 50 years (mean age, 40 years), 84% for those aged 50–70 years (mean age, 63 years), and 89% for those aged > 70 years (mean age, 76 years). On balance, the model predictions do not seem at all inconsistent with the Swedish data at 14 years. Having shown that the model gives sensible predictions at 14 years in comparison with available data, we have extrapolated to 60 years to show the predicted revision-free survival. The results of this extrapolation are shown in *Figure 5*.

Cost and benefit of the baseline prosthesis

The first stage in the analysis of the potential cost-effectiveness of new prostheses is to estimate the cost and benefit associated with standard prostheses (that is, the baseline cost and benefit). If the estimated linear increasing hazard function is incorporated into the Markov model together with the other probability and cost data detailed above, estimates of the cumulative 60-year expected cost and QALYs of THR can be calculated. These results are shown in *Table 7,* and are detailed for the various sub-groups of patient. Older patients undergoing a primary THR will have a lower expected cost because they are less likely to require a revision. This is because they may die before a revision is necessary and they are likely to be less mobile thus putting less pressure on a prosthesis. Conversely, younger patients accrue greater long-term costs because they are likely to live longer and to have a more active life, thus requiring more revisions of THR.

Costs and benefits of the new prostheses

A two-way threshold analysis relating to an average patient cohort has been used to explore the question of how much more effective a new prosthesis needs to be, relative to the baseline, to be

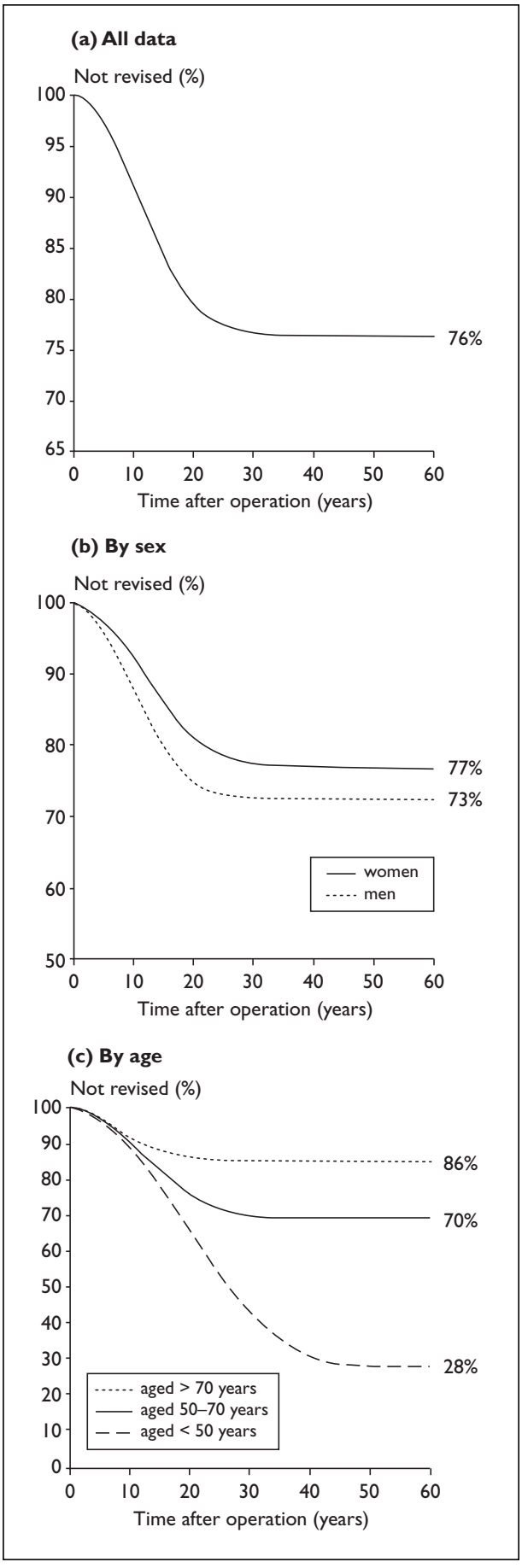


FIGURE 5 *Long-term revision-free survival predicted (a) from all data, (b) for men and women and (c) by age at THR*

TABLE 7 *The baseline results of the model: estimated costs, life expectancy and QALYs of establish prostheses over 60 years for various sub-groups*

Patient group	Expected costs (£) (Discounted at 6%)	Life expectancy (years) Discounted at 6%	Life expectancy (years) Undiscounted	Expected QALYs Discounted at 6%	Expected QALYs Undiscounted
All patients (mean age 68 years; 71% female)	4804	9.19	14.37	8.39	13.12
Women (mean age 67 years; Stanmore only)	4746	9.89	16.15	9.06	14.79
Men (mean age 66 years; Stanmore only)	4963	8.69	13.25	7.9	12.04
Age < 50 at primary THR (mean age 40 years; 63% female)	5799	14.87	36.70	13.57	33.32
Age 50–70 years at primary THR (mean age 63 years; 66% female)	4951	10.28	17.31	9.4	15.80
Age > 70 years at primary THR (mean age 76 years; 73% female)	4535	9.79	6.99	8.93	6.37

considered cost-effective given its additional acquisition cost. The results of this analysis are shown in *Figure 6*. The horizontal axis of the graph shows the effectiveness of the new prosthesis in terms of the percentage reduction in revision rate relative to the baseline established prosthesis. The vertical axis shows the acquisition cost of the new prosthesis as a percentage factor of the acquisition cost of established prostheses, which is taken as the cost of the standard Charnley. The three functions shown in the graph represent the three critical thresholds discussed above:

1. the point at which the overall health-service cost of the new prosthesis is equal to the baseline cost despite its higher acquisition cost (cost neutral)
2. the point at which the new prosthesis has a higher overall health-service cost and each extra QALY it generates costs £6500
3. the point at which the new prosthesis has a higher overall health-service cost and each extra QALY it generates costs £10,000.

Given the choice of one of these thresholds to define cost-effectiveness, points to its north-west show combinations of added acquisition cost and lower revision rate that would be insufficient for the new prosthesis to be considered cost-effective. Conversely, points to the south-east of the relevant threshold represent combinations of additional acquisition cost and reduced revision rate that would be sufficient for the new implant to be cost-effective.

Figure 6 can be used to estimate revision rates required for any new prosthesis to be cost-effective. Two scenarios can be used to illustrate this approach. Firstly we can consider a new **cementless** prosthesis approximately 300% more costly than the Charnley. Secondly we can consider a new **cemented** prosthesis with an acquisition cost of 150% of the Charnley. (These relative costs are typical of newer cementless and cemented prostheses.) From *Figure 6* it is clear that the new cementless prosthesis is very unlikely to be cost neutral, however much it reduces the revision rate. However it might still be considered cost-effective if reductions in the revision rate are greater than about 50% of the baseline (for a threshold cost per QALY of £6500) or 40% of the baseline (for a threshold cost per QALY of £10,000). By contrast, the new cemented prosthesis would need to show only a 30% reduction in the revision rate compared with a standard prosthesis to be cost neutral, a 15% reduction for a cost per QALY of no more than £6500, and just over a 10% reduction if its cost-effectiveness is to be less than £10,000 per QALY.

The results of the threshold analyses of the two hypothetical new prostheses relative to the standard Charnley prosthesis are shown for the patient

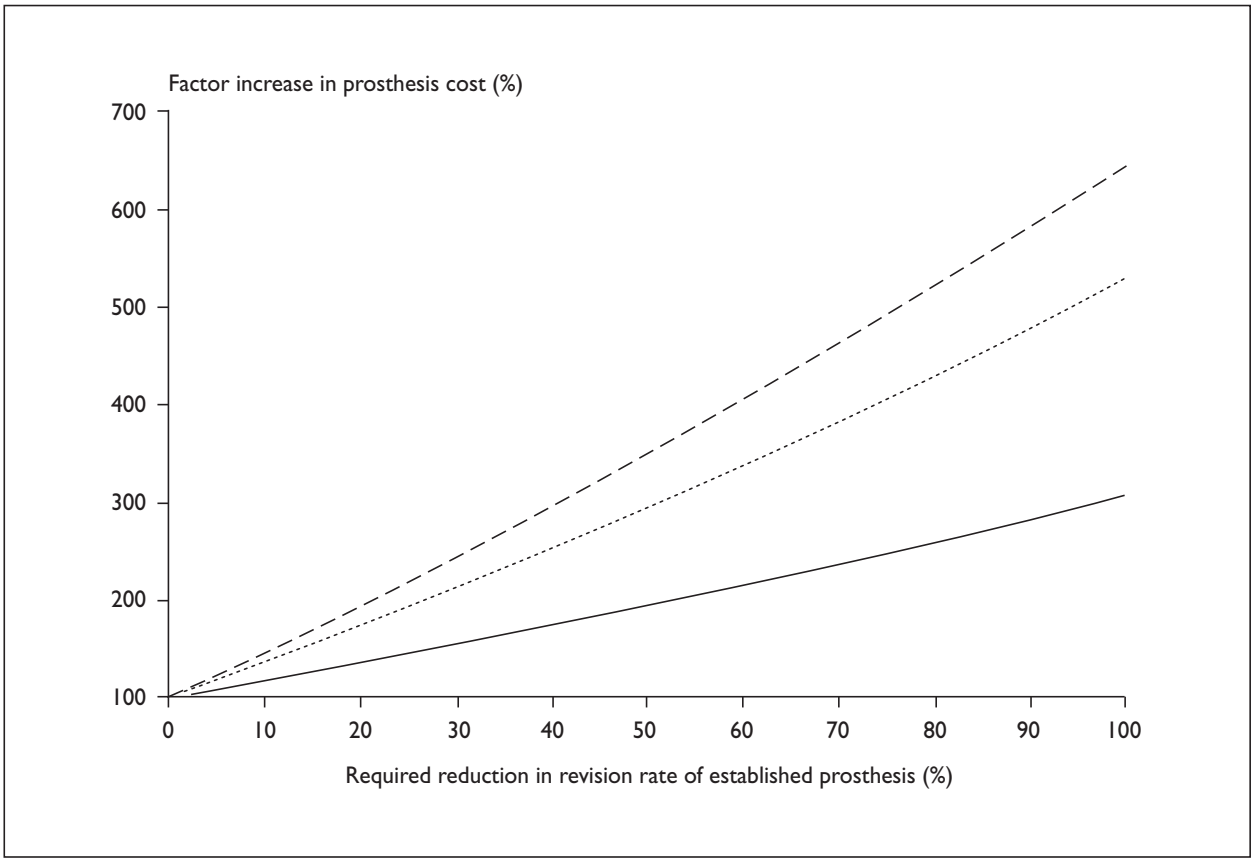


FIGURE 6 *Two–way threshold analysis* to indicate how more effective (in terms of reducing the revision rate) a new prosthesis has to be to justify its additional aquisition cost (——, cost-neutral threshold; ·····, £6500 per QALY threshold; – – –, £10,000 per QALY threshold)*

* *The analysis relates to an average cohort of whom 60% are women and the mean age is 65 years.*

sub-groups in *Tables 8* and *9*. The tables detail the percentage reduction in revision rate that is needed to achieve cost-effectiveness in each of the key sub-groups. For those sub-groups that experience a lower revision risk – in particular older patients compared with younger ones, and women compared with men – a greater percentage reduction in revision rate is required for a new prosthesis to be considered cost-effective. This is because for these sub-groups the **scope** for improvement from new technologies is less.

Sensitivity analysis

Any economic modelling of the sort presented here is subject to uncertainty in many of its parameters. Sensitivity analysis is used to assess the extent to which the principal results of the analysis are robust to changes in uncertain parameters.[145] As a starting point for the sensitivity analysis, it is assumed that a new prosthesis is 300% or 150% more costly than a standard prosthesis and that the reduction in revision rate is 51% or 14%, respectively, so that both the new cementless prosthesis and the new cemented prosthesis have a cost-effectiveness ratio of £6500 per QALY. For each of the parameters of the model, an elasticity is calculated which measures the impact of a change in the parameter on the cost-effectiveness ratio. The elasticity is given by

$$\varepsilon_y = \frac{\Delta y / y}{\Delta z / z}$$

where y is the input parameter and z is the outcome of interest.

Table 10 presents the baseline value of each of the parameters, together with their elasticities, for the model based on data for all patients. The greater the elasticity, the more sensitive the cost-effectiveness ratio to changes in that parameter. Elasticities greater than 1 indicate that a percentage change in the input parameter results in a greater percentage change in the cost-effectiveness ratio.

TABLE 8 *The maximum revision risk that a new prosthesis, costing 300%[*] more than the standard Charnley, has to achieve to be considered cost-effective[†]*

	Reduction required in revision rate relative to established prosthesis		
Patient sub-group	**Cost neutral**	**£6500 per extra QALY**	**£10,000 per extra QALY**
All patients (mean age 68 years; 71% female)	98%	51%	41%
Women (mean age 67 years; Stanmore only)	Undefined[‡]	54%	43%
Men (mean age 66 years; Stanmore only)	84%	44%	35%
Age < 50 at primary THR (mean age 40 years; 63% female)	53%	27%	21%
Age 50–70 years at primary THR (mean age 63 years; 66% female)	85%	44%	35%
Age > 70 years at primary THR (mean age 76 years; 73% female)	Undefined	76%	61%

[*] *Approximately equivalent to the cost of many newer cementless prostheses.*[2]
[†] *Values are percentage reductions in the baseline revision risk of the established prosthesis.*
[‡] *Undefined = it is impossible for a prosthesis to be cost-effective in this context.*

TABLE 9 *The maximum revision risk that a new prosthesis, costing 150%[*] more than the standard Charnley, has to achieve to be considered cost-effective[†]*

	Reduction required in revision rate relative to established prosthesis		
Patient sub-group	**Cost neutral**	**£6500 per extra QALY**	**£10,000 per extra QALY**
All patients (mean age 68 years; 71% female)	28%	14%	11%
Women (mean age 67 years; Stanmore only)	30%	14%	11%
Men (mean age 66 years; Stanmore only)	24%	12%	9%
Age < 50 at primary THR (mean age 40 years; 63% female)	15%	7%	6%
Age 50–70 years at primary THR (mean age 63 years; 66% female)	24%	12%	9%
Age > 70 years at primary THR (mean age 76 years; 73% female)	41%	20%	16%

[*] *Approximately equivalent to the cost of many newer cemented prostheses.*[2]
[†] *Values are percentage reductions in the baseline revision risk of the established prosthesis.*

TABLE 10 *Illustrative sensitivity analysis for both a typical new cementless prosthesis or a typical cemented prosthesis**

Parameter description	Baseline value	Elasticity†
Mortality rate associated with surgery		
Primary THR	1%	0.02
Revision THR	1%	–0.06
Revision rate (constant hazard)	4%	–0.31
Cost of procedure		
Primary THR	£3686	0
Revision THR	£5290	–0.92
Quality of life weight for:		
'no pain'	1.00	–0.96
'slight pain'	0.69	–0.17
'moderate pain'	0.38	0.03
'severe pain'	0.19	0.09
Pain		
Proportion of patients in no vs. slight pain following successful operation	80%	–0.30
Proportion of patients in moderate vs. severe pain following hip failure	15%	0.02
Typical new cementless prosthesis		
Additional cost of new prosthesis	300%	2.07
Revision risk reduction of new prosthesis	51%	–2.88
Typical new cemented prosthesis		
Additional cost of new prosthesis	150%	5.76
Revision risk reduction of new prosthesis	14%	–1.94

* *Note from* Tables 8 *and* 9 *that this gives a cost per QALY ratio of £6500.*
† *Calculated as the percentage change in outcome (cost per QALY ratio) over percentage change in input parameter. Quoted figures based on 1% change in input parameter.*

It is clear from *Table 10* that the two most important parameters in the model are the cost and the effectiveness of a new prosthesis relative to the standard prosthesis. Most of the remaining parameters have low elasticities, indicating that they are unlikely to be key parameters in terms of determining cost-effectiveness. The two exceptions are the cost of the revision operation and the quality-of-life weights for no pain, both of which have an elasticity approaching 1. Note that the cost of the primary operation (excluding the cost of the prosthesis itself) has an elasticity of 0: this is because this cost is accrued for all patients irrespective of the prosthesis they have fitted. (Of course, this result may change if a new prosthesis significantly alters the time required for an operation.)

It is acknowledged that there may be substantial variations in lengths of stay and other important cost-generating events between hospital units and also changes over time in clinical practice, such as greater use of bone grafting, which may affect the structure of costs in ways that sensitivity analysis in preliminary modelling cannot fully address.

Chapter 7

Discussion of economic modelling

Economic evaluations in the literature

There have been very few formal economic evaluations of THR. Liang and colleagues[146] undertook a 'before and after' study of 45 consecutive osteoarthritis patients undergoing total joint replacement. They found that HRQL improved appreciably by 6 months after surgery, but that costs did not fall greatly compared with before surgery. However, the absence of a genuine control group, the relatively short follow-up, and the absence of comparable cost-effectiveness estimates for other procedures preclude a clear conclusion on the cost-effectiveness of THR from this study.

A Canadian group[147,148] report the results of an economic evaluation undertaken alongside a randomised trial which involved 250 osteoarthritis patients and compared cemented and uncemented prostheses for THR. The study collected resource use and health-state valuation data directly from patients. One year after surgery, the mean cost per patient was Can$975 in the group in which the cemented prosthesis was used and Can$1297 in the group fitted with the uncemented prosthesis (1988 prices). However, the study found no statistically significant difference in health state values between the two groups, which suggests that the cemented prosthesis is more cost-effective. The authors also compared pre-surgery and post-surgery cost and health-state value data from their study and estimated that the incremental cost per QALY of THR (compared with no THR) was between Can$8031 and Can$27,139, depending on assumptions about the duration of benefit. They concluded that, compared with other uses of health-care resources, THR is highly cost-effective. However, the study was based on a very short follow-up.

Perhaps the most rigorous economic evaluation of THR was undertaken by Chang and colleagues[149] who used a detailed decision analytic model and data taken systematically from the published literature to estimate the lifetime costs and QALYs per patient of THR compared with non-surgical management. They found that THR is less likely to be considered cost-effective in men and as age increases. The model suggested that, for a 60-year-old woman with functionally significant but not dependent hip osteoarthritis, THR would actually save money. In a man aged 85 years or older, the incremental cost per additional QALY of THR was estimated at US$4600 (1991 prices).

Possible economic implications of the current analysis

The purpose of our analysis is not to estimate the cost-effectiveness of THR *per se*. Rather, it is to explore some of the economic implications of the growth in the number of new implants available in the UK. Given that these implants invariably have higher acquisition costs than established prostheses, it is sensible to ask how much more effective they need to be, in terms of lower revision rates, to be considered good value for money.

The analysis presented here shows that new prostheses costing approximately 300% more than established ones will probably be unable to generate reductions in THR revision risk that are sufficient to produce overall cost savings. If purchasers are willing to incur additional costs in return for extra benefits to patients from more expensive prostheses, then the likelihood of a new prosthesis being considered cost-effective will depend partly on the sub-group of patient being considered.

Suppose, for example, that a purchaser is prepared to fund as cost-effective a new prosthesis that costs more overall than established prostheses and has an incremental cost per additional QALY of no more than £6500. As shown in *Table 8,* to meet this requirement a new prosthesis that costs 300% more than the standard Charnley prosthesis will have to generate a 27% reduction in the revision rate achieved by established prostheses in patients aged < 50 years, but a reduction of 76% is needed for patients aged 70 years and older. This is partly because many elderly patients undergoing primary THR will die well before a revision is necessary. In contrast, for a new prosthesis that costs 150% more than the standard Charnley prosthesis (*Table 9*), only a 12% reduction in the revision rate for older patients (aged 50–70 years) is required for the prosthesis to be cost-effective. For younger patients (aged < 50 years) the required reduction in revision rate is only 7%. Such levels of reduction are far more plausible.

This sort of analysis can feed into healthcare decision making at various levels. If a new prosthesis has to halve the revision rate achieved by standard implants to prove cost-effective, clinicians and purchasers may consider this unfeasible and continue with the established prosthesis. Conversely, if the required reduction is much smaller – say 5% – and, based on experience, the new implant seems to have achieved this improvement then they may consider its use to be justified.

Preliminary modelling can also be of value in planning research and development. Given the dearth of long-term survival data on the newer prostheses,[2,120] but the evidence that these implants are being used in the NHS,[17] there may be an inclination on the part of funders of research and development to support trials to establish whether the newer implants are more effective. However, if preliminary modelling indicates that a new prosthesis would have to cut revision rates by 50% to prove cost-effective, it may be concluded that such a trial would be a waste of limited research resources because such a difference simply is not feasible. Alternatively, if only a 5% reduction in revision rates is required to achieve value for money, a trial might be considered worthwhile and the threshold estimate would assist in the calculation of the required sample size. The modelling in *Table 9* suggests that, for example, a quite small (7%) reduction in revision rate is all that is required for a new cemented prosthesis costing 150% of the standard Charnley prosthesis in the case of younger patients. A trial to investigate such improvements could well prove worthwhile.

The implant manufacturers, too, can be usefully informed by preliminary economic modelling. In setting the price of a new prosthesis, the manufacturer has control of a major determinant of the prosthesis' cost-effectiveness. If an estimate has been made of the improvement in revision rate that is feasible with a new implant, the modelling will indicate at what price the prosthesis would be considered cost-effective while making a reasonable return on investment. If manufacturers had adopted this sort of approach in the past, it is likely that many of the new implants would never have reached the market.

A major difficulty with preliminary economic modelling is that whereas it is easy to define cost-effectiveness in general terms it is difficult in practice. If a new implant is only considered cost-effective if it is cost neutral relative to the established prosthesis, models using threshold analysis would be straightforward, and this is probably why earlier work of this type on THR has focused on this definition of cost-effectiveness.[130,131] However, it is quite feasible for a new implant to be more expensive overall and still to be more cost-effective than established techniques, as long as it generates more benefits to patients, and those benefits are valued sufficiently.

In principle, the use of QALYs in economic analysis allows the benefits of all healthcare interventions to be expressed in common units and to be related to their costs. Hence, the concept of ranking the uses of healthcare resources according to their incremental costs per additional QALY has been suggested as a tool for purchasers.[150] In economic evaluation a major area of uncertainty is where in that ranking to set the cut-off point below which an intervention is not considered cost-effective. In practice, this threshold will vary between purchasers and over time. The two thresholds used here (£6500 and £10,000) are, therefore, indicative only.

The analysis presented here indicates that the price of a prosthesis represents an important element of the overall cost of THR. Many of the newer implants, especially the cementless ones, cost over £1000.[2] Even allowing for the possible saving in the cost of cement, this represents at least a 300% higher cost than that of an established prosthesis. With this additional cost, it is unlikely that new implants will ever save money compared with established prostheses because to do so they would have to reduce revision rates by about 90% in comparison with established prostheses. If purchasers are willing to accept an increase in costs as long as additional benefits for patients are generated, to be cost-effective new implants with this sort of additional acquisition cost will have to reduce revision rates to between 50% and 60% of that of established implants. If this level of reduction is considered feasible, trials should be undertaken to confirm it. If not, the results of economic modelling suggest that the role of the newer implants in the NHS should be questioned.

Chapter 8

Conclusions and recommendations for further research

Conclusions

This review has sought to identify from an extended search of available published research those studies which provide the clearest and least biased evidence of the relative performance of different prostheses used for THR in the NHS. The evidence only relates to primary THR and provides no clear inferences for revision THR, in which the issues of bone loss and the emergence of bone impaction grafting raise different questions of cost-effectiveness of prostheses.

Although 11 RCTs were found that addressed the issue, these trials did not provide a firm basis for any conclusions about the relative performance of different prostheses. It was not possible to perform any kind of meta-analysis of the data from RCTs for the simple reason that no two trials examined the same set of prostheses. Moreover, in most trials the sample size was small and patients had not been followed up for the length of time required for failure of the prosthesis to occur with any frequency. Two RCTs were found with somewhat larger sample sizes and longer periods of follow-up. Garellick and colleagues[37] randomised 410 patients to receive either a Charnley or a Spectron prosthesis and have so far reported the results over 5.6 years. Marston and colleagues[12] randomised 413 patients to receive either a Charnley or a Stanmore prosthesis and have followed up patients for 6.5 years so far. To date publications from neither research group have reported any significant differences between prostheses. However, the follow-up periods may be rather too short for problems yet to have arisen. In general, the small amount of data from RCTs provides no clear evidence regarding relative medium to longer term performance of prostheses.

There is continual innovation in basic biomedical and bio-engineering science in relation to THR. This will continue to provide new and important understanding. However there is no immediate prospect of major changes in the knowledge base of THR, in the sense that the authors of this review are not aware of major RCTs currently underway that will transform our understanding of the relative cost-effectiveness of existing prostheses.

Because of the extremely limited nature of the evidence from RCTs, particular attention was given to the evidence provided by 18 studies which had in common the features that they were comparative and concurrent (i.e. studying more than one prosthesis being used over a common period). It is not easy to estimate the extent to which data from such non-randomised studies address problems of bias and confounding arising from other variables relating to patient, surgeon, or surgical centre that may be responsible for differences between prostheses.

A major limitation in the available literature generally was the poor level of outcome assessment. Substantial risks of bias arise from reliance on evidence such as rate of revision surgery or surgical opinion of outcome. Above all, only exceptionally has the literature on THR provided evidence of outcomes as assessed by patients. This is a fundamental flaw.

In terms of numbers of patients observed, the evidence from comparative observational studies is very largely dominated by the Swedish National Hip Arthroplasty Register initiated in 1979.[6] This register provides about 80% of all of the patients observed in comparative observational studies included in this review. The most recent report of this Register, which appeared after we had carried out the analyses for the current review, is based on an even larger sample of over 130,000 primary THR operations.[127] The next most important contribution is the Norwegian Register[52,53] which provides about 14% of patients reported in the comparative observational studies found for this review. In the absence of convincing evidence from RCTs the Swedish and Norwegian registers provide the most useful evidence that could be found of the relative performance of different prostheses. The Swedish Register in particular has the merits of a large sample size, long periods of follow-up, and the likelihood that results generalise well to all THR in Sweden. As well as providing evidence of the comparative performance of prostheses, the Register has been used to show the impact of types of cement, surgical method and hospital on revision rates.[127] Such observational data have

to be used to estimate likely effects when no better data exist.[126]

For exploratory purposes in this review, the evidence available from all observational as well as randomised studies were finally combined to provide an overall estimate of the failure rates of different prostheses and the scale of differences between prostheses. It is quite definitely not intended that this meta-analysis should represent an authoritative overall estimate of the relative performance of different prostheses for THR available in the NHS. The analysis combines data gathered by studies with very diverse methodologies, and with uneven standards of reporting of results. There are large differences between studies in patient characteristics, and in the periods observed (and therefore in the levels of expertise of surgeons in the use of particular prostheses). There were very substantial problems of comparing outcomes across studies and assumptions had to be made in order to standardise outcomes in the meta-analysis. Although combining data in this way had the effect of increasing the sample size for estimating failure rates for different prostheses, the extent to which a variety of possible biases may influence such results is impossible to estimate. This exploratory meta-analysis represents – rather like the economic modelling also reported in this study – an attempt to estimate the possible scale of differences in performance of prostheses, but **it cannot be considered a precise estimate of the relative performance of any single prosthesis**.

The Charnley prosthesis has been available the longest. It is by far the most frequently studied of prostheses and has data for longer periods of follow-up based on a much larger number of patients than any other prosthesis. Data from the Swedish Register[6] that were available for this review showed it performed very well. This has been confirmed in more recent (1996) results for the Register which show no significant differences in performance after 15 years between the three most successful prostheses: the Charnley, Lubinus IP and CAD.[127] Amongst studies of any size only in the Norwegian Register is there evidence of a prosthesis performing clearly better than the Charnley.[52] In that study a significantly more favourable survival rate is reported for the Spectron/ITH in this study, but the results are based on just 2.7 years of follow-up.

Given the weakness of all evidence identified for the current review, the relative success of the Charnley prosthesis has to be expressed tentatively. Nevertheless it can be said that the Charnley is the most investigated of all prostheses, and the evidence on it is consistently favourable, with no other prosthesis clearly performing consistently more favourably. Unfortunately an important limitation on this statement is the fact that the design of the Charnley prosthesis – like that of other prostheses – has changed over the years of its use. Much of the evidence assessed in this review therefore does not necessarily apply to the Charnley prosthesis currently in use in the NHS.

If we are cautious in our expression of how positive the evidence is for the Charnley prosthesis, statements of the relative success of other prostheses have to be even more guarded. The Lubinus has performed indistinguishably from the Charnley in the Swedish Register, but it is not widely used in the NHS. It is a matter of curiosity that it has not been adopted more often by UK orthopaedic surgeons given the relatively favourable evidence available. From the meta-analysis, the Exeter appeared to have a relatively favourable revision rate. In the most recent data from the Swedish register,[127] the Exeter prosthesis with polished finish has a more favourable revision rate than the Charnley prosthesis. Overall, there is far less observational evidence for the Exeter and it has not been included in any RCTs found for this review. Thus there is evidence suggesting that Exeter polished may, along with Charnley, be associated with favourable long term outcomes, but the extent of the evidence is not so substantial.

There is little substantial evidence available for the Stanmore prosthesis. The meta-analysis suggested that it may be associated with a revision rate not very markedly higher than that of the Charnley or Exeter prosthesis. A recent overview concluded more emphatically in favour of the Stanmore.[122] This was in part influenced by the evidence of the RCT of Marston and colleagues[12] to which reference has already been made. To the extent that the Charnley is a 'gold standard' for comparative purposes, the inability of Marston and colleagues to detect any difference between the Charnley and Stanmore prostheses may be considered favourable evidence for the latter.

At the other extreme, the available evidence – almost entirely observational – has occasionally identified prostheses with a particularly poor record. The Christiansen fell out of use because of very high revision rates observed in the Swedish national register.

More generally, non-cemented prostheses are less commonly used in all the countries for which

there is evidence. However they are particularly used for younger patients, who may in the future come forward for THR in growing numbers. Amongst non-cemented prostheses, HA-coated and hemispheric prostheses seem to perform more favourably, although their relatively recent introduction means shorter observation of potential problems. Some threaded metal-backed prostheses (e.g. Ti-Fit) were found to perform less well than other cementless prostheses in the analysis of the Norwegian register's data set of cementless prostheses.[53] This is consistent with the evidence from the meta-analysis by Yahiro and colleagues[121] which compared revision rates of threaded acetabular cups with cemented and porous ingrowth acetabular cups and found significantly higher revision rates for threaded cups.

For the majority of the prostheses introduced into the NHS, the evidence from the present review is entirely consistent with previously expressed judgements,[2,122] in concluding that there is little or no systematic evidence about outcomes to support their adoption by orthopaedic surgeons. Recently introduced prostheses usually cost considerably more than established prostheses. The price of the Charnley was quoted as £282 in a survey of manufacturers published in 1995.[2] The same source reported that, of 24 cementless prostheses, 16 cost more than £1000 each. The results of the economic modelling carried out in this review (which must be regarded with caution because of the assumptions made and evidence used) suggested that more expensive prostheses, such as the majority of cementless prostheses, would need to be associated with substantial improvements in the overall revision rate to justify their higher prices. It is quite clear that there is no evidence for such improvements in outcomes of more recently adopted prostheses. Indeed, the total costs to society associated with poorer results of THR using some newer prostheses can be high.[151] The scope for achieving substantial improvements in older patients that justify the higher costs of prostheses is not great because of the reduced life expectancy of patients. It might also be argued that even in younger patients the 27% reduction in revision rates required by the modelling reported in this review is not very plausible.

Implications of the findings

Use of prostheses in the NHS

It might be argued that the range of prostheses made available for use in the NHS should be restricted to the small number (two or three) prostheses identified by this review as having relatively favourable supportive evidence. Surgeons would have very restricted ability to use other prostheses. In our view, the evidence available does not justify such a severe restriction because it does not unambiguously identify the prostheses that should be selected. Surgeons also need to be comfortable with the materials and techniques that they use. Therefore although this review has concluded that a small number of well-established prostheses such as the Charnley have a somewhat firmer evidence base, it would not be appropriate, given the current limited state of knowledge, to conclude that the use of other prostheses should be discontinued. On the basis of the data found for this review it would seem appropriate to suggest that orthopaedic surgeons should consider the rationale for selection of prostheses very carefully if they prefer to use prostheses for which there is no strong evidence base.

Equally, it does seem consistent with the evidence provided here to argue that the greater the costs of prostheses, the greater the onus should be on manufacturers and other parties concerned to justify those costs in terms of evidence of cost-effectiveness before a prosthesis is adopted for use in the NHS.

Recommendations for research

It has been argued that national prospective registers or dedicated trials are needed to provide independent evidence to purchasers, clinicians and patients of the relative performance of prostheses.[1,18] The findings of the present review support such arguments. There are very few independent data with which to evaluate the new prostheses which are being increasingly used in the NHS. With minor exceptions, for this report we have had to resort to the registers of Sweden and, to a lesser extent, Norway to obtain evidence to inform the review. This is far from satisfactory, not only because the profile of use of prostheses in those countries differs from that in the UK, but also because of the unknown scale of differences in surgical practice between countries. A UK register of prostheses would clearly have greater power to distinguish levels of performance of prostheses, and also to identify other factors influencing outcomes, because a much larger population is provided for by the NHS than is possible in Sweden and Norway. A UK register could be in a number of formats, ranging from a minimal register of THRs that identified the type of prosthesis used, through

to a format with some longitudinal monitoring of major failure in the form of revision surgery and, in the most ambitious form, longitudinal monitoring of patient-based outcomes and other key variables. Obviously the costs of a register would increase with the scale of its format.

The case for a register also needs to be considered in terms of the potential benefits of informing orthopaedic surgery and society generally about determinants of outcome other than prostheses, such as the role of surgical expertise and surgical technique. Rather than recommending the setting up of a register *per se,* instead we urge that the case for a register should be evaluated very seriously. The case for a register is not clear-cut because of all of the limitations of drawing clear inferences from observational evidence which have been illustrated by this review. A register may only draw attention to the most poorly performing prostheses.

Randomised trials will produce more precise estimates of the relative cost-effectiveness of prostheses, and we recommend that their role be as seriously examined as that of the register. It may well prove impossible to justify trials of the more expensive cementless prostheses because, taking revision surgery as the proxy indicator of outcome for such trials, the extent of the reduction in revision rate they would need to achieve is so large and implausible. On the other hand, the preliminary economic modelling reported here suggests that prostheses costing 150% more than the Charnley prosthesis could well be found to be cost-effective from the evidence of trials because only modest reductions in revision surgery would need to be achieved. In particular, our modelling points to trials of different prostheses for use in younger patients as being more likely to deliver evidence of plausible levels of improvement. The case for such trials particularly needs to be considered.

Although it has often been observed that there is no tradition of RCTs in orthopaedic surgery, and the practical and other obstacles are substantial, recent successful implementation of RCTs in relation to THR[12,38] suggests that there are no insurmountable problems. The main problem is that recent examples of RCTs were almost certainly considerably under-powered. To produce large sample sizes, and so that trials in this field do not become prohibitively expensive, greater thought needs to be given to the feasibility of multicentre trials of THR. One of the most labour intensive and therefore most expensive aspects of surgical trials is the clinical assessment of outcome. It is increasingly clear that whereas surgical opinions of key aspects of outcome of THR are not accurately or reliably assessed, patient-based outcome measures of function and HRQL may provide data that are not only more standardised, reliable and validated but also more relevant and appropriate.[152–154] They may have the vital additional merit of rendering large clinical trials far more feasible. A simple patient-based measure of outcome has been developed and validated for use in THR and is currently the primary outcome for a national observational audit of THR in England.[155,156] As our review has demonstrated, there is an urgent need for clearer and more authoritative evidence of the relative value to patients of the proliferating range of prostheses available and in use in the NHS. We recommend that patient-based outcome measures be central to such evaluation.

Acknowledgement

We are indebted to the referees for their perseverance in reading the report and the quality of their comments.

References

1. Newman K. Total hip and knee replacements: a survey of 261 hospitals in England. *J R Soc Med* 1993;**86**:52–79.
2. Murray D, Carr A, Bulstrode C. Which primary total hip replacement? *J Bone Joint Surg Br* 1995;**77**:520–7.
3. Murray D, Carr A, Bulstrode C. Survival analysis of joint replacements. *J Bone Joint Surg Br* 1993;**75**:697–700.
4. Older J, Butorac R. Charnley low friction arthroplasty (LFA): a 17–21 year follow-up study. *J Bone Joint Surg Br* 1992;**74** Suppl 3:251.
5. Ritter M, Campbell E. Long-term comparison of the Charnley, Muller, and Trapezoidal-28 total hip prostheses. *J Arthrop* 1987;**2**:299–308.
6. Malchau H, Herberts P, Ahnfelt L. Prognosis of total hip replacement in Sweden. Follow-up of 92,675 operations performed in 1978–1990. *Acta Orthop Scand* 1993;**64**:497–506.
7. Owen T, Moran C, Smith S, Pinder I. Results of uncemented porous-coated anatomic total hip replacement. *J Bone Joint Surg Br* 1994;**76**:258–62.
8. Harris W, Sledge C. Total hip and total knee replacement. *New Engl J Med* 1990;**323**:725–31.
9. Williams M, Frankel S, Nanchanal K, Coast J, Donovan J. Epidemiologically based needs assessment. Total hip replacement. DHA Project: Research programme commissioned by the NHS Management Executive. London: Crown Publisher; 1992. Report 2a: 1987.
10. Sutherland C, Wilde A, Borden L, Marks K. A ten-year follow-up of one hundred consecutive Muller curved stem total hip replacement arthroplasties. *J Bone Joint Surg Am* 1982;**64**:970–81.
11. Wejkner B, Stenport J. Long-term results of bilateral Charnley total hip arthroplasty. *J Arthroplasty* 1988;**3**:305–8.
12. Marston R, Cobb A, Bentley G. Stanmore compared with Charnley total hip replacement. A prospective study of 413 arthroplasties. *J Bone Joint Surg Br* 1996;**78**:178–84.
13. Ahnfelt I, Herbert P, Malchau H. Prognosis of total hip replacement. A Swedish multicenter study of 4664 revisions. *Acta Orthop Scand* 1990;**238** Suppl:1–26.
14. Gross M. A critique of the methodologies used in clinical studies of hip-joint arthroplasty published in the English-language orthopaedic literature. *J Bone Joint Surg Am* 1988;**70**:1364–71.
15. Laupacis A, Rorabeck C, Bourne R, Feeny D, Tugwell P, Sim D. Randomized trials in orthopaedics: why, how and when? *J Bone Joint Surg Am* 1989;**71**:535–43.
16. Morris R. Evidence-based choice of hip prostheses. *J Bone Joint Surg Br* 1996;**78**:691–3.
17. Bulstrode C, Murray D, Carr A, Pynsent P, Carter S. Designer hips. *BMJ* 1993;**306**:732–3.
18. Sochart D, Long A, Porter M. Joint responsibility: the need for a national arthroplasty register. *BMJ* 1996;**313**:66–7.
19. Gillespie W, Pekarsky B, O'Connell D. Evaluation of new technologies for total hip replacement: economic modelling and clinical trials. *J Bone Joint Surg Br* 1995;**77**:528–33.
20. Pynsent P, Carter S, Bulstrode C. The total cost of hip-joint replacement: a model for purchasers. *J Public Health Medicine* 1996;**18**:157–68.
21. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;**309**:1286–91.
22. NHS Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness. CRD Report 4. York: NHS CRD, University of York; 1996.
23. Oxman A. Checklists for review articles. *BMJ* 1994;**309**:648–51.
24. Callahan C, Drake B, Heck D, Dittus R. Patient outcome following tricompartmental total knee replacement. *JAMA* 1994;**271**:1349–57.
25. Harris W. Traumatic arthritis of the hip after dislocation in acetabular fractures: treatment by mould arthroplasty. *J Bone Joint Surg Am* 1969;**51**:737–55.
26. Kavanagh B, Fitzgerald R. Clinical and roentgenographic assessment of total hip arthroplasty: a new hip score. *Clin Orthop* 1985;**193**:133–40.
27. Charnley J. Long term results of low friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br* 1972;**54**:61–3.
28. Callaghan J, Dysart S, Savory C, Hopkinson W. Assessing the results of hip replacement; a comparison of five different rating systems. *J Bone Joint Surg Br* 1990;**72**:1008–9.

29. Reigstad A, Brandt M, Hetland K. Total hip replacement with Muller prosthesis and ICLH double cup. 2–6 year results of a prospective comparative study. *Arch Orthop Trauma Surg* 1986;**105**:175–82.

30. Wykman A, Olsson E, Axdorph G, Goldie I. Total hip arthroplasty: a comparison between cemented and press-fit non-cemented fixation. *J Arthroplasty* 1991;**6**:19–29.

31. Godsiff S, Emery R, Heywood-Waddington M, Thomas T. Cemented versus uncemented femoral components in the Ring hip prosthesis. *J Bone Joint Surg Br* 1992;**74**:822–4.

32. Jacobsson S, Djerf K, Gillquist J, Hammerby S, Ivarsson I. A prospective comparison of Butel and PCA hip arthroplasty. *J Bone Joint Surg Br* 1993;**75**:624–9.

33. Reigstad A, Rokkum M, Bye K, Brandt M. Femoral remodeling after arthroplasty of the hip. *Acta Orthop Scand* 1993;**64**:411–16.

34. Søballe K, Toksvig-Larsen S, Gelineck J, *et al.* Migration of hydroxyapatite coated femoral prostheses. *J Bone Joint Surg Br* 1993;**75**:681–7.

35. Karrholm J, Malchau H, Snorrason F, Herberts P. Micromotion of femoral stems in total hip arthroplasty. *J Bone Joint Surg Am* 1994;**76**:1692–705.

36. Onsten I, Carlsson A, Ohlin A, Nilsson JA. Migration of acetabular components, inserted with and without cement, in one-stage bilateral hip arthroplasty. *J Bone Joint Surg Am* 1994;**76**:185–93.

37. Garellick G, Malchau H, Herberts P. Charnley versus Spectron – a randomized, prospective study on cemented hip arthroplasty using contemporary technique. *Acta Orthop Scand* 1995;**66** Suppl:25–6.

38. Mulliken B, Nayak N, Bourne R, Rorabeck C, Bullas R. Early radiographic results comparing cemented and cementless total hip arthroplasty. *J Arthroplasty* 1996;**11**:24–32.

39. Sudman E, Havelin L, Lunde O, Rait M. The Charnley versus the Christiansen total hip replacement operation. *Acta Orthop Scand* 1983;**54**:545–52.

40. Djerf K, Wahlstrom O. Total hip replacement. Comparison between the McKee-Farrar and Charnley prostheses in a 5-year follow-up study. *Acta Orthop Trauma Surg* 1986;**105**:158–62.

41. Jacobsson S, Djerf K, Wahlstrom O. A comparative study between McKee-Farrar and Charnley arthroplasties with long-term follow up periods. *J Arthroplasty* 1990;**5**:9–14.

42. Ritter M. The cemented acetabular component of a total hip replacement: all polyethylene versus metal backing. *Clin Orthop* 1995;**311**:69–75.

43. Agins H, Salvati E, Ranawat C, Wilson P, Pellicci P. The nine to fifteen year follow-up of one-stage bilateral total hip arthroplasty. *Orthop Clin North Am* 1988;**19**:517–30.

44. Carlsson A, Lindberg H, Sanzen L. Loosening of the socket in a 35-mm snap-fit prosthesis and the Charnley hip prosthesis. *Clinical Orthop* 1988;**228**:63–8.

45. Collis D. Long-term results of an individual surgeon. *Orthop Clin North Am* 1988;**19**:541–50.

46. Wilson-MacDonald J, Morscher E. Comparison between straight- and curved-stem Müller femoral prostheses: 5- to 10-year results of 545 total hip replacements. *Arch Orthop Trauma Surg* 1989;**109**:14–20.

47. Ritter M, Keating M, Faris P, Brugo G. Metal-backed acetabular cups in total hip arthroplasty. *J Bone Joint Surg Am* 1990;**72**:672–7.

48. Timperley J, Bannister G, Gie G, Ring P, Ling R. Effect of cup geometry and the presence of cement on acetabular component fixation. *Arch Orthop Trauma Surg* 1992;**111**:301–4.

49. Müller W, Friederich N, Laubert P. Five to ten year follow-up results of non-cemented total hip replacement. *Acta Orthopaedica Belg* 1993;**59**:297–303.

50. Huracek J, Spirig P. The effect of hydroxyapatite coating on the fixation of hip prostheses. A comparison of clinical and radiographic results of hip replacement in matched pair study. *Arch Orthop Trauma Surg* 1994;**113**:72–7.

51. Johnsson R, Franzen H, Nilsson L. Combined survivorship and multivariate analyses of revisions in 799 prostheses. *J Bone Joint Surg Br* 1994;**76**:439–43.

52. Espehaug B, Havelin L, Engesaeter L, Vollset S, Langeland N. Early revision among 12,179 hip prostheses. A comparison of 10 different brands reported to the Norwegian arthroplasty register 1987–1993. *Acta Orthop Scand* 1995;**66**:487–93.

53. Havelin L, Vollset S, Engesaeter L. Revision for aseptic loosening of uncemented cups in 4,352 primary total hip prostheses. *Acta Orthop Scand* 1995;**66**:494–500.

54. Hwang S, Park J. Cementless total hip arthroplasty with AML, PCA and HGP prostheses. *Int Orthop (SICOT)* 1995;**19**:77–83.

55. Jacobsson S, Ivarsson I, Djerf K, Wahlstrom O. Stem loosening more common with ITH than Lubinus prosthesis. *Acta Orthop Scand* 1995;**66**:425–31.

56. Ranawat C, Deshmukh R, Peters L, Umlas M. Prediction of the long-term durability of all-polyethylene cemented sockets. *Clin Orthop* 1995;**317**:89–105.

57. Weidenheim L, Mikhail M, Nelissen R, Bauer T. Cemented collarless (Exeter-CPT) versus cementless collarless (PCA) femoral components. *J Arthroplasty* 1995;**10**:592–7.

58. Salvati E, Wilson P, Jolley M, Vakili F, Aglieti P, Brown G. A ten-year follow-up study of our first one hundred consecutive Charnley total hip replacements. *J Bone Joint Surg Am* 1981;**63**:753–67.

59. Stauffer R. A ten year follow-up study of total hip replacement: with particular reference to roentgenographic loosening of the components. *J Bone Joint Surg Am* 1982;**64**:983–90.

60. Johnston R, Crowninshield R. Roentgenologic results of total hip arthroplasty. A ten year follow-up study. *Clin Orthop* 1983;**181**:92–8.

61. Sharp D, Porter K. The Charnley total hip arthroplasty in patients under age 40. *Clin Orthop* 1985;**201**:51–5.

62. Eftekhar N, Tzitzikalakis G. Failures and reoperations following low-friction arthroplasty of the hip. *Clin Orthop* 1986;**211**:65–78.

63. Kim Y, Ahn B, Ko C, Lee J, Kwak B, Yoon Y. Arthroplasty using the Charnley prosthesis in old tuberculosis of the hip. *Clin Orthop* 1986;**211**:116–21.

64. Jinnah R, Amstutz H, Tooke M, Dorey F, Dalseth T. The UCLA Charnley experience. *Clin Orthop* 1986;**211**:164–71.

65. Brady L, McCutchen J. A ten year follow-up study of 170 Charnley total hip arthroplasties. *Clin Orthop* 1986;**211**:51–4.

66. Dall D, Grobbelaar C, Learmonth I, Dall G. Charnley low friction arthroplasty of the hip. *Clin Orthop* 1986;**211**:85–90.

67. Halley D, Wroblewski B. Long-term results of low friction arthroplasty in patients 30 years of age or younger. *Clin Orthop* 1986;**211**:43–50.

68. Hedeboe J, Kjaer J, Gudmundsson G. Failures in total hip replacement: a long-term follow-up. *Acta Orthop Belg* 1986;**52**:405–14.

69. Wroblewski B. 15–21 year results of the Charnley low-friction arthroplasty. *Clin Orthop* 1986;**211**:30–5.

70. Older J. Low-friction arthroplasty of the hip. A ten to twelve year follow-up study. *Clin Orthop* 1986;**211**:36–42.

71. Dall D, Learmonth I, Dall G, Record C. Long-term results of Charnley low-friction arthroplasty of the hip. *S Afr Med J* 1988;**73**:89–92.

72. Linde F, Jensen J, Pilgaard S. Charnley arthroplasty in osteoarthritis secondary to congenital dislocation or subluxation of the hip. *Clin Orthop* 1988;**227**:164–71.

73. Wejkner B, Stenport J. Charnley total hip replacement: a ten to 14 year follow-up study. *Clin Orthop* 1988;**231**:113–19.

74. Kavanagh B, Dewitz M, Ilstrup D, Stauffer R, Coventry M. Charnley total hip arthroplasty with cement: fifteen year results. *J Bone Joint Surg Am* 1989;**71**:1496–503.

75. Hartofilakidis G, Stamos K, Ioannidis T. Fifteen years' experience with Charnley low-friction arthroplasty. *Clin Orthop* 1989;**248**:48–56.

76. Loudon J, Older M. Subsidence of the femoral component related to long-term outcome of hip replacement. *J Bone Joint Surg Br* 1989;**71**:624–8.

77. Hozack W, Rothman R, Booth R, Balderston R, Cohn J, Pickens G. Survivorship analysis of 1041 Charnley total hip arthroplasties. *J Arthroplasty* 1990;**5**:41–7.

78. Skeie S, Lende S, Sjoberg E, Vollset S. Survival of the Charnley hip in coxarthrosis. A 10–15 year follow-up of 629 cases. *Acta Orthop Scand* 1991;**62**:98–101.

79. Wroblewski B, Taylor G, Siney P. Charnley low-friction arthroplasty: 19–25 year results. *Orthopaedics* 1992;**15**:421–4.

80. Karachalios T, Hatoflakidis G, Zacharakis N, Tsekoura M. A 12 to 18 year radiographic follow-up study of Charnley low friction arthroplasty: the role of the center of rotation. *Clin Orthop* 1993;**296**:140–7.

81. Joshi A, Porter M, Trail J, Hunt L, Murphy J, Hardinge K. Long-term results of Charnley low-friction arthroplasty in young patients. *J Bone Joint Surg Br* 1993;**75**:616–23.

82. Hodgkinson J, Maskell A, Ashok P, Wroblewski B. Flanged acetabular components in cemented Charnley hip arthroplasty. *J Bone Joint Surg Br* 1993;**75**:464–7.

83. Wroblewski M, Siney P. Charnley low-friction arthroplasty in the young patient. *Clin Orthop* 1992;**285**:45–7.

84. Kobayashi S, Eftekhar N, Terayama K. Predisposing factors in fixation failure of femoral prostheses following primary Charnley low-friction arthroplasty: a 10- to 20-year follow-up. *Clin Orthop* 1994;**306**:73–83.

85. Garellick G, Herberts P, Stromberg C, Malchau H. Long-term results of Charnley arthroplasty: a 12–16 year follow-up study. *J Arthroplasty* 1994;**9**:333–40.

86. Neumann L, Freund K, Sorenson KH. Long-term results of Charnley total hip replacement: review of 92 patients at 15 to 20 years. *J Bone Joint Surg Br* 1994;**76**:245–51.

87. Onsten I, Besjakov J, Carlsson A. Improved radiographic survival of the Charnley prosthesis in rheumatoid arthritis and osteoarthritis. *J Arthroplasty* 1994;**9**:3–8.

88. Sullivan P, MacKenzie J, Callaghan J, Johnston R. Total hip arthroplasty with cement in patients who are less than fifty years old: a sixteen to twenty-two year follow-up study. *J Bone Joint Surg Am* 1994;**76**:863–9.

89. Callaghan J. Cemented arthroplasty: a long look back. *Orthopaedics* 1995;**18**:803.

90. Almby B, Hierton T. Total hip replacement: a ten year follow-up of an early series. *Acta Orthop Scand* 1982;**53**:397–406.

91. Reikera O. Ten year follow-up of Muller hip replacements. *Acta Orthop Scand* 1982;**53**:919–22.

92. Dunn W, Hamilton L. Muller curved-stem total hip arthoplasty. *South Med J* 1986;**79**:698–701.

93. Stockley I, McLean L, Gross A. Clinical and radiographic results of the Muller straight stem used as a press-fit. *J Arthroplasty* 1992;**7**:477–82.

94. Santori F, Mancini A, Manili F, Falez F, Vitullo A, Anzini M. The cementless PCA primary total hip system. *Ital J Orthop Traumatol* 1992;**18**:287–95.

95. Heekin D, Callaghan J, Hopkinson W, Savory C, Xenos J. The porous-coated anatomic total hip prosthesis, inserted without cement. Results after 5–7 years in a prospective study. *J Bone Joint Surg Am* 1993;**75**:77–91.

96. Hungerford D, Krackow K. Clinical experience with the PCA total hip replacement: femoral component – minimum five year follow-up. *Acta Orthop Belg* 1993;**59** Suppl 1:182.

97. Wellen P, Demuynck M, Haentjens P, Valkeneer O, Casteleyn P, Berghe I, *et al.* Total hip arthroplasty with the porous-coated anatomic (PCA) prosthesis: the femoral component. *Acta Orthop Belg* 1993;**59**:282–6.

98. Kim Y, Kim V. Uncemented porous-coated anatomic total hip replacement: results at six years in a consecutive series. *J Bone Joint Surg Br* 1993;**75**:6–14.

99. Bourne R, Rorabeck C, Ghazal, Lee M. Pain in the thigh following total hip replacement with a porous-coated anatomic prosthesis for osteoarthrosis. *J Bone Joint Surg Am* 1994;**76**:1464–70.

100. Billotti J, Zimmerman M, Pizzurro J, Mango T, Billings J, Parsons R. The porous-coated anatomic (PCA) total hip arthroplasty: a review of 73 uncemented cases with 2 year follow-up. *Orthopaedics* 1995;**18**:37–43.

101. Andrew T, Berridge D, Thomas A, Duke R. Long term review of Ring total hip arthroplasty. *Clin Orthop* 1985;**201**:111–23.

102. Nunn D. The Ring uncemented plastic-on-metal total hip replacement. *J Bone Joint Surg Br* 1988;**70**:40–4.

103. Albrecht-Olsen P, Owen-Falkenberg T, Burgaard P, Andersen P. Nine year follow-up of the cementless Ring hip. *Acta Orthop Scand* 1989;**60**:77–80.

104. Bryant M, Mollan R, Nixon J. Survivorship analysis of the Ring hip arthroplasty. *J Arthroplasty* 1991;**6** Suppl:S5–10.

105. Harper G, Bull T, Cobb A, Bentley G. Failure of the Ring polyethylene uncemented acetabular cup. *J Bone Joint Surg Br* 1995;**77**:557–61.

106. Tillberg B. Total hip arthroplasty using the McKee and Watson-Farrar prosthesis. *Acta Orthop Scand* 1982;**53**:103–7.

107. August A, Aldam C, Pynsent P. The McKee-Farrar hip arthroplasty: a long-term study. *J Bone Joint Surg Br* 1986;**68**:520–7.

108. Visuri T. Long-term results and survivorship of the McKee-Farrar total hip prosthesis. *Arch Orthop Trauma Surg* 1987;**106**:368–74.

109. Jantsch S, Schwagerl W, Zenz P, Semlitsch M, Fertschak W. Long-term results after implantation of McKee-Farrar total hip prostheses. *Arch Orthop Trauma Surg* 1991;**110**:230–7.

110. Schmalzried T, Harris W. The Harris-Galante porous-coated acetabular component with screw fixation. *J Bone Joint Surg Am* 1992;**74**:1130–9.

111. Kim Y, Kim V. Results of the Harris-Galante cementless hip prosthesis. *J Bone Joint Surg Br* 1992;**74**:83–7.

112. Galante J. Five year HGP femoral component. *Acta Orthopaedica Belg* 1993;**59** Suppl:153–4.

113. Dobbs H. Survivorship of total hip replacements. *J Bone Joint Surg Br* 1980;**62**:168–73.

114. Van der Schaaf D, Deutman R, Mulder T. Stanmore total hip replacement: a nine to ten year follow-up. *J Bone Joint Surg Br* 1988;**70**:45–8.

115. Alsema R, Deutman R, Mulder T. Stanmore total hip replacement: a nine to ten year follow-up. *J Bone Joint Surg Br* 1988;**70**:45–8.

116. Olsson S, Jernberger A, Tryggo D. Clinical and radiological long-term results after Charnley-Muller total hip replacement. *Acta Orthop Scand* 1981;**52**:531–42.

117. Pavlov P. A 15-year follow-up study of 512 consecutive Charnley-Muller total hip replacements. *J Arthroplasty* 1987;**2**:151–5.

118. Fowler J, Gie G, Liang R. Experience with the Exeter total hip replacement since 1970. *Orthop Clin North Am* 1988;**19**:477–89.

119. Rockburn P, Olsson S. Loosening and bone resorption in Exeter hip arthroplasties: a review at a minimum of five years. *J Bone Joint Surg Br* 1993;**75**:865–8.

120. Cowley D. Prostheses for primary total hip replacement: a critical appraisal of the literature. *Int J Technol Assess Health Care* 1995;**11**:770–8.

121. Yahiro M, Gantenberg J, Nelson R, Lu H, Mishra N. Comparison of the results of cemented, porous-ingrowth, and threaded acetabular cup fixation. *J Arthroplasty* 1995;**10**:339–50.

122. NHS Centre for Reviews and Dissemination. Total hip replacement. *Effective Health Care* 1996; **2**(7):1–12.

123. Van der Linden V. Pitfalls in randomized surgical trials. *Surgery* 1980;**87**:258–62.

124. Stirrat G, Farndon J, Farrow S, Dwyer N. The challenge of evaluating surgical procedures. *Ann R Coll Surg Engl* 1992;**74**:80–4.

125. Rudicel S, Esdaile J. The randomized clinical trial in orthopaedics: obligation or option? *J Bone Joint Surg Am* 1985;**67**:1284–93.

126. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;**312**:1215–18.

127. Malchau H, Herberts P. Prognosis of total hip replacement. Presented at 63rd Annual Meeting of the American Academy of Orthopaedic Surgeons; 1996; Atlanta, GA, USA.

128. Sculpher MJ, Drummond MF, Buxton MJ. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy* 1997;**2**:26–30.

129. Doubilet P, Weinstein MC, McNeil BJ. Use and misuse of the term "cost effective" in medicine. *New Engl J Med* 1986;**314**:253–6.

130. Daellenbach HG, Gillespie WJ, Crosbie P, Daellenbach US. Economic appraisal of new technology in the absence of survival data – the case of total hip replacement. *Soc Sci Med* 1990;**31**:1287–93.

131. Gillespie WJ, Pekarsky B, O'Connell DL. Evaluation of new technologies for total hip replacement: economic modelling and clinical trials. *J Bone Joint Surg Br* 1995;**77**:528–33.

132. Britton AR, Murray DW, Bulstrode CJ, *et al.* A long term comparison of the Charnley and Stanmore total hip replacements. *J Bone Joint Surg Br* 1996;**78**:802–8.

133. Sonnenberg FA, Beck JR. Markov models in medical decision making. *Med Decis Making* 1993;**13**:322–38.

134. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ* 1994;**3**:95–104.

135. Seagroatt V, Tan HS, Goldacre M, Bulstrode C, Gill L. Elective total hip replacement; incidence, emergency readmission rate, and postoperative mortality. *BMJ* 1991;**303**:1431–5.

136. Central Statistical Office annual abstract of statistics. London: HMSO; 1995.

137. Review body for nursing staff, midwives, health visitors and professions allied to medicine (Chairman Bryan Rigby) (1996). Thirteenth Report. London: HMSO.

138. Review body on doctors' and dentists' remuneration (Chairman CB Gough) (1996). Twenty-Fifth Report. London: HMSO.

139. Drummond MF, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press, 1987.

140. HM Treasury economic appraisal in central government: a technical guide for government departments. London: HMSO; 1991.

141. Patrick DL, Erickson P. Health Status and Health Policy. Allocating Resources to Health Care. New York: Oxford University Press; 1993.

142. Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, *et al.* The effect of elective total hip replacement on health-related quality of life. *J Bone Joint Surg Am* 1993;**75**:1619–26.

143. Torrance GW. Measurement of health state utilities for economic appraisal – a review. *J Health Econ* 1986;**5**:1–30.

144. Laupacis A, Feeny D, Detsky AS, Tugwell P. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J* 1992;**146**:473–81.

145. Briggs AH, Sculpher MJ, Logan RPH, Aldous J, Ramsay M, Baron J. Cost effectiveness of screening for and eradication of Helicobacter pylori in management of dyspeptic patients under 45 years of age. *BMJ* 1996;**312**:1321–5.

146. Liang MH, Cullen KE, Larson MG, Thompson M, Schwartz J, Fossel A, *et al.* Cost-effectiveness of total joint arthroplasty in osteoarthritis. *Arthritis Rheum* 1986;**29**:937–43.

147. Bourne RB, Rorabeck Ch, Laupacis A, Feeny D, Wong C, Tugwell P, *et al.* A randomized clinical trial comparing cemented to cementless total hip replacement in 250 osteoarthritic patients: the impact on health related quality of life and cost effectiveness. *Iowa Orthop J* 1994;**14**:108–14.

148. Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, *et al.* Costs of elective total hip arthroplasty during the first year: cemented versus noncemented. *J Arthroplasty* 1994;**9**:481–7.

149. Chang RW, Pellissier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA* 1996;**275**:858–65.

150. Maynard A. The production of health and health care. *J Econ Stud* 1985;**10**:31–45.

151. Furnes A, Lie SA, Engesaeter LB. The economic impact of total hip replacement surgery. *Acta Orthop Scand* 1996;**67**:115–21.

152. O'Boyle C, McGee H, Hickey A, O'Malley K, Joyce C. Individual quality of life in patients undergoing hip replacement. *Lancet* 1992;**339**:1088–91.

153. Amadio P. Outcomes measurement: more questions; some answers. *J Bone Joint Surg Am* 1993;**75**:1583–4.

154. Cleary P, Reilly D, Greenfield S, Mulley A, Wexler L, Frankel F, *et al.* Using patient reports to assess health-related quality of life after total hip replacement. *Qual Life Res* 1993;**2**:3–11.

155. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg Br* 1996;**78**:185–90.

156. Dawson J, Fitzpatrick R, Murray D, Carr A. Comparison of measures to assess outcomes in total hip replacement. *Qual Health Care* 1996;**5**:81–8.

157. Detsky A, Naylor C, O'Rourke K, McGeer A, Abbe K. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol* 1992;**45**:255–65.

158. Oxman A, Chalmers I, Clarke M, Enkin M, Schulz K, Starr M. The Cochrane Collaboration handbook: section VI: preparing and maintaining systematic reviews. Oxford: UK Cochrane Centre; 1994.

159. Morris R. A statistical study of papers in the Journal of Bone and Joint Surg Br 1984. *J Bone Joint Surg Br* 1988;**70**:242–6.

160. Chalmers T, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, *et al.* A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials* 1981;**2**:32–49.

161. Britton A, Murray D, Bulstrode C, McPherson K, Denham R. Loss to follow-up: does it matter? *Lancet* 1995;**345**:1511–12.

162. Kelly M, Wadsworth J. What price inconclusive clinical trials? *Ann R Coll Surg Engl* 1993;**75**:145–6.

Appendix 1

A survey of current use of prostheses by NHS orthopaedic surgeons

To identify current patterns of use of prostheses in THR by UK orthopaedic surgeons, a survey was designed and carried out.

Methods

A draft questionnaire was drawn up, intended to obtain basic information about preferences regarding prostheses in THR. The questionnaire was then tested by asking six orthopaedic surgeons to fill it out and then assess its relevance and acceptability. A revised version of the questionnaire was then tested on two further orthopaedic surgeons. The focus of the questionnaire was upon currently used prostheses, recent changes in use, reasons for change, and views about the role of prostheses compared with other factors in influencing outcomes of THR.

Sampling

The Register of Fellows of the British Orthopaedic Association was used as a sampling frame. Every third surgeon listed as resident in the UK was sent a questionnaire and explanatory letter. Non-respondents were sent a reminder letter after 4 weeks.

A total of 336 surgeons returned the questionnaire (81% response rate). Of these, 302 surgeons had performed primary THR in the last year. They are the primary focus of the survey for this review and the denominator for results. Their average age was 47.1 years (range, 35–70 years). All were of consultant grade. They were asked to estimate the number of THRs they had performed in the year before the survey. About one-third (33.8%) reported performing 30 or fewer, whereas 47% reported performing more than 40 THR operations in the past year.

Use of prostheses

By far the most commonly used prosthesis was the Charnley (53% of respondents; *Table 1*). The only other prostheses used by more than 10% of respondents were the Exeter and Müller.

Comment

The dominant role of the Charnley prosthesis is consistent with the evidence of Newman[1] who found that it was by far the most commonly used prosthesis in NHS hospitals in England in 1991 (reported as in use by 74% of hospitals performing THR). Also consistent with Newman's survey is the much lower frequency with which cementless prostheses are reported as being used. In our survey, as in Newman's, the Furlong was reported as the most frequently used cementless prosthesis. The one major difference between the two surveys is that the Ring cementless prosthesis was reported by Newman as in use by almost as many hospitals as reported the use of the Furlong, whereas at the time of the current survey less than 1% of surgeons reported the use of the Ring cementless prosthesis.

Appendix 2

MEDLINE search strategy

2. RANDOMIZED-CONTROLLED TRIAL in PT
3. RANDOMIZED-CONTROLLED-TRIALS/all subheadings
4. RANDOM-ALLOCATION (Term allows no subheadings)
5. DOUBLE-BLIND-METHOD (Term allows no subheadings)
6. SINGLE-BLIND-METHOD (Term allows no subheadings)
7. #2 or #3 or #4 or #5 or #6
8. TG=ANIMAL not (TG=HUMAN and TG=ANIMAL)
9. #7 not #8
10. CLINICAL-TRIAL in PT
11. explode CLINICAL-TRIALS/all subheadings
12. (CLIN* near TRIAL*) in TI
13. (CLIN* near TRIAL*) in AB
14. (SINGL* or DOUBL or TREBL* or TRIPL*) near (BLIND* or MASK*)
15. (#14 in TI or (#14 in AB)
16. PLACEBOS/all subheadings
17. PLACEBO* in TI
18. PLACEBO* in AB
19. RANDOM* in TI
20. RANDOM* in AB
21. RESEARCH-DESIGN/all subheadings
22. #10 or #11 or #12 or #13 or #15 or #16 or #17 or #18 or #19 or #20 or #21
23. 69331 #8
24. #22 not #23
25. HIP PROSTHESIS/all subheadings
26. PROSTHESIS-FAILURE/all subheadings
27. CEMENTATION /
28. REOPERATION/all subheadings
29. explode PROSTHESIS-DESIGN/all subheadings
30. PROSTHESIS-RELATED INFECTIONS/ all subheadings
31. PROSTHESIS-FITTING/all subheadings
32. explode COHORT STUDIES (Term allows no subheadings)
33. explode OUTCOME-AND-PROCESS ASSESSMENT-HEALTH-CARE/all subheadings
34. PATIENT-SATISFACTION/all subheadings
35. CLINICAL-TRIAL in PT
36. MULTICENTER-STUDIES/all subheadings
37. INTERVENTION-STUDIES(Term allows no subheadings)
38. explode EVALUATION-STUDIES
39. explode TREATMENT-OUTCOME (Term allows no subheadings)
40. #25 or #26 or #27 or #28 or #29 or #30 or #31
41. #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39
42. #9 or #24 or #41
43. #42 and #40

Appendix 3

Ranked listing of journals from MEDLINE

Sample years: 1985, 1990, 1994

1. *Clinical Orthopaedics and Related Research*
2. *Journal of Bone and Joint Surgery* (American edition)
3. *Journal of Bone and Joint Surgery* (British edition)
4. *Acta Orthopaedica Scandinavica*
5. *Acta Orthopaedica Scandinavica Supplementum*
6. *Acta Orthopaedica Belgica*
7. *Journal of Arthroplasty*
8. *Zeitschrift fur Orthopadie und ihre Grenzgebiete*
9. *Orthopedics*
10. *Archives of Orthopaedic and Trauma Surgery*
11. *International Orthopaedics*

Appendix 4

EMBASE search strategy

1. CLINICAL STUDY @ EX
2. CASE REPORT 'KMAJOR, KMINOR
3. 1-2
4. DOUBLE BLIND PROCEDURE @ KMAJOR, KMINOR
5. SINGLE BLIND PROCEDURE @ KMAJOR, KMINOR
6. FOLLOW UP @ KMAJOR, KMINOR
7. LONG TERM @ TI, AB, KWDS
8. ARTHROPLASTY @ KMAJOR, KMINOR
9. ACETABULUM FRACTURE @KMAJOR, KMINOR
10. HIP FRACTURE @ KMAJOR, KMINOR
11. HIP INJURY @ KMAJOR, KMINOR
12. HIP SURGERY @ EX
13. HUMAN @ KMAJOR, KMINOR
14. 3, 4, 5, 6, 7
15. 8, 9, 10, 11, 12
16. 14 + 15
17. HIP REPLACEMENT* @ TI, AB, KWDS
18. HIP ARTHROPLAST* @ TI, AB, KWDS
19. HIP PROSTHE* @ TI, AB, KWDS
20. ACETABUL* @TI, AB, KWDS
21. BI-POLAR @ TI, AB, KWDS
22. HEMI-ARTHROPLAST* @ TI, AB, KWDS
23. HEMIARTHROPLAST* @ TI, AB, KWDS
24. RANDOM* @ TI, AB, KWDS
25. DOUBLE BLIND @TI, AB, KWDS
26. SINGLE BLIND @ TI, AB, KWDS
27. TRIPL*BLIND @ TI, AB, KWDS
28. TREBL*BLIND @TI, AB, KWDS
29. SINGLE MASK* @ TI, AB, KWDS
30. DOUBLE MASK* @ TI, AB, KWDS
31. TRIPLE* MASK* @ TI, AB, KWDS
32. TREBL*MASK* @ TI, AB,KWDS
33. CONTROL*TRIAL* @ TI, AB, KWDS
34. CONTROL* STUD* @ TI, AB, KWDS
35. FOLLOW UP @ TI, AB, KWDS
36. 17, 18, 19, 20, 21, 22, 23
37. 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35
38. 36 + 37
39. 16, 38
40. 13 + 39

Appendix 5

Ranked listing of journals from EMBASE

Sample years: 1992–1994

1. *Clinical Orthopaedics and Related Research*
2. *Journal of Bone and Joint Surgery* (British edition)
3. *Acta Orthopaedica Scandinavica*
4. *Journal of Arthroplasty*
5. *Journal of Bone and Joint Surgery* (American edition)
6. *International Orthopaedics*
7. *Archives of Orthopaedic and Trauma Surgery*
8. *Journal of the Japanese Orthopaedic Association*
9. *Acta Anaesthesiologica Scandinavica*
10. *International Journal of Radiation Oncology, Biology, Physics*

Appendix 6

Assessment of methodological quality of trials

Introduction

At the beginning of the study, it was decided that an analysis of the quality of evidence obtained for this review be undertaken in order to examine whether any 'false positive' or 'false negative' conclusions might arise because of inclusion in the review of less rigorously conducted studies.[157] An informal inspection of the quality of studies obtained from the search phase of the review made it clear that there were unlikely to be any clear positive or negative conclusions from the review. In addition, there were unlikely to be many generalisations about the presence or absence of effects of specific prostheses based on the highest quality of evidence available (i.e. RCTs). It was also clear that there was no scope for meta-analysis of RCTs because each trial addressed a unique comparison of prostheses. On the other hand it was thought important to provide an assessment of the quality of evidence in this field for its own sake, rather than to determine inclusion criteria for studies for the review or to weight studies in the analysis of results. In particular, an assessment of the quality of studies in this field should provide indications of how future research on prostheses could be conducted more effectively. It was therefore decided to review all RCTs found in the search and a sample of other studies.

Methods

There are a number of different check-lists available for the assessment of RCTs, but as there is no 'gold-standard' measure of a good trial there is therefore no method of determining the best check-list.[157,158] It was decided that a check-list was needed that could be applied to non-randomised studies as well as to RCTs. A check-list was drawn up based on a previous check-list developed by one of the current investigators (RM) to assess the quality of studies appearing in an orthopaedic research journal.[159] This check-list is highly relevant to the assessment of orthopaedic research but had not been developed specifically to assess RCTs. Therefore the investigators also drew on a commonly cited check-list for RCTs.[160] The resulting instrument was then piloted on four studies by three investigators (JD, RF, ES) and a discussion was held to draw up a revised draft in the light of the pilot.

In the main study, each paper was assessed by three raters (one orthopaedic surgeon, and two non-surgeon researchers working in the field of orthopaedics). Raters were drawn from a panel of three orthopaedic surgeons (C. Bulstrode, AC, DM) and four non-surgical researchers who have worked on outcomes of orthopaedic surgery (JD, RF, RM, ES). In this way each paper was rated by at least one expert in orthopaedic surgery and two experts in research methods. Papers were scored blind to others' ratings. A paper was given its final score on the basis of a majority 'vote' (e.g. two out of three raters agreeing that a study did not have a clear definition of primary outcome). No majority agreed rating occurred when, for example, on the item about a clear definition of primary outcome, one rater's opinion was 'yes', another's was 'no' and the third's was 'not applicable'. An indication of the extent of agreement between raters is that of 1053 ratings required over the 39 studies, there was no majority agreement on only 56 (5.3%).

Thirty-nine studies were rated. They were selected in the following way. All 12 RCTs were included in the assessment (11 studies cited in the main review, together with an RCT of method of fixation that was subsequently excluded from the review). An equal number (12) of comparative observational studies were selected from the 18 studies included in the main review. Finally, 15 observational studies of single prostheses were selected. Studies were stratified into three approximately equal periods: 1980–1984, 1985–1989 and 1990–1995, and five studies were randomly selected from each of the three periods.

To provide a simple standardised expression of the quality of studies, groups of studies (RCTs, comparative studies, single prosthesis studies) as a whole are scored in each of the six areas assessed:

- clarity of study question and definition of outcome
- description of prosthesis and fixation
- description of study sample
- adequacy of randomisation
- duration and completeness of follow-up
- statistical and analytical considerations.

Scores are the sum of positive ('yes') scores as a proportion of possible positive scores (with 'not applicable' removed from the denominator). An entirely arbitrary classification was made on the basis of the score: any criterion with a score of 75% and above was classified as 'good quality', a score between 74% and 50% was taken to indicate 'intermediate quality', and a score below 50% was deemed to indicate 'poor quality'. Scores for individual assessment items are shown in *Tables 11–13*.

Results

RCTs

The 12 RCTs were rated very favourably in terms of two dimensions: clarity of study question and of definition of outcome, and description of prosthesis and fixation method, with positive scores of 90% and 92% respectively *(Table 14)*.

RCTs were of 'intermediate quality' in terms of description of the study sample, duration and completeness of follow-up and statistical and analytical considerations. There was very little evidence of co-morbidity assessed in patients. This may not be considered a serious failing since patients have to be sufficiently fit to undergo THR. More seriously, only five out of 12 trials gave a clear description of inclusion and exclusion criteria. This makes it difficult for readers to know exactly to whom trial results apply. In terms of duration and completeness of follow-up, no trial analysed the consequences of loss to follow-up. This can be a serious source of potential bias in orthopaedic clinical trials.[161] Probably the most serious single deficit in the RCTs is that no trial was based on any justification of sample size such as power calculation. This is of major concern. Trials that are too small risk committing type II errors of missing a difference that truly exists. It is likely that power calculations would reveal that to detect relatively small differences in infrequent problems between prostheses, trials of at least 1000 are required.[2,16,162] In this case all of the trials found in the search may well be underpowered. This is probably why only one of the 12 trials was rated as of adequate sample size to detect failure.

The RCTs in the study were scored overall as 'poor' in terms of control of bias in study design. If the items in *Table 11* are examined more closely, it is not clear how feasible or how important it is in orthopaedic surgical trials to mask patients to their study assignment, or indeed for study statisticians and investigators to be blind to the identity of arms of the trial. It is of more concern that in five out of 12 trials the investigators did not check baseline values to examine the success of the randomisation.

Comparative observational studies

As with the RCTs, the comparative observational studies were rated favourably in terms of the first two of the six general criteria: clarity of study question and definition of outcome, and description of prosthesis and fixation method. Nevertheless, it is of concern that, from more specific items *(Table 12),* four out of 12 reports did not give an adequate description of the method of fixation. Since the comparative studies with concurrent controls form the most important source of evidence of relative performance of prostheses, such basic aspects of a study need to be described more accurately.

The comparative studies were rated overall as intermediate in terms of quality with regard to duration and completeness of follow-up. As with RCTs there was hardly any analysis of the consequences of loss to follow-up. Three out of 12 studies did not provide any explanation for their loss to follow-up and four out of 12 did not make clear the duration of observation to follow-up assessment.

The very low score for control of bias in study design requires comment. It could reasonably be argued that since these were non-experimental studies, this criterion is inappropriate. Raters gave low scores rather than 'not applicable' because it is at least debatable whether, if comparative observational studies are to continue to be central to orthopaedic surgical research, issues regarding blinding may still be considered as methods of reducing bias. Certainly it is of even greater importance for non-randomised designs to examine baseline values for patients being compared.

Comparative observational studies were overall rated as poor in terms of describing the study sample. A majority of studies failed to describe adequately inclusion and exclusion criteria, basic characteristics of the sample and disease/diagnostic information *(Table 12)*. Studies were also 'poor' in terms of statistical and analytical considerations. As with RCTs, the most notable feature is the failure to justify sample size (and therefore the likelihood that samples were inadequate). It is of interest that Malchau and colleagues[6] estimate from the data in the Swedish register that to have adequate power to detect a difference in revision rate between Charnley and CAD over 5 years would require a sample of 19,180 patients.

TABLE 11 *Methodological assessment of sample of RCTs of prostheses. Agreed ratings for 12 RCTs*

Assessment item	Yes	No	Unable to judge	Not applicable	No agreed rating
1. Clarity of study question and definition of outcome					
Is the purpose of the study clearly stated?	12	–	–	–	–
Is the definition of prosthesis failure clear?	11	–	–	–	1
Is there a clear definition of primary outcome(s)?	11	–	–	–	1
Are standardised outcome measures used?	10	2	–	–	–
Are the outcome measures used appropriate for the purpose of the study?	10	2	–	–	–
2. Description of prosthesis and method of fixation					
Is the prosthesis design adequately described?	11	1	–	–	–
Is the method of fixation adequately described?	11	1	–	–	–
3. Description of study sample					
Is the method of selection of the sample adequately described?	9	3	–	–	–
Are study exclusion and inclusion criteria stated?	5	2	5	–	–
Is the baseline sample clearly described in terms of basic characteristics (age, sex etc.)?	9	3	–	–	–
Is the study sample sufficiently homogeneous in terms of disease/diagnosis?	10	1	–	–	1
Is the study sample sufficiently homogeneous in terms of co-morbidity?	1	–	10	–	1
4. Control of bias in study design					
Is the method of randomisation adequate?	6	1	4	–	1
Is a method of masking the patient to the intervention allocated stated?	2	8	1	–	1
Were outcome assessors blind to intervention allocation?	4	4	4	–	–
Are baseline values for groups compared?	6	5	–	–	1
5. Duration and completeness of follow-up					
Are intervals between surgery and follow-up assessment clearly stated?	12	–	–	–	–
Are reasons for loss of patients at follow-up stated?	8	2	–	1	1
Are those lost to follow-up compared to rest of sample?	–	8	–	–	4
Is there an appropriate length of follow-up?	9	2	–	–	1
6. Statistical and analytical considerations					
Has the study sample size been justified?	–	12	–	–	–
Are the data clearly presented?	9	3	–	–	–
Was the data analyst masked to interventions?	–	–	12	–	–
Has type of statistical test and actual probability value been stated?	9	3	–	–	–
Are statistical tests appropriate to study?	7	4	–	–	1
Is the sample on which failures are assessed adequate?	1	3	1	6	1
Are conclusions justified by evidence?	9	1	–	–	2

TABLE 12 *Methodological assessment of sample of RCTs of prostheses. Agreed ratings for 12 comparative observational studies*

Assessment item	Yes	No	Unable to judge	Not applicable	No agreed rating
1. Clarity of study question and definition of outcome					
Is the purpose of the study clearly stated?	11	1	–	–	–
Is the definition of prosthesis failure clear?	10	2	–	–	–
Is there a clear definition of primary outcome(s)?	9	3	–	–	–
Are standardised outcome measures used?	6	4	–	–	2
Are the outcome measures used appropriate for the purpose of the study?	7	1	–	–	4
2. Description of prosthesis and method of fixation					
Is the prosthesis design adequately described?	10	2	–	–	–
Is the method of fixation adequately described?	8	4	–	–	–
3. Description of study sample					
Is the method of selection of the sample adequately described?	8	4	–	–	–
Are study exclusion and inclusion criteria stated?	3	7	–	–	2
Is the baseline sample clearly described in terms of basic characteristics (age, sex etc)?	5	7	–	–	–
Is the study sample sufficiently homogeneous in terms of disease/diagnosis?	3	5	3	–	1
Is the study sample sufficiently homogeneous in terms of co-morbidity?	–	1	11	–	–
4. Control of bias in study design					
Is the method of randomisation adequate?	–	–	–	12	–
Is a method of masking the patient to the intervention allocated stated?	–	–	–	12	–
Were outcome assessors blind to intervention allocation?	–	4	1	7	–
Are baseline values for groups compared?	–	12	–	–	–
5. Duration and completeness of follow-up					
Are intervals between surgery and follow-up assessment clearly stated?	7	4	–	–	1
Are reasons for loss of patients at follow-up stated?	6	3	–	1	2
Are those lost to follow-up compared to rest of sample?	1	9	–	1	1
Is there an appropriate length of follow-up?	9	1	–	–	2
6. Statistical and analytical considerations					
Has the study sample size been justified?	1	11	–	–	–
Are the data clearly presented?	9	3	–	–	–
Was the data analyst masked to interventions?	–	6	1	1	4
Has type of statistical test and actual probability value been stated?	9	3	–	–	–
Are statistical tests appropriate to study?	5	5	1	–	1
Is the sample on which outcomes are assessed adequate?	4	5	2	–	1
Are conclusions justified by evidence?	8	2	–	–	2

TABLE 13 *Methodological assessment of sample of observational studies of single prostheses. Agreed ratings for 15 observational studies*

Assessment item	Yes	No	Unable to judge	Not applicable	No agreed rating
1. Clarity of study question and definition of outcome					
Is the purpose of the study clearly stated?	14	1	–	–	–
Is the definition of prosthesis failure clear?	15	–	–	–	–
Is there a clear definition of primary outcome(s)?	14	1	–	–	–
Are standardised outcome measures used?	13	2	–	–	–
Are the outcome measures used appropriate for the purpose of the study?	12	1	–	–	2
2. Description of prosthesis and method of fixation					
Is the prosthesis design adequately described?	15	–	–	–	–
Is the method of fixation adequately described?	14	1	–	–	–
3. Description of study sample					
Is the method of selection of the sample adequately described?	7	8	–	–	–
Are study exclusion and inclusion criteria stated?	3	12	–	–	–
Is the baseline sample clearly described in terms of basic characteristics (age, sex etc)?	–	–	–	15	–
Is the study sample sufficiently homogeneous in terms of disease/diagnosis?	8	2	2	1	2
Is the study sample sufficiently homogeneous in terms of co-morbidity?	–	–	15	–	–
4. Control of bias in study design					
Is the method of randomisation adequate?	–	–	–	15	–
Is a method of masking the patient to the intervention allocated stated?	–	–	–	15	–
Were outcome assessors blind to intervention allocation?	–	–	–	15	–
Are baseline values for groups compared?	–	–	–	15	–
5. Duration and completeness of follow-up					
Are intervals between surgery and follow-up assessment clearly stated?	13	2	–	–	–
Are reasons for loss of patients at follow-up stated?	6	8	–	–	–
Are those lost to follow-up compared to rest of sample?	–	9	–	4	2
Is there an appropriate length of follow-up?	14	1	–	–	–
6. Statistical and analytical considerations					
Has the study sample size been justified?	–	15	–	–	–
Are the data clearly presented?	6	9	–	–	–
Was the data analyst masked to interventions?	–	–	–	15	–
Has type of statistical test and actual probability value been stated?	4	2	–	7	2
Are statistical tests appropriate to study?	3	3	–	5	4
Is the sample on which outcomes are assessed adequate?	9	4	1	–	–
Are conclusions justified by evidence?	10	3	–	–	2

TABLE 14 *Summary scores for methodological quality of studies of THR*

Criteria	RCTs	Comparative observational studies	Single-prosthesis observational studies
	Positively scored (%)	Positively scored (%)	Positively scored (%)
Clarity of study question and definition of outcome	54/60 (90)	49/60 (82)	68/75 (91)
Description of prosthesis and fixation method	22/24 (92)	18/24 (75)	29/30 (97)
Description of study sample	34/60 (57)	19/60 (32)	18/59 (31)
Control of bias in study design	18/48 (38)	0/17 (0)	na* (–)
Duration and completeness of follow-up	29/47 (62)	23/46 (50)	33/56 (59)
Statistical and analytical considerations	35/78 (45)	36/83 (43)	32/78 (41)
Total scores	**198/317 (62)**	**145/290 (50)**	**180/298 (60)**

* *na = not applicable*

Single prosthesis studies

In summary, the 15 studies of single prostheses have merits and faults almost identical to those of the comparative observational studies. They had particularly poor scores for a number of items in relation to description of the study sample and statistical and analytical considerations *(Tables 13* and *14)*.

Conclusion

The original purpose of this methodological review was to take account of the quality of evidence that might influence inferences made about the relative effectiveness of prostheses. As it became clear that hardly any strong inferences could be made about any prostheses, except informally with regard to Charnley, it was not necessary to stratify evidence for prostheses by the quality or rigour of the evidence. Nevertheless this review provides strong support for the view that there is no strong evidence base for choice of hip prosthesis at present. There are very few RCTs and those that have been performed are based on small sample sizes and with a few exceptions have not observed outcomes over a sufficiently long time period. Numerous methodological limitations surround what constitutes the vast majority of the evidence (i.e. observational studies). Others researchers have noted the lack of appropriate survival analysis available accurately to estimate the relative performance of prostheses, and the virtually complete absence of measures of outcomes of THR that matter to patients.[3,153,154] This review draws attention to a number of other problems, particularly lack of power but also other design and analytical limitations.

Health Technology Assessment panel membership

This report was identified as a priority by the Acute Sector Panel.

Acute Sector Panel

Chair: Professor John Farndon, University of Bristol †

Professor Senga Bond, University of Newcastle-upon-Tyne †

Professor Ian Cameron, Southeast Thames Regional Health Authority

Ms Lynne Clemence, Mid-Kent Health Care Trust †

Professor Francis Creed, University of Manchester †

Professor Cam Donaldson, University of Aberdeen

Mr John Dunning, Papworth Hospital, Cambridge †

Professor Richard Ellis, St James's University Hospital, Leeds

Mr Leonard Fenwick, Freeman Group of Hospitals, Newcastle-upon-Tyne †

Professor David Field, Leicester Royal Infirmary †

Ms Grace Gibbs, West Middlesex University Hospital NHS Trust †

Dr Neville Goodman, Southmead Hospital Services Trust, Bristol †

Professor Mark P Haggard, MRC †

Mr Ian Hammond, Bedford & Shires Health & Care NHS Trust

Professor Adrian Harris, Churchill Hospital, Oxford

Professor Robert Hawkins, University of Bristol †

Dr Gwyneth Lewis, Department of Health †

Dr Chris McCall, General Practitioner, Dorset †

Professor Alan McGregor, St Thomas's Hospital, London

Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London

Professor Jon Nicholl, University of Sheffield †

Professor John Norman, University of Southampton

Dr John Pounsford, Frenchay Hospital, Bristol †

Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham †

Professor Gordon Stirrat, St Michael's Hospital, Bristol

Dr William Tarnow-Mordi, University of Dundee

Professor Kenneth Taylor, Hammersmith Hospital, London

Diagnostics and Imaging Panel

Chair: Professor Mike Smith, University of Leeds †

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London *

Professor Andrew Adam, UMDS, London †

Dr Pat Cooke, RDRD, Trent Regional Health Authority

Ms Julia Davison, St Bartholomew's Hospital, London †

Professor Adrian Dixon, University of Cambridge †

Mr Steve Ebdon-Jackson, Department of Health †

Professor MA Ferguson-Smith, University of Cambridge †

Dr Mansel Hacney, University of Manchester

Professor Sean Hilton, St George's Hospital Medical School, London

Mr John Hutton, MEDTAP International Inc., London

Professor Donald Jeffries, St Bartholomew's Hospital, London †

Dr Andrew Moore, Editor, *Bandolier* †

Professor Chris Price, London Hospital Medical School †

Dr Ian Reynolds, Nottingham Health Authority

Professor Colin Roberts, University of Wales College of Medicine

Miss Annette Sergeant, Chase Farm Hospital, Enfield

Professor John Stuart, University of Birmingham

Dr Ala Szczepura, University of Warwick †

Mr Stephen Thornton, Cambridge & Huntingdon Health Commission

Dr Gillian Vivian, Royal Cornwall Hospitals Trust †

Dr Jo Walsworth-Bell, South Staffordshire Health Authority †

Dr Greg Warner, General Practitioner, Hampshire †

Methodology Panel

Chair: Professor Martin Buxton, Brunel University †

Professor Anthony Culyer, University of York *

Dr Doug Altman, Institute of Health Sciences, Oxford †

Professor Michael Baum, Royal Marsden Hospital

Professor Nick Black, London School of Hygiene & Tropical Medicine †

Professor Ann Bowling, University College London Medical School †

Dr Rory Collins, University of Oxford

Professor George Davey-Smith, University of Bristol

Dr Vikki Entwistle, University of Aberdeen †

Professor Ray Fitzpatrick, University of Oxford †

Professor Stephen Frankel, University of Bristol

Dr Stephen Harrison, University of Leeds

Mr John Henderson, Department of Health †

Mr Philip Hewitson, Leeds FHSA

Professor Richard Lilford, Regional Director, R&D, West Midlands †

Mr Nick Mays, King's Fund, London †

Professor Ian Russell, University of York †

Professor David Sackett, Centre for Evidence Based Medicine, Oxford †

Dr Maurice Slevin, St Bartholomew's Hospital, London

Dr David Spiegelhalter, Institute of Public Health, Cambridge †

Professor Charles Warlow, Western General Hospital, Edinburgh †

* Previous Chair
† Current members

continued

continued

Pharmaceutical Panel

Chair: Professor Tom Walley, University of Liverpool †

Professor Michael Rawlins, University of Newcastle-upon-Tyne*

Dr Colin Bradley, University of Birmingham

Professor Alasdair Breckenridge, RDRD, Northwest Regional Health Authority

Ms Christine Clark, Hope Hospital, Salford †

Mrs Julie Dent, Ealing, Hammersmith & Hounslow Health Authority, London

Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Tim Elliott, Department of Health †

Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire

Dr Felicity Gabbay, Transcrip Ltd †

Dr Alistair Gray, Health Economics Research Unit, University of Oxford †

Professor Keith Gull, University of Manchester

Dr Keith Jones, Medicines Control Agency

Professor Trevor Jones, ABPI, London †

Ms Sally Knight, Lister Hospital, Stevenage †

Dr Andrew Mortimore, Southampton & SW Hants Health Authority †

Mr Nigel Offen, Essex Rivers Healthcare, Colchester †

Dr John Posnett, University of York

Mrs Marianne Rigge, The College of Health, London †

Mr Simon Robbins, Camden & Islington Health Authority, London †

Dr Frances Rotblat, Medicines Control Agency †

Mrs Katrina Simister, Liverpool Health Authority †

Dr Ross Taylor, University of Aberdeen †

Dr Tim van Zwanenberg, Northern Regional Health Authority

Dr Kent Woods, RDRD, Trent RO, Sheffield †

Population Screening Panel

Chair: Professor Sir John Grimley Evans, Radcliffe Infirmary, Oxford †

Dr Sheila Adam, Department of Health*

Ms Stella Burnside, Altnagelvin Hospitals Trust, Londonderry †

Dr Carol Dezateux, Institute of Child Health, London †

Dr Anne Dixon Brown, NHS Executive, Anglia & Oxford †

Professor Dian Donnai, St Mary's Hospital, Manchester †

Dr Tom Fahey, University of Bristol †

Mrs Gillian Fletcher, National Childbirth Trust †

Professor George Freeman, Charing Cross & Westminster Medical School, London

Dr Mike Gill, Brent & Harrow Health Authority †

Dr JA Muir Gray, RDRD, Anglia & Oxford RO †

Dr Anne Ludbrook, University of Aberdeen †

Professor Alexander Markham, St James's University Hospital, Leeds †

Professor Theresa Marteau, UMDS, London

Dr Ann McPherson, General Practitioner, Oxford †

Professor Catherine Peckham, Institute of Child Health, London

Dr Connie Smith, Parkside NHS Trust, London

Dr Sarah Stewart-Brown, University of Oxford †

Ms Polly Toynbee, Journalist †

Professor Nick Wald, University of London †

Professor Ciaran Woodman, Centre for Cancer Epidemiology, Manchester

Primary and Community Care Panel

Chair: Dr John Tripp, Royal Devon & Exeter Healthcare NHS Trust †

Professor Angela Coulter, King's Fund, London *

Professor Martin Roland, University of Manchester *

Dr Simon Allison, University of Nottingham

Mr Kevin Barton, East London & City Health Authority †

Professor John Bond, University of Newcastle-upon-Tyne †

Ms Judith Brodie, Age Concern, London †

Dr Nicky Cullum, University of York †

Professor Shah Ebrahim, Royal Free Hospital, London

Mr Andrew Farmer, Institute of Health Sciences, Oxford †

Ms Cathy Gritzner, The King's Fund †

Professor Andrew Haines, RDRD, North Thames Regional Health Authority

Dr Nicholas Hicks, Oxfordshire Health Authority †

Professor Richard Hobbs, University of Birmingham †

Professor Allen Hutchinson, University of Sheffield †

Mr Edward Jones, Rochdale FHSA

Professor Roger Jones, UMDS, London

Mr Lionel Joyce, Chief Executive, Newcastle City Health NHS Trust

Professor Martin Knapp, London School of Economics & Political Science

Dr Phillip Leech, Department of Health †

Professor Karen Luker, University of Liverpool

Professor David Mant, NHS Executive South & West †

Dr Fiona Moss, Thames Postgraduate Medical and Dental Education †

Professor Dianne Newham, King's College London

Professor Gillian Parker, University of Leicester †

Dr Robert Peveler, University of Southampton †

Dr Mary Renfrew, University of Oxford

Ms Hilary Scott, Tower Hamlets Healthcare NHS Trust, London †

* Previous Chair
† Current members

Health Technology Assessment 1998; Vol. 2: No. 20

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.soton.ac.uk/~hta

ISSN 1366-5278