DEVELOPMENT OF ASSESSMENT IN HIP ARTHROPLASTY REVIEW

LINDSAY KAY SMITH

A thesis submitted as partial fulfilment of the requirements of the University of the West of England, Bristol for the degree of Doctor of Philosophy

Faculty of Health and Life Sciences

December 2011

Table of contents

Chapter 1 Introduction

Chapter 2 Background

Chapter 3   Image Interpretation

Chapter 4   Development of a radiographic assessment tool

Chapter 5 Clinical study

Chapter 6   General discussion

Chapter 7   Conclusions   165

References   167

Appendices   190

List of Tables

List of tables continued…

List of Figures

Objectives

The overriding aim of this PhD thesis was to develop research skills through studies which would be of benefit to others. The research study was entitled, 'Development of assessment in hip arthroplasty review' and a number of objectives were addressed. An initial objective was to establish the current methods by which a failing total hip arthroplasty is assessed in order to inform the process of review. The second objective was to demonstrate a method by which the skill of interpretation of x-ray images of hip arthroplasty could be acquired in order to inform the training of future non-medical health professionals involved in arthroplasty review. The third objective employed basic research skills to develop a clinical tool for use in combination with x-rays in hip arthroplasty review. The final objective was to conduct a clinical study which explored the relationship between some of the tools commonly used in the assessment of hip arthroplasty in order to add to the scientific evidence base for arthroplasty review.

Abstract

This thesis describes the development of aspects of assessment in hip arthroplasty review. Although many hip replacements provide pain relief and improved function, periodic review is advised to assess the state of the joint in order to identify a failing hip arthroplasty.

A literature search was conducted to establish methods of assessing failing hip arthroplasty and the findings are summarised. There was a lack of standardisation but an emphasis on the need for review because failing hip arthroplasty is frequently asymptomatic.

The review process may be conducted by medical or non-medical members of the orthopaedic team. A lack of formal educational programmes in arthroplasty review has led to innovative ways of non-medical health professionals achieving the required competency. One of these methods is described to show the development of a skill in radiographic image interpretation for hip replacements. Image interpretation is an important component of hip arthroplasty review and includes the measurement of osteolytic lesions, a phenomenon generally considered to be caused by the wear particles produced from the articulating surfaces of the artificial joint. A simple, radiographic tool was developed to measure these irregularly shaped lesions and its testing is described. The tool was found to be valid and reliable when used interchangeably between any professional who is part of an orthopaedic team conducting hip arthroplasty review.

Finally, a clinical study of hip arthroplasty was conducted to explore the relationship between changes in a patient reported outcome measure (PROM) and radiographic changes over the same period of time. All study participants had a hip replacement approximately seven years earlier and were at a stage (mid-term) when signs of deterioration of the hip joint often appear. The results showed that radiographic changes were not predicted by changes in the PROM over the same period.

This thesis illustrates a training model for non-medical health professionals to acquire the skill of radiographic image interpretation and employs basic research to develop a simple and reliable radiographic tool for use in hip arthroplasty review. It shows that, for patients reviewed at mid-term, it is important to include an x-ray because a joint-specific PROM is not able to predict the radiographic changes around a hip replacement. This information adds to the scientific evidence for assessment in hip arthroplasty review. It is of potential benefit to patients through the improvement of current surveillance methods and future planning of hip arthroplasty review.

Acknowledgements

Terms and abbreviations

§	Section
ACPA	Arthroplasty Care Practitioners' Association
ALVAL	Aseptic lymphocyte-dominated vasculitis associated lesion
ANOVA	Analysis of variance
A-P	Antero-posterior
ARMD	Adverse reaction to metal debris
BMS	Between subjects means square
BOA	British Orthopaedic Association
CASP	Critical appraisal skills programme
CI	Confidence interval
cms	Centimetres
cm^2	Centimetres squared
CSP	Chartered Society of Physiotherapy
CT	Computed tomography
°	Degrees
DH	Department of Health
DOPS	Direct observation of procedural skills
EMS	Mean square experimental error
EQ-5D	EuroQol 5 dimension questionnaire
ESP	Extended scope practitioner
EWBL	Evidencing work based learning
HEI	Higher education institution
HHS	Harris Hip Score
ICC	Intraclass correlation coefficient
IL1	Interleukin 1
IRMER	Ionising radiation (Medical exposure) regulations
JMS	Between observers mean square

MARS	Metal artefact reduction sequencing MRI
MDC	Minimal detectable change
mm	Millimetres
mm^2	Millimetres squared
MRI	Magnetic resonance imaging
NHS	National Health Service
NJR	National Joint Registry of England and Wales
nm	nanometres
OHS	Oxford Hip Score
PROM	Patient reported outcome measure
RANK	Receptor activator of nuclear transcription factor
RANKL	Receptor activator of nuclear transcription factor-kappa B ligand
RCN	Royal College of Nurses
RCR	Royal College of Radiologists
SEM	Standard error of the measurement
SF-12	Short form 12 questionnaire
SF-36	Short form 36 questionnaire
SW	South West
SWOT	Strengths, Weaknesses, Opportunities, Threats
THA	Total hip arthroplasty
TNFα	Tumour necrosing factor alpha
TTO	Time trade off
UHMWPE	Ultrahigh molecular weight polyethylene
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
WGH	Weston General Hospital
WOMAC	Western Ontario and McMaster University Osteoarthritis index
WMS	Within subjects means square
μm	micrometres

Hip-op Rappity Rap

> I'm sitting here on the 14th floor
> In the limb replacement corridor
> My X-ray of some months before
> Confirms just why this leg's so sore
> And I've begged the femur specialist
> To eradicate my starboard list
> 'Please operate, make this ache desist'
> He said, 'Take these pills, join the waiting list'

He put me down / he put me down / he put me down / he put me down
For a

Hip-op rappity rap
I can't stand straight, I'm a lop-sided chap
Hip-op rappity roo
My career is over as a kangaroo
Hip op rappity squeaks
It's not much fun when your pelvis creaks
Hip-op rappity ouch
Only three more years of pain to go,
Only three more years of pain to go...

> They'd have to treat me quick if I had something mean
> Like Egyptian typhoid or a ruptured spleen
> But all I've got is a permanent lean
> I make Long John Silver look like Torvill and Dean
> They said 'If you go private there's a room for you
> With colour television and a tree-top view
> Or why not do it yourself at B & Q
> With a saw and a ratchet and some super glue?'

I couldn't pay / I couldn't pay / I couldn't pay / I couldn't pay
For a

Hip-op rappity rap …

By Stewart Henderson (©1989, used with permission)

Chapter 1

Introduction

The 'Development of assessment in hip arthroplasty review' is a study about total hip replacement. This is an orthopaedic procedure in which the natural hip is replaced with a man-made one in order to reduce pain and improve function for the recipient.

1.1 Hip arthroplasty

1.1.1 Hip anatomy

The hip joint is a 'ball and socket' joint in which the round head of the femoral bone articulates with the concave socket of the acetabulum in the pelvis (see Figure 1.1). The bony surfaces are covered with articular cartilage and the acetabulum is extended by a fibro cartilaginous rim (acetabular labrum) producing a close fitting joint. It is lubricated by synovial fluid which is contained by the capsule surrounding the joint. Internally, the capsule is lined by synovial membrane and externally, it is intimately blended with three ligaments which reinforce the stability of the joint (Warwick and Williams 1973).

The shape of the hip joint allows movement in three planes produced by muscles originating on the innominate bone (hemi pelvis) or in the lumbar region and inserting on the femur, the long bone of the thigh. The movements are termed flexion-extension from anterior to posterior (forwards and backwards), abduction-adduction (side to side) and internal and external

Rotation (Warwick and Williams 1973). Movement at the hip is an essential component of ambulation and other everyday activities.

Figure 1.1. Illustration of the right hip joint
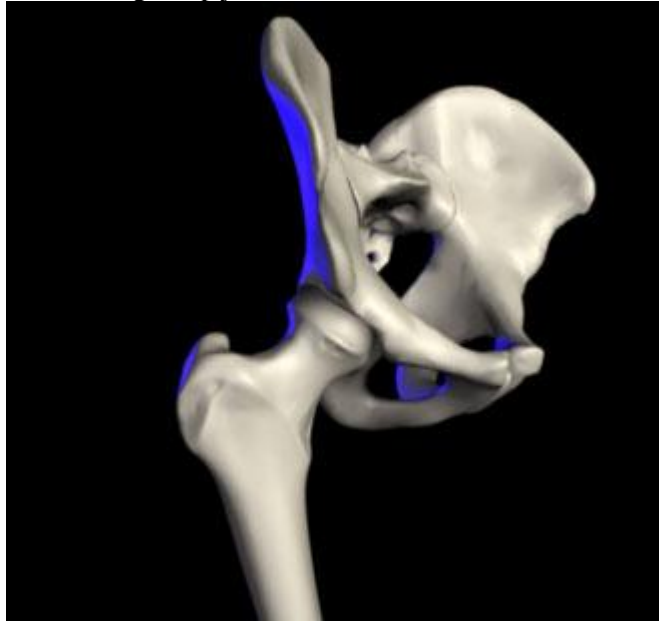
Image used under the terms of the GNU Free Documentation License Version 1.2.

1.1.2 Total hip replacement

Total hip replacement, also known as total hip arthroplasty (THA), is usually an elective procedure in which an orthopaedic surgeon will replace the articulating surfaces of the joint. A routine procedure will involve dislocation of the hip, then removal of the femoral head and is followed by preparation of the acetabulum to receive an artificial cup; the shaft of the femur is then prepared to receive a suitably sized femoral stem (NHS Choices Information 2010). The joint is re-assembled with trial prostheses in situ to check the articulation before the final prostheses are inserted, with or without cement.

1.1.3 Indications for total hip replacement

The initial indications of a deteriorating hip joint are pain in and around the joint, sometimes referred to the knee, and a loss of function. These symptoms prompt a person to seek medical

help. The symptoms are most often due to degenerative changes within the joint with a loss of synovial cartilage and damage to the underlying bone (Crawford Adams and Hamblen 2001). The symptoms are sometimes due to a disease process such as rheumatoid arthritis or a metabolic disorder affecting the quality of the bone; sometimes there is a history of congenital or developmental abnormalities (Crawford Adams and Hamblen 2001). Trauma to the pelvis or in the upper femoral region may also require treatment with a THA.

1.1.4 History of total hip replacement

The use of THA first became popular in the 1970's following extensive work by Sir John Charnley at the Wrightington Hospital in Lancashire, England (Brand 2010). The design most commonly used at that time was a metal femoral head and stem with a polyethylene acetabular cup. The head and stem were of one piece, and both the femoral stem and the acetabular cup were cemented into place (Crawford Adams and Hamblen 2001). Since that time, many new designs of THA and new materials have been introduced, some more successful than others (Maloney et al. 1999, Furnes et al. 2001, Norton et al. 2002, Ong et al. 2002, Callaghan et al. 2004, Utting et al. 2008, Wroblewski et al. 2009, Huddleston et al. 2010).

1.1.5 Types of THA

The different types of THA currently available can be broadly grouped into four categories – cemented, uncemented, hybrid and reverse hybrid (National Joint Registry 2010). Cemented THA consists of a femoral stem and an acetabular cup both cemented into place (see Figure 1.2) whereas an uncemented THA involves no cement. The hybrid THA has a cemented femoral stem and an uncemented acetabular cup, and the reverse hybrid is the other way around with only the acetabular cup cemented into place.

The materials used for the femoral component consist of a metallic stem and either a metallic or ceramic head which may be of one piece with the stem or modular (Crawford Adams and Hamblen 2001). The femoral component can be cemented or uncemented. If the latter, it will be designed to fit the medullary cavity in the femur and may be coated with materials which encourage the surrounding bone to grow onto it (Della Valle and Paprosky 2002).

There is considerable variation in the acetabular cup: it may be all polyethylene, which is cemented into place or an uncemented metallic shell with a separate liner of ceramic or polyethylene material (McWilliams and Parker 2008). The shells are often coated to encourage bony ongrowth, as with an uncemented stem. There are also metallic acetabular cups which are designed to articulate with a metal femoral head to create a metal-on-metal bearing (Huo et al. 2010). There are many different THA systems on the market, each manufacturer using a slightly different design or material (National Joint Registry 2010). For suitably selected patients, there is an option for a hip resurfacing procedure (see Figure 1.2) which preserves more of the patient's bone (Huo et al 2010), and an emerging range of short stemmed femoral components in THA with the same objective in mind.

Figure 1.2. An x-ray image of a cemented metal on polyethylene total hip replacement in the right hip and a resurfacing hip prosthesis in the left hip
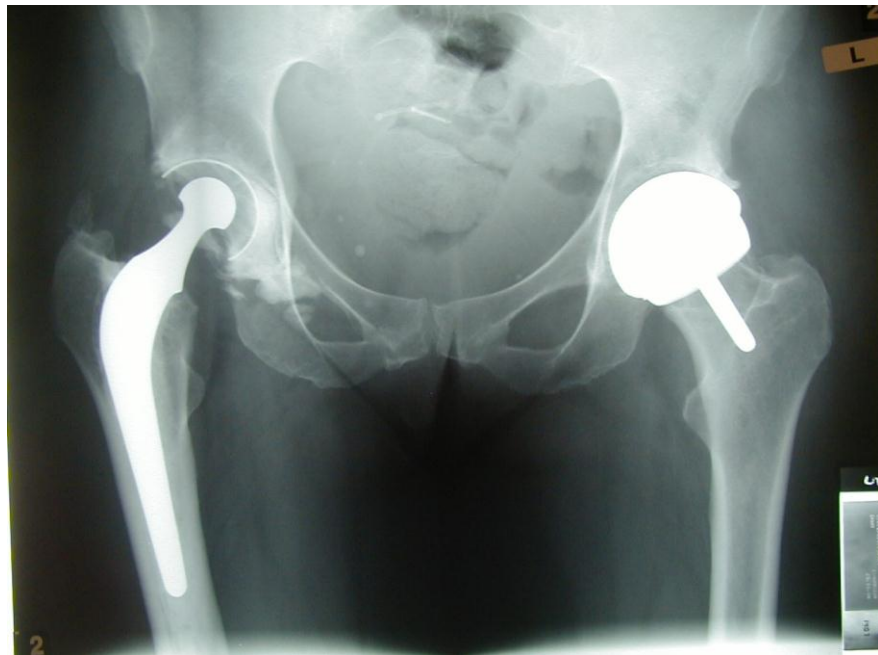
1.1.6 Failure of THA

The replacement of a hip joint with man-made components has become a highly successful procedure which can restore the patient to a more active lifestyle than pre-operatively and consequently, improve their general health (Malchau et al. 2005). There is a small percentage of patients who do not experience the immediate relief from pain that is expected but most patients with THA will get many years of use from the joint (Soderman et al. 2001, National Joint Registry 2010). However, in time it may begin to loosen or there may be changes in the surrounding bone which threaten the integrity of the artificial joint.

At present, there are more than 65,000 primary hip replacements performed annually in England and Wales alone and approximately ten percent will require revision surgery during the life time of the recipient (National Joint Registry 2010). Some of these patients will experience symptoms which prompt them to seek medical help but others will be unaware of the changes that are taking place until there is significant underlying deterioration of the bone (British Orthopaedic Association 2006). Hip arthroplasty review is designed to monitor the state of the hip replacements beyond the immediate post-operative period in order to identify patients who may need revision surgery.

1.2 Current health environment

1.2.1 National Health Service

In the United Kingdom (UK), the National Health Service was set up in 1948 to provide healthcare for all. It is currently undergoing major changes in an effort to reduce costs as the economic burden of providing for a growing population with greater choice of treatments becomes unsustainable. There has been a suggestion of reducing the availability of elective procedures, including hip and knee replacements, which has provoked comment from the President of the Royal College of Surgeons of England who emphasised that introducing delay for such successful procedures would increase future health problems (Campbell 2011).

In addition to the reduction in available funds, the delivery of care by medical staff in the UK has been further impeded by a European directive (Department of Health 2004a). This resulted in a reduction of working hours and has particularly impacted the junior doctors and the support that they can provide to their senior colleagues. Junior orthopaedic doctors are an essential part of the orthopaedic team and as such, contribute to the delivery of surgical interventions. There is an increasing need for other health professionals to perform some of the tasks traditionally undertaken by these doctors.

1.2.2 Regulatory requirements

Against the background of limited resources, there is increased pressure for individual orthopaedic units to produce good results as evaluated with patient reported outcome measures (PROMs). In April 2009 it became mandatory in England to collect pre- and post-operative PROMs for all total hip and total knee replacements (Department of Health 2008). This national audit has been introduced to monitor and compare orthopaedic units on their delivery of common elective procedures based on the patient reports of progress at six months after the primary surgery. This information has recently become widely available, thereby increasing pressure on hospital units to be able to interpret their results in a meaningful way (The NHS Information Centre 2011). This requires knowledge of the local patient population and ongoing review of arthroplasty patients to monitor the long term outcomes.

1.2.3 Joint registers

The Swedish Arthroplasty Register has been in existence since 1979 and other countries have subsequently established registries including, more recently, the United States of America (USA) (Malchau et al. 2005, American Association of Orthopedic Surgeons 2010, Karrholm 2010). These registries collect data at the time of the initial surgery and from any subsequent surgery but not in the intervening period. Data from registries provides a rich source of high quality, evidence-based information on medical treatment. There are now efforts to combine data from multiple registries for analysis although the originator of the Swedish Registry has expressed some doubts about pooling the data from a range of countries with different languages and lack of a single robust system for tracking citizens (Rapp 2011).

In an attempt to address the lack of information between primary surgery and revision, the Swedish Registry piloted a project to collect PROMs and radiographic data at six and ten years after the joint replacement (Malchau et al. 2005). However, the compliance was poor and so the project was abandoned (Karrholm 2011). This reinforces the importance of arthroplasty review as a means of collecting data between primary and revision surgery to add to the scientific evidence base for this treatment.

1.3 Current practice

1.3.1 Assessment

The current recommendations for long term follow up of THA in the UK are that each patient should be seen at one and five years, and at further five year intervals (British Orthopaedic Association 2006). These recommendations suggest that the assessment should include a history, a clinical examination and an x-ray of the hip but a survey of UK hospitals suggested that practice is varied (Darrah 2006). Many of the larger orthopaedic units in England and Wales have conducted some long term follow up but it has not usually taken place in smaller units, although there are exceptions in both cases. In Scotland, the process of follow up is well established and information relating to any hospital episode is recorded. Data are then extracted and used for analysis by the Scottish Arthroplasty Project (NHS National Services Scotland 2010).

The use of 'virtual surveillance' is a new method for arthroplasty review in England which has only been established in a few orthopaedic units, one of which is the University Hospital Coventry (Hugill et al. 2010). With this system, THA patients are sent a questionnaire and asked to attend a local hospital for an x-ray of their hip. The information from the questionnaire is reviewed in conjunction with the x-ray but without the patient present. If the results imply that there may be adverse changes or if they are unclear, the patient is given an appointment to attend an orthopaedic clinic.

1.3.2 Health professionals

Historically, follow-up of THA patients was done by medical members of the orthopaedic team but the pressure of reduced medical hours plus increased workload has resulted in non-medical health professionals being involved in the assessment process (British Orthopaedic Association 2006, Aiken et al. 2009). However, the training of these staff has not been standardised and often reflects local need. In response to this, the Arthroplasty Care Practitioners' Association (ACPA) was formed in 2006 with support from the British Hip Society. This national group of health professionals (predominantly physiotherapists and orthopaedic nurses) has worked with the

CHAPTER 1. INTRODUCTION

Department of Health, the British Orthopaedic Association, the British Hip Society and the British Association for Surgery of the Knee to define the skills needed for health care professionals working in arthroplasty (British Orthopaedic Association 2010). The definitions were needed to reflect the advanced level at which these practitioners are working, beyond their basic professional training. The British Orthopaedic Association has now established a formal alliance with ACPA with the intention that there will be representatives of both organisations involved in discussions about the future training of practitioners and care of arthroplasty patients (British Orthopaedic Association 2010).

The development of non-medical health professionals into arthroplasty practitioners is increasing both in the UK and elsewhere (Aiken et al. 2007). ACPA has recently hosted visits by physiotherapists from Australia and Canada who have subsequently returned to their respective countries to start arthroplasty review services (see Appendix I). The role of arthroplasty practitioners is expanding and they can be found in pre-operative, peri-operative and post-operative clinical situations. They maintain their professional affiliation but work in an extended scope of practice alongside medical staff. Further work is needed to develop the training of future arthroplasty practitioners.

1.4   Purpose of the study

The aim of the 'Development of assessment in hip arthroplasty review' was to add to the scientific evidence base for hip arthroplasty review. The background was the need for surveillance of an increasing number of total hip replacements being performed annually and the reduced capacity of the orthopaedic surgeons to provide this service. There was no standardisation of the methods used for arthroplasty surveillance at the time and there was widespread evidence of non-medical health professionals performing this work.

The aim of the research was addressed through a number of objectives. Initially, an objective was to establish the current methods used to assess failing THA as part of arthroplasty review. The second objective was to illustrate the academic principles underlying acquisition of a skill in interpretation of x-ray images of THA by a non-medical health professional. The third objective employed basic research skills to develop a clinical tool for use in combination with x-rays in hip arthroplasty review. The final objective was to conduct a clinical study which explored the relationship between some of the tools commonly used in the assessment of THA in order to add to the scientific evidence base for arthroplasty review.

Each of the objectives was addressed through a chapter in the thesis, as summarised below.

Chapter 2: Background

A search was undertaken of current orthopaedic literature related to failing hip arthroplasty and the results were discussed with reference to hip arthroplasty review.

Chapter 3: Image Interpretation

The method by which a non-medical health professional acquired the skill of image interpretation in arthroplasty was described and discussed with reference to training of future arthroplasty practitioners.

Chapter 4: Development of a radiographic tool

The concept of a morphometric grid was developed into a simple clinical tool for the measurement of osteolytic lesions seen on x-ray images of hip arthroplasty. Initial testing of the reliability and validity of this tool was described.

Chapter 5: Clinical study

A clinical study was conducted in which a cohort of patients with hip arthroplasty was reviewed to explore the relationship between changes in a patient reported outcome measure and the changes seen on x-ray images.

Chapter 6: Discussion

Implications of the research are discussed with suggestions for further research

1.4.1 Potential value of the research

There was potential value in this work to improve patient outcomes, to add to the scientific base for arthroplasty review and to facilitate future research. The potential benefit to the patient was through support for arthroplasty review in order to provide an option for early revision surgery where appropriate. In addition, the development of a clinical tool was potentially beneficial by providing a visual method to demonstrate adverse changes on x-ray, thus increasing a patient's understanding and involvement in any decision about further treatment.

One of the contributions to the scientific base was the development of a clinical tool to be used by orthopaedic health professionals with x-ray images of THA. A further contribution was the exploration of relationships between PROMs and x-ray images to increase the scientific evidence for the criteria used in arthroplasty review. In addition, the documented evidence of skill acquisition by a non-medical health professional provided a framework for comparison with alternative training methods to develop the skills of future arthroplasty practitioners.

Finally, the study had potential to facilitate future research for the benefit of patients, health professionals and health providers. The information gathered through the individual chapters generated a number of different ideas for future research which are discussed in Chapter 6.

Chapter 2

Background

This chapter explores the current information about failing hip replacements and the assessment process. It includes an overview of the different methods by which THA can fail and detailed discussion of these mechanisms in an aseptic (non-infected) hip. A literature review of assessment methods for failing hip replacements was conducted and the results are presented and discussed.

2.1 Failing hip replacement

2.1.1 Types of failure

The success of a total hip replacement is dependent on a variety of factors which include the patient's expectations, the diagnosis and comorbidities, the surgeon's skill, the type of prosthesis and the subsequent stresses on the replaced joint (Duffy et al. 2005, Mancuso et al. 2009). For many patients, the relief from pain and improvement in function achieved after surgery will satisfy their expectations and they will have many years of use from the hip arthroplasty (Duffy et al. 2005). However, a proportion of patients will continue to experience problems with the operated hip after surgery or develop problems over time and careful evaluation is required to assess the need for further treatment. In some cases, the treatment will be a revision hip replacement.

The indications for revision surgery provide a useful insight into why a THA fails (see Table 2.1). The results shown below are from the Swedish Hip Replacement Register (Malchau et al 2002) and the Australian National Joint Replacement Register (Australian Orthopaedic Association 2010). In these registries, surgeons may record one reason for revision in each case whereas in the National Joint Registry (NJR) of England and Wales, a surgeon is permitted to enter more than one reason for revision which prevents a direct comparison. The summary produced by the NJR for the years 2006 to 2009 (25222 revisions in total) showed that 70% of revision procedures included aseptic loosening or lysis as a reason, 23% listed 'pain' and 7% were under the category of infection (National Joint Registry 2010). Revision for aseptic loosening or lysis is the largest category in all registries with over 50% of revision surgery being completed for this reason. The symptoms associated with aseptic loosening develop over a period of time and may not be associated with pain whereas the other indications listed for revision tend to be symptomatic. Further detail of aseptic failure will be presented in a subsequent section as it is the central phenomenon of arthroplasty review in the longer term.

Table 2.1. Indications for hip revision procedures from the Swedish and Australian joint registries (Malchau et al. 2002, Australian Orthopaedic Association 2010)

Diagnosis	Australia	Sweden
	2004-2009 (% of 31335 hips)	1979-2000 (% of 14081 hips)
Aseptic loosening/lysis	55.4	75.3
Pain	1.7	0.3
Dislocation/subluxation	14.4	5.8
Periprosthetic fracture	9.0	5.1
Infection	11.7	7.6
Malalignment	0.4	
Fractured component	2.1	1.5
Incorrect sizing/technical error	0.1	3.0
Wear acetabulum	2.7	0.9
Metal sensitivity	0.4	
Leg length discrepancy	0.3	
Instability	0.3	
Acetabular erosion	0.3	
Avascular necrosis	0.1	
Heterotopic bone	0.1	
Tumour	0.1	
Other	0.8	0.4
Missing		<0.1

The lower proportion of revision surgery for aseptic loosening in Australia than in Sweden or England and Wales may be due to the wider options available to surgeons in Australia when recording the category for revision and the larger proportions of revision performed for other reasons, such as dislocation and infection. This, in turn, may be explained by differences such as procedural preferences and the type and make of prostheses commonly used in the relevant country during the specific time period analysed. Further exploration of the underlying data would be required to fully understand these differences.

The category listed as 'infection' includes both primary deep infection and secondary infection introduced by another route at a later time. A high standard of perioperative care is essential to ensure that the infection risk is minimized, and the use of clean air operating rooms with vertical laminar flow, prophylactic antibiotics and an aseptic surgical technique are well established methods (Archibeck et al. 2001). An infection may be superficial, affecting only the wound and surrounding tissues, or deep, when it affects the prostheses of the joint replacement. Antibiotic treatment is used to treat a superficial wound infection but a deep infection frequently requires a revision of the THA although long term antibiotic treatment is sometimes employed, particularly in the elderly and medically unfit.

From the patient's perspective, realistic pre-operative expectations are important for satisfaction with the outcome although unrealistic expectations will not cause a hip replacement to fail (Duffy et al. 2005, Huo et al. 2010). If there is a complaint of persistent pain or poor function after THA, further investigation may be necessary. This will include some or all of the following: obtaining a thorough medical history from the patient, physical examination, appropriate imaging studies (plain x-ray, nuclear imaging, ultrasound, computed tomography [CT] and/or magnetic resonance imaging [MRI]) and laboratory tests of blood and joint aspirate (Duffy et al. 2005). Ultimately, if non-surgical management is unsuccessful, some patients are offered a revision of their THA for unexplained pain because no other source of the problem can be identified (Malchau et al. 2002, Wroblewski et al. 2007, National Joint Registry 2010).

The surgeon's skill and choice of implant are also essential to the success of a THA because the precise positioning and appropriate pairing of the components, with reference to the patient's needs, can affect the survival of the implant (Malchau et al. 2002, Dobzyniak et al. 2006, Wroblewski et al. 2007). Congenital or developmental conditions present challenges for

component selection and fixation and survival of the components is not as good as for other conditions (Chougle et al. 2005). Comorbidities may affect the surrounding bone and predispose it to fracture or loss of bone density with subsequent loosening of the components.

Fractures of the bone adjacent to the hip replacement (peri-prosthetic fracture) may occur intra-operatively or at any time postoperatively. The causes include low-impact trauma from a fall, particularly in the elderly, or high-impact trauma, as in a road traffic accident. In addition, osteopoenia, osteoporosis, osteolysis, unfavourable biomechanics and abnormal stresses on the bone all have the potential to cause a fracture. Depending on the site and severity, diagnosis may require CT and/or MRI in addition to plain radiographs, especially when the base (floor) of the acetabulum is involved (Cahir and Toms 2009).

Fracture of a component, either the acetabular cup or the femoral stem or head, is less common than peri-prosthetic fracture. Other than following a traumatic incident, stem fractures may occur in a distally fixed stem with lack of proximal fixation (secure at the lower end only) whereas fracture of the cup may follow excessive wear of the polyethylene (Toms et al. 2009). The use of ceramic femoral heads or ceramic liners for an acetabular cup provides another possible source of component fracture as this material is brittle. It has recently been found that acetabular cups of highly cross-linked ultrahigh molecular weight polyethylene are less fracture resistant than non cross-linked polyethylene and stress concentrations may lead to fracture failure (Furmanski et al. 2009).

Further indications for revision surgery of a mechanical nature include instability of a component, repeated dislocation or subluxation (incomplete dislocation) of the femoral component or residual discrepancy in leg length (Huo et al. 2010). Occasionally, malalignment of a prosthesis produces symptoms such as an iliopsoas tendonitis due to retroversion of the acetabular cup (Duffy et al. 2005). If conservative treatment fails, a revision operation may be required.

Failure of a hip replacement is often described in terms of when it occurs. Following discharge from hospital, the first four to five years are the early postoperative period. From five to six years after surgery until nine or ten years is 'mid-term' or the intermediate period (Jacobs et al. 2007, Lusty et al. 2007). Beyond ten years is often referred to as long term follow up (British Orthopaedic Association 2006, Horne et al. 2007). These terms are not precisely defined

as they have developed alongside the longevity of the prostheses. Reports are now available of some patients with implants still in place after 38 years, which shows that previous expectations of ten to fifteen years were conservative (Wroblewski et al. 2007).

2.1.2 Aseptic failure

Aseptic failure accounts for the largest proportion of revision in any joint registry and includes the sub-categories of aseptic loosening, wear of articulating surfaces and osteolysis. All these processes take place over time with the patient unaware of changes until they significantly threaten the survival of the joint (Zicat et al. 1995, Maloney et al. 1999, Duffy et al. 2005). They are not apparent in the early postoperative period but usually begin to appear at midterm (Malchau et al. 2005, Wroblewski et al. 2007, Hallan et al. 2010).

2.1.2.1 Aseptic loosening

Aseptic loosening refers to the debonding of a THA component from the surrounding bone and is a major cause of failure in hip replacements in the long term. It may be caused by repeated mechanical stresses on the joint or as a result of the production of wear particles (see §2.1.2.2). It is diagnosed using x-ray images to assess any migration of the prosthesis (movement from its original position in any direction) or development of radiolucencies not seen post-operatively (darkened areas adjacent to the component indicating absence of bone). Conventionally, a distance of 2mm width of radiolucency has been used by orthopaedic surgeons and radiologists as the defining measure of significance (O'Neill and Harris 1984, Duffy et al. 2005, Toms et al. 2009) and is still commonly used for both femoral and acetabular evaluation (Spangehl et al. 1999, Archibeck et al. 2001, Capello et al. 2003, Ito et al. 2004, Mabry et al. 2004, Perka et al. 2004, Incavo et al. 2008, Kim 2008, Utting et al. 2008). The 2mm measurement has become the critical distance as a non-progressive radiolucency of <2mm is usually attributed to a stable fibrous reaction between the bone and the cement whereas >2mm is more likely to be due to infection or aseptic loosening (McBride and Prakash 2011).

Radiolucencies may be located between the prosthesis and cement, or between the cement and the surrounding bone, or adjacent to an uncemented prosthesis. The first of these groups is of particular concern unless it is between the shoulder of the prosthesis and the cement in a femoral

stem which has been designed to subside (Williams et al. 2002, Duffy et al. 2005, Toms et al. 2009). A radiolucency that is progressing (widening or lengthening on serial x-rays) is of more concern than one which remains static. The presence of radiolucency alone does not confirm debonding but if the length is extensive and it is progressive, there is a high statistical association with loosening (Toms et al. 2009). The only definite indication of loosening is migration of the prosthesis (Duffy et al. 2005, Toms et al. 2009).

2.1.2.2 Wear

The mechanical articulation between a polyethylene acetabular cup or cup liner and a metal femoral head produces wear of the polyethylene due to the differential hardness. The wear particles generated migrate away from the articulation and into surrounding tissues stimulating an inflammatory response, the extent of which is dependent on the size, concentration and composition of the particles (Matthews et al. 2000). Polyethylene wear particles of less than a micrometre in size are thought to be particularly bio reactive as they stimulate the macrophage action of phagocytosis. This may lead directly to loosening due to the formation of granulomatous fibrous tissue around the joint (a pseudomembrane) as a result of a chronic inflammatory response (Holt et al. 2007). The other important biological response to wear debris particles is osteolysis of the adjacent bone (see §2.1.2.3).

The wear properties of polyethylene have improved over time with changes in the production treatment processes and with the type of sterilisation employed (Jacobs et al. 2007). The original polytetrafluorethylene was replaced by ultra high molecular weight polyethylene (UHMWPE) and more recently, highly cross-linked polyethylene has been introduced (McCalden et al. 2005, Jacobs et al. 2007). UHMWPE is subjected to ionizing radiation to achieve the cross-linking and the mid-term in vivo studies of this highly cross-linked polyethylene suggest that the wear rate is lower than for UHMWPE (Olyslaegers et al. 2008, McCalden et al. 2009). However, there are concerns that it produces larger volumes of smaller particles which may be more bio reactive and longer term studies are needed to assess the performance (Holt et al. 2007, Jacobs et al. 2007, Geerdink et al. 2009, Mu et al. 2009). Further developments have led to a polyethylene impregnated with vitamin E (by diffusion) to reduce the number of free radicals formed by the irradiation process used to produce the cross-linking. This

change may facilitate low wear rates and lower bio reactive response without compromising the other properties of the material (Jacobs et al. 2007, Jarrett et al. 2010). In a laboratory study, a vitamin E impregnated polyethylene was shown to demonstrate improved fracture resistance (Oral et al. 2008).

Production of wear particles is not unique to metal-on-polyethylene articulations. Hip replacements incorporating metal-on-metal or ceramic-on-ceramic materials will also produce wear particles although in much lower volumes (Holt et al. 2007). Alumina particles from ceramic wear may stimulate osteolysis but at a lower rate than when polyethylene is involved (Hernigou et al. 2009). Metal-on-metal particles are even smaller (20-90 nm) and these nanoparticles do not stimulate the extent of bony osteolytic response seen in metal-on-polyethylene articulations but other effects have been noted. In cases of excessive metallic wear, a local response to the increased level of metallic particles (metallosis) and/or metal ions may stimulate hypoxia/tissue necrosis and a periprosthetic soft tissue reaction with an increased fluid collection described as a pseudotumour (Pandit et al. 2008). There is evidence of a hypersensitivity type reaction in the periprosthetic tissues of some patients known as an aseptic lymphocyte-dominated vasculitis associated lesion (ALVAL) (Davies et al. 2005, Willert et al. 2005). The presence of elevated metal ions in whole blood may have other effects which are under investigation (Bhabra et al. 2009, Hart et al. 2009, Parry et al. 2010). These terms have recently been grouped together until more details of the underlying pathology are known and have been termed an 'adverse reaction to metal debris' (ARMD) (Langton et al. 2010).

2.1.2.3 Osteolysis

Osteolysis represents the destruction of bone, most commonly in response to excessive wear particles. The particles migrate away from the articulating surface via the effective joint space (Zicat et al. 1995) to periprosthetic bone where an inflammatory response is stimulated with release of inflammatory cytokines. These cytokines, in particular tumour necrosing factor α (TNFα) and interleukin1 (IL1), affect the normal balance of bone turnover which is carried out by the osteoblast cells (which synthesise new bone and bone matrix) and the osteoclasts (which are responsible for bone destruction and resorption). A process occurs at cellular level through cytokine stimulation of a substance known as RANKL (receptor activator of nuclear transcription

factor-kappa B ligand) from the surface of the osteoblasts. RANKL binds to the substance RANK (receptor activator of nuclear transcription factor-kappa B) found on osteoclast precursor cells which causes them to mature into an active cell or to stimulate active bone resorption. This leads to an imbalance of the local skeletal homeostasis with an increase of bone destruction and resorption known as osteolysis (Holt et al. 2007).

Two patterns of osteolysis are recognised – linear and expansile. The former develops along the prosthesis-bone interface, leading to aseptic loosening whereas expansile lesions are more balloon-shaped and extend away from the prosthesis (Dumbleton et al. 2002, Elke et al. 2003, Toms et al. 2009). It is these expansile lesions which threaten the stability of a THA with the potential for periprosthetic fracture (Paprosky et al. 2001, Kitamura et al. 2006b, Needham et al. 2008). Not all osteolytic lesions lead to failure of the THA but a progressive lesion is of concern and monitoring is highly recommended (Maloney et al. 1999, Dumbleton et al. 2002, British Orthopaedic Association 2006, Wroblewski et al. 2007). In the past, the measurement of lesion size was not standardised and many were reported by location only or with an approximation of size (Dumbleton et al. 2002). The introduction of a simple method for quantitative assessment of these lesions would greatly assist effective review of THA.

A link between wear rate of polyethylene and volume of osteolysis has been proposed. It is suggested that a wear rate of less than 0.1mm per year is unlikely to produce osteolysis and that at less than 0.05mm per year, osteolysis does not occur (Dumbleton et al. 2002). This led to the concept of a wear rate threshold. However, more recently, that has been challenged and the concept of dose-response relationship was introduced which includes the accumulation of wear debris over time as well as an annual wear rate (Emms et al. 2010). The implication is that osteolysis may develop around a prosthesis that has been in situ for many years even if the wear rate is low. Another challenge to the wear threshold is presented by the newer polyethylenes with a lower wear rate but increased bio-reactivity of the wear debris particles (see § 2.1.2.2).

2.1.3 Failure by component type

There are recognised patterns of failure associated with the different types of THA. The acetabular component, if cemented, commonly fails by aseptic loosening whereas uncemented cups are threatened by expansile osteolytic lesions although this is sometimes preceded by linear

aseptic loosening (Zicat et al. 1995, Dumbleton et al. 2002, Malchau et al. 2002). If asymmetry of the cup or cup liner is observed (eccentric wear), then osteolysis should be suspected as a reaction to the debris produced by the magnitude of the polyethylene wear (Dumbleton et al. 2002).

Uncemented femoral stems are designed to fill the proximal medullary cavity of the femur and seal it against migration of particulate wear debris via the effective joint space. If wear debris enters the femur around the prosthesis, loosening may be caused by osteolysis or by the formation of granulomatous fibrous tissue which prevents osseointegration (Zicat et al. 1995). The success of the effective 'sealing' determines the extent of the changes around the femoral component which, if present, may progress from a linear to an expansile pattern (Engh et al. 1987, Zicat et al. 1995). Aseptic loosening will occur if the prosthesis does not osseointegrate with the femur.

Cemented femoral stems are subject to aseptic loosening and osteolysis by the same mechanisms as uncemented stems. Particulate wear debris may migrate along the prosthesis-cement interface or into the fibrous layer commonly found between the cement and the bone. Osteolysis is frequently observed in the medial femoral neck where the particles naturally gravitate from the head-cup articulation but may also appear as linear patterns around the femur or as localized expansile lesions (Zicat et al. 1995). The increased use of collarless, polished tapered stems, which are designed to subside within the cement and seal the proximal metaphysis, has reduced the incidence of osteolysis seen in the femur (Yates et al. 2008).

2.1.4 Summary

There are a number of different factors that contribute towards the failure of THA but most revision surgery is conducted for aseptic failure of the prosthesis. The early development of aseptic failure is usually asymptomatic and the changes will not prompt the patient to seek medical advice until extensive. Other methods of failure of hip arthroplasty, such as infection, dislocation or fracture of the device, will be associated with pain and medical intervention will be prompted by the patient.

The silent and slow development of the mechanisms of aseptic failure can be identified by periodic review of the patients (Maloney et al. 1999, Dumbleton et al. 2002, British Orthopaedic

Association 2006, Wroblewski et al. 2007) and it is this principle which forms the basis of existing THA surveillance. The symptoms do not commonly appear in the early postoperative period but gradually increase from mid-term onwards. In order to establish the current situation with regard to the follow up of hip arthroplasty, a literature search of existing assessment methods was undertaken.

2.2 Literature review

2.2.1 Aim

The literature review was conducted to critically examine the methodology of existing orthopaedic follow-up of THA. The aim was to examine the range of methods used to identify failing hip arthroplasty and to discover any deficiencies that might be addressed by further research. The infected (or septic) THA was not included in the search as patients with an infected joint experience pain and will naturally seek medical help. The main purpose of routine surveillance is to identify the aseptically failing THA that may need further treatment as the patient is usually unaware of underlying changes or of the implications of any symptoms that they experience. The question posed to inform the literature review was: *How is an aseptically failing hip replacement identified and how is it assessed?* The search terms were selected to retrieve data relating only to aseptic failure and to capture the range of assessment methods used in THA surveillance from mid-term onwards.

2.2.2 Materials and methods

2.3.2.1 Data sources

Five electronic databases were used to conduct the search: Medline, EMBASE, CINAHL, AMED and BNI. The first two were via the Ovid gateway and the latter three via EBSCOhost. All studies or reviews which investigated or reported clinical and/or radiological outcomes after total hip replacement, published from 2000 to 2010 and available in English language were included. The limitation on language was pragmatic as all countries with established hip joint registries will publish some or all of their results in English and these are the countries in which follow up is more widely practised. It was noted that the search retrieved papers from a wide range of countries including France, Greece, and Japan, so the search still reflected international practice. The records included all types of total hip replacement.

There were two sets of search terms used with Boolean operators for each database – the first set were 'replacement' AND 'fail*' AND 'aseptic' (anywhere in the text) AND 'hip' in the title. The second set was similar but the word 'arthroplasty' was substituted for 'replacement'.

The use of both terms - 'arthroplasty' and 'replacement' – was intended to capture a wider selection of literature as there was a trend for national preferences with North American based authors favouring the use of 'arthroplasty' and British based authors the use of 'replacement'. Duplicates were removed and the search was limited to articles with full text available as all major orthopaedic journals had adopted full text availability during the time period being considered. This was a pragmatic approach due to the large number of articles available.

2.2.2.2 Study selection

The abstracts of all records were screened for eligibility against the aim of the study. Some were updates on previously reported cohorts of patients and in these cases, only the most recent article was included (Callaghan et al. 2004, Callaghan et al. 2009, Della Valle et al. 2009) although the older articles were retrieved for a description of method if needed.

2.2.2.3 Eligibility criteria

The aim of the literature search was to identify the range of assessment methods used in THA from mid-term onwards and studies with follow-up of less than four years after surgery were not included.   Animal studies, laboratory studies, detailed chemical analysis of materials and descriptions of surgical technique were also excluded as they were not relevant to the routine follow up of human subjects with hip arthroplasty.   Some studies, such as case reports or outcomes from small subgroups of patients, were excluded as the evidence may not have been applicable to the general THA population.

2.2.2.4 Quality assessment

The level of evidence for each study was assessed using the most recent tool published by the Centre for Evidence Based Medicine at the University of Oxford (Centre for Evidence Based Medicine 2010). This consists of levels one (highest) to five (lowest) depending on the type of research question and the strength of evidence obtained to support the stated results. This tool has an advisory note on the importance of individual judgement of quality of evidence in addition to the grading process. Consequently, the Critical Appraisal Skills Programme (CASP) tool for

cohort studies was applied to studies in which the level of evidence was unclear in order to evaluate the quality (Critical Appraisal Skills Programme 2010).

2.2.2.5 Data extraction

The data extracted were stored in a spreadsheet which included the study design, the number and type of hip replacements, length of follow up, proportion revised, methods of clinical review, outcome scores, methods of radiographic review, the definition of failure, the reasons for failure and key findings of the study (see Appendix VII). During this process, four further studies were identified from the reference lists and included. One of these was published in 1999 but was retained as it was considered to add important evidence to the literature review.

2.2.3 Results

2.2.3.1 Review process

A total of 863 citations were identified through the five databases. The process of reduction to the final 70 full text articles can be seen in Figure 2.1. During the process of data extraction, it was apparent that there were a number of methods of radiographic assessment which had been described in earlier studies, in total, an additional 35 studies. These were subsequently retrieved and details of the methods of radiographic assessment were extracted and stored in a second spreadsheet (see Appendix VIII). This provided supplementary information with which to interpret the 70 studies included in the literature review and allowed a comparison of radiographic assessment methods.

Figure 2.1. Flow diagram of selection process of the articles for literature review

```
┌─────────────────────────┐
│ Records identified by the │
│     search strategy      │
│       (n = 863)          │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Records after duplicates │
│        removed           │
│       (n = 574)          │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Records available in English │
│        language          │
│       (n = 498)          │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐         ┌────────────────────────────────┐
│ Records screened with full │──────▶│ Records excluded: not eligible or earlier │
│    text available        │         │ reports of studies with later publications │
│       (n = 224)          │         │            (n = 158)           │
└─────────────────────────┘         └────────────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Articles included     │
│       (n = 66)           │
└─────────────────────────┘
            │                        ┌────────────────────────────────┐
            │◀───────────────────────│ Articles added from reference lists │
            │                        │            (n = 4)             │
            ▼                        └────────────────────────────────┘
┌─────────────────────────┐
│ Final articles reviewed  │
│       (n = 70)           │
└─────────────────────────┘
```
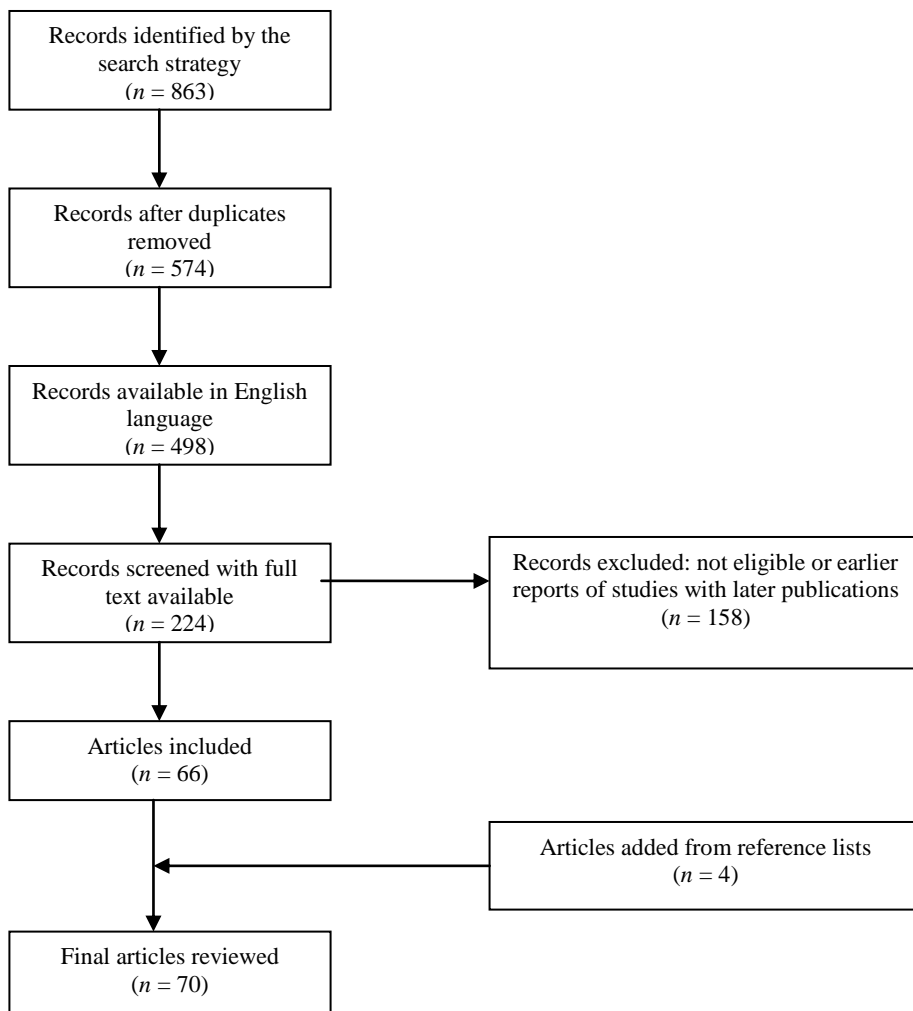
2.2.3.2 Quality assessment

The quality and type of papers that were reviewed is summarised in Table 2.2. The level of evidence was assigned with respect to the aim of the literature review which was to identify the methodology used in surveillance to identify the failing THA. The six studies based on registry data provided the strongest evidence while the one study described as a systematic review was of poor quality with no description of how the study was conducted or of the papers reviewed, and was consequently downgraded. The majority of papers were follow-up studies graded at level

three although some were given a lower grade when analysed using the CASP tool. In particular, the three studies graded level three minus showed evidence of bias and one at level four lacked sufficient follow up. All studies were retained in the review.

Table 2.2. Type of study and level of evidence of 70 reviewed papers

Type of study	No. of papers	Level of evidence
Registry study	6	2
Non-randomized controlled cohort	5	3
Follow-up study	50	3
	3	3 minus
	1	4
Historically controlled study	1	4
Cross sectional study	1	4
Systematic review	1	4
Opinion	2	5

2.2.3.3 Studies included

The median sample size of the 70 papers reviewed was 103 participants (mode 49, range 12 to 17,409). The age of the participants ranged from 12 to 96 years with a median age of 54 years. The median length of follow-up was 10 years (range 5 to 35) and the phrase 'mid-term' or 'intermediate' was used in six studies to refer to the 6 to 9 year postoperative period (Spangehl et al. 2001, Kim et al. 2003, Revell et al. 2006, Sugano et al. 2007, Ollivere et al. 2009, Emms et al. 2010). There were a broad range of countries (17) represented as shown below in Table 2.3.

Table 2.3. Countries represented in the 70 papers reviewed

Country	No. of papers (%)
Austria	2 (3)
Canada	2 (3)
Finland	4 (6)
France	1 (1.5)
Germany	2 (3)
Greece	3 (4)
Italy	1 (1.5)
Japan	4 (6)
Northern Ireland	1 (1.5)
Norway	2 (3)
Scotland	1 (1.5)
South Korea	4 (6)
Sweden	1 (1.5)
Switzerland	2 (3)
The Netherlands	1 (1.5)
England	17 (24)
USA	22 (30)

The data extracted included the methods used in assessment of a failing hip arthroplasty which were broadly grouped under the headings of clinical assessment, outcome scores and radiographic assessment.

2.2.3.4 Clinical assessment

Clinical assessment referred to any method that required a face-to-face situation with the patient and an evaluation by a health professional (surgeon completed outcome scores are addressed in the following section §2.2.3.5). A physical examination was used in two papers but no details or results were given. Leg length was measured in three studies in order to discuss restoration of anatomy and the use of walking aids was assessed in one study. There was one study in which hip abductor muscle strength was measured and two studies where haematology tests were used. There was only one study in which the Charnley classification of lower limb problems was applied (Charnley 1972).   This simple system, completed by the health professional, is used to categorise the extent of lower limb disability based on unilateral hip, bilateral hip or multiple joint disease and whether or not arthroplasty has taken place.

2.2.3.5 Outcome scores

There was a range of 14 outcome scores employed in a total of 53 papers; there were 17 (25%) papers without an outcome measure. Multiple use of scores in some papers resulted in a total of 77 uses and the frequency for each score can be seen in Table 2.4.

Eight of the scores were validated for use in THA and there were five surgeon and nine patient completed questionnaires. The only validated hip specific and patient reported outcome measure identified in the literature review was the Oxford Hip Score (OHS) although the Harris Hip Score (HHS) and Merle d'Aubigné and Postel scores were more frequently used. The geographic distribution of the four most frequently used scores is shown in Table 2.5.

In addition to the 14 specific outcome scores, there was a variety of methods used to record pain in the region of the hip in eight of the studies. Five of these employed a categorical score and three used a visual analogue score (VAS) although the question asked of the patient was not standardised across the studies. The results were statistically analysed in the three papers which had a numerical result recorded on a VAS and descriptive statistics were used to summarise the data in the others.

Table 2.4. Record of the outcome scores used for THA review (total of 77 uses)

Questionnaire	No. of papers (%)	Type of completion	Validated for use in THA
EuroQol[1]	3 (4)	Patient	Yes
Further surgery question	1 (1)	Patient	No
Harris Hip Score[2]	27 (35)	Surgeon	Yes
HSS hip score[3]	1 (1)	Surgeon	No
Iowa score[4]	1 (1)	Surgeon	No
Johnston questionnaire[5]	1 (1)	Surgeon	Yes
Merle d'Aubigné & Postel[6]	14 (18)	Surgeon	No
Oxford Hip Score[7]	10 (13)	Patient	Yes
Satisfaction score	5 (7)	Patient	No
SF 12[8]	1 (1)	Patient	Yes
SF 36[9]	1 (1)	Patient	Yes
Tegner activity score[10] (knee activity rating scale)	2 (3)	Patient	No
UCLA[11]	2 (3)	Patient	Yes
WOMAC[12]	8 (11)	Patient	Yes
Total	77 (100)		

[1] EuroQol questionnaire (EuroQol Group 2010) [2] Harris Hip Score (Harris 1969) [3] Hospital for Special Surgery hip score (Wilson et al. 1972) [4] Iowa score (Johnston et al. 1969) [5] Johnston questionnaire (Johnston et al. 1990) [6] Merle d'Aubigne and Postel (Merle D'Aubigne et al. 1954) [7] Oxford Hip Score (Dawson et al. 1996a) [8] Medical Outcomes Study short form 12 (Ware et al. 1996) [9] Medical Outcomes Study short form 36 (Ware et al. 1992) [10] Tegner activity score (Tegner et al. 1985) [11] University of California in Los Angeles activity scale (Amstutz et al. 1984) [12] Western Ontario and McMaster University Osteoarthritis index (Bellamy et al. 1988)

Table 2.5. Geographic distribution of most frequently used outcome scores

Score	Total	UK	N. America	Europe (excluding UK)	Far East
Harris Hip Score	27	5 (4 with OHS)	7 (1 with OHS)	9	6
Merle d'Aubigné and Postel	14	3 (2 with OHS)	3	6	2
Oxford Hip Score	10	9	1	0	0
WOMAC	8	0	5	3	0

2.2.3.6 General radiographic assessment of THA

There were 59 out of the 70 papers that included some form of radiographic assessment and it was from these that the 35 papers were identified with detail of radiographic review methods (§2.2.3.1). Of the 11 which did not include radiographic assessment, four were registry studies; one was a systematic review; two were follow up studies which used only questionnaires; one follow up study was retrospective and one was rated of poorer quality; and two papers were those written as an expert opinion.

The reporting of radiolucencies around components was usually defined by the zone in which they were found and was mentioned in 49 out of 59 papers (83%). The zones were described according to Gruen et al (1979) and Johnson et al (1990) for the femur and DeLee and Charnley (1976) for the acetabulum. Any width measurement was in millimetres (mm) whereas length was described as a proportion of the zone in which it was found.

Loosening of components was specifically mentioned in 17 (28%) of the papers although there were at least four different methods of assessing this in THA (Harris et al. 1982, Hodgkinson et al. 1988, Dall et al. 1992, Martell et al. 1993). The Dall method was used in two studies, the Harris method in five, the Hodgkinson method in four and the Martell method in one paper. All these methods are predominantly based on assessing the extent and width of radiolucencies and a change in position of the components (see Appendix VIII).

Osteolysis was recorded in 42 studies (71%) with reference to three definitions (Zicat et al. 1995, Joshi et al. 1998, Archibeck et al. 2001). The location was recorded by zone but there was

wide variety in technique used to record the size of lesions with no standardization of the methodology. Some were purely descriptive with lesions classified as linear or expansile; others classified lesions as 'small' or 'large' but the definition of these sizes varied between authors (Archibeck et al. 2001, Ito et al. 2004). An estimation of the area of lesions in mm^2 was attempted in some studies using a measurement of maximum length and width and was extrapolated to estimate volume with a mathematic formula in others.

Heterotopic ossification, the formation of bone in soft tissues, was assessed using the Brooker (1973) classification in 12 (20%) of the studies. Five (9%) papers included a descriptive assessment of trabecular density or absorption of bone which is indicative of the response to stress or absence of stress (Della Valle and Paprosky 2002).

2.2.3.7 Radiographic assessment of the acetabular cup

Issues that were specific to the acetabular component, the cup, included the measurement of wear which was mentioned in 23 (39%) of the studies. There were six different techniques listed and five studies in which the authors used their own method (see Table 2.6).

Table 2.6. Methods of measurement of wear of acetabular component

Method of measurement of wear	No. of papers
Charnley and Halley 1975	3
Dorr and Wan 1996	1
Griffiths et al 1978	2
Latimer and Laschiewicz 1996	1
Livermore et al 1990	6
Shaver et al 1997	5
Author's own method	5

The migration of the cup, vertically and medially, was measured in 22 (37%) out of 70 studies with one article referring to a previous method (Massin et al. 1989). The inclination of the cup opening in relation to a horizontal line through the pelvis was measured in 13 (22%) of the studies.

2.2.3.8 Radiographic assessment of the femoral stem

The most commonly occurring measurements recorded for the femoral component were the subsidence in mm (16 studies, 28%) and the alignment of the stem in relation to the femoral axis (14 studies, 24%). The measurement of the change in the centre of rotation and/or the femoral offset (horizontal distance from the long axis of the femur to the centre of the femoral head) were included in a few studies, four (7%) in total.

The response of the bone by cortical hypertrophy (thickening of cortical bone) was recorded in four (7%) studies and proximal femoral bone resorption was recorded in ten (17%) studies. Specific mention of pedestal formation, another bony response seen at the tip of uncemented stems, was recorded in two (4%) of the studies. However, this particular sign would also have been included in the ten (17%) studies that referred to assessment of cementless stems using the criteria of Engh et al (1990). This method assesses osseointegration of the stem through a combination of bony trabecular pattern, lack of reactive and radiolucent lines, absence of subsidence and evidence of distal stability.

Eleven (19%) of the studies graded the quality of the cement mantle around the femoral stem using either the Barrack method (7 studies), the Schmalzried method (2 studies) or the Mulroy method (2 studies) (Barrack et al. 1992, Schmalzried et al. 1993, Mulroy et al. 1995). There was also one paper which referred to a method of classifying the debonding seen adjacent to the shoulder of a cemented Charnley prosthesis (Berry et al. 1998).

2.2.3.9 Failure

The definition of failure of the THA varied between the 70 studies. Revision of a component for any reason other than infection was classed as failure in 34 studies (49%). However, some of these studies also included a diagnosis of aseptic loosening as failure even if the component had not been revised. The diagnosis of aseptic loosening of the acetabular cup or the femoral stem differed between studies but could be broadly divided into similar categories for the two components (see Table 2.7). There was one study in which stem fracture or bending was specifically listed as failure of the THA (Ong et al. 2002) and one in which cup fracture was listed (Muller et al. 2003).

Table 2.7. Criteria used to diagnose aseptic loosening of THA components

Acetabular loosening	Stem loosening
36 (54%) studies	18 (26%) studies
Radiolucencies in all three antero-posterior zones >2mm or >1mm	Radiolucencies >2mm in > two zones or 50-100% of circumference
Migration >2mm or >3mm or >4mm or >5mm	Subsidence >2mm or >3mm or >4mm
Change in angle of inclination by 5° or 8°	Change in varus or valgus by 2° or 3°
Expanding osteolysis	Expanding osteolysis
Excessive wear of cup	Cement mantle crack
	Lack of signs of osseointegration plus signs of distal instability (cementless)

2.2.3.10 Personnel

The studies were authored by orthopaedic surgeons in all cases although there were a number of other professions involved in the research. There were three papers which included scientists, nine with statisticians, five with allied health professionals or nurses, one with a radiologist and one with an epidemiologist in addition to ten which included basic researchers.

2.2.4 Discussion

The purpose of the literature review was to establish the range of methods commonly used in hip arthroplasty follow up and how they are used to identify aseptically failing hip replacements. The search produced a total of 70 papers for final review and data extraction, the majority being follow-up studies.

The search strategy limited records to English language and full text availability but the range of countries (17) was representative of the majority of orthopaedic associations across the world. The grey literature was not searched specifically although the British Orthopaedic guidelines were consulted (British Orthopaedic Association 2006) and there was some access to unpublished papers in which poor methodology or lack of generalisability were cited as the reasons for rejection. An additional 35 papers were retrieved from references to radiographic methodology in the 70 reviewed papers; they were used for supplementary information in the analysis of radiographic assessment.

2.2.4.1 Quality of evidence

The quality of the evidence was highest from the registry studies which were based on large amounts of data with the analysis carried out by a statistician or epidemiologist or a surgeon with academic links. Many of the other studies achieved a moderate level of evidence due to the methodological consistency in follow-up rather than rigorous scientific content. There were six studies of lower quality which were nonetheless retained in the review as they contributed to the understanding of a failing THA and methods of assessment.

The aim of the literature review was to identify the current methods used in assessment of THA and consequently, the level of evidence of a paper did not influence the extraction of this information. However, some of the studies provided recommendations for future surveillance techniques and, in these cases, the quality of the paper influenced the discussion that follows in this section. The lack of significant numbers of high quality papers to establish an orthopaedic evidence base is not unique to arthroplasty (Smith et al. 2009) and may improve with increasing use of validated tools and a multidisciplinary approach to research. Although all the studies were conducted by members of the orthopaedic medical profession, it was evident that there was a

range of personnel included in the authorship. This trend is likely to continue with the inclusion of non-medically qualified clinicians as part of the multidisciplinary orthopaedic team and with more stringent requirements for statistical analysis in orthopaedic publications (Petrie 2006, British Orthopaedic Association 2010).

2.2.4.2 Follow-up interval

The length of post surgery follow-up recorded in the included studies ranged from five to 35 years with an average of ten years (median). Earlier post-operative studies were excluded as irrelevant and the results indicate the time at which follow up is frequently used in evaluating the performance of the prostheses. The 'mid-term' description was used up to ten years post-operatively and beyond that, 'long term' was the common description of follow up. Current UK recommendation is for follow up to take place at one, five and ten years and thereafter as indicated (British Orthopaedic Association 2006). From the patient's perspective, an early follow-up provides reassurance (Bolz et al. 2010) even when they are unaware of possible changes. Some authors suggest that the mid-term review is important but does not give a guaranteed indication of the long-term survival of the prosthesis (Ollivere et al. 2009, Emms et al. 2010).

2.2.4.3 Clinical assessment

The components of clinical review that were recorded were minimal with tests for leg length and/or abductor muscle strength only mentioned in four papers. It is likely that any clinical interface would have included some type of measurement, such as a range of motion. However, a lack of reliability or high ceiling values often precludes the inclusion of such information unless a carefully standardised procedure is employed (Bryant et al. 1993, Soderman and Malchau 2001).

2.2.4.4 Outcome scores

Although pain was recorded as an outcome in eight studies, there was no standardisation for the method of obtaining this information and only three used the results for statistical inference. A majority of the studies included some type of comprehensive outcome score (81%) although less than half (32 out of 77 uses) were patient completed scores. Traditional orthopaedic assessment

was completed by the surgeon and the same measure was applied throughout long term follow-up although some studies included a newer outcome tool at subsequent assessment points (Norton et al. 2002, Williams et al. 2002, Pospischill et al. 2005, Incavo et al. 2008, Utting et al. 2008, Wangen et al. 2008, Santori et al. 2010). Since the 1990's, PROMs have been introduced and used in orthopaedic assessment (Ware et al. 1992, Dawson et al. 1996a, Ware et al. 1996, Bellamy et al. 1988, EuroQol Group 2010). They provide a measurement of the health status from the patient's perspective and do not require a clinical interface. The simplicity of application and independence from the surgeon are considered to be an advantage to the patient (Learmonth and Cavendish 2005). In addition, they can be administered by post and are currently being used to evaluate health service delivery in the UK (Department of Health 2008). The use of PROMs is gradually increasing (Harvie et al. 2005, Malchau et al. 2005) but the adoption of this process is time dependent and is illustrated by the proportion of surgeon completed scores in the literature review.

The most frequently used scores were the HHS, Merle d'Aubigné and Postel, the OHS and the WOMAC. There are differences between these scores other than the surgeon or patient completion issue. The Merle d'Aubigné and Postel score assesses pain, range of movement and walking ability, each on a scale of zero to six, and the HHS adds function and absence of deformity. The HHS and Merle d'Aubigné were both designed for a surgeon to assess the outcome of a THA. The OHS was designed for the patient to record their progress from before to after a THA whereas WOMAC is a more general tool designed for use in a disease specific group (osteoarthritis) affecting the hip or knee. The geographic distribution of the score usage provides an interesting insight. The HHS was developed in the USA, the Merle D'Aubigné in France, the OHS in England and WOMAC in Canada. The results showed a trend to select a tool developed in the same geographic region (see Table 2.5). There was a slight tendency for more frequent use of the HHS in Europe than in North America but this may be due to the large numbers of papers published over the years which refer to the HHS (making it more familiar to researchers) and pressure from some publishers to include it in long term follow-up (The Journal of Bone and Joint Surgery, American Volume 2010).

Aside from the historical and geographical influences on the choice of outcome score, it is now recommended that the psychometrics of a tool be considered (Pynsent et al. 2004). This

refers to the validity, reliability and responsiveness to change. The validity is the ability of the tool to measure the construct for which it has been developed and is strengthened by repeated testing in a specific population (Peat et al. 2002, Pynsent et al. 2004). The reliability is the ability to produce consistent data when used repeatedly on the same population and the responsiveness is the ability of the tool to measure meaningful change (Sim and Wright 2000). The HHS and Merle d'Aubigné score were both developed before psychometric testing was identified as important although there has been some retrospective testing of the HHS (Soderman and Malchau 2001). However, the statistical tests reported by Soderman and Malchau (2001) were correlations which are not recommended in psychometric tests (Altman and Bland 1983, Morris 1997, Pynsent et al. 2004). In addition, the authors do not recommend use of the range of motion and the deformity components of the HHS due to unreliability (Soderman and Malchau 2001).

In contrast, the OHS and WOMAC have both undergone rigorous testing in the development process, and subsequently, and have been shown to be valid and reliable in the THA population (Dawson et al. 1996a, Dawson et al. 1996b, Bellamy et al. 1988, Fitzpatrick et al. 2000, Nilsdotter et al. 2001). The responsiveness of both scores in the THA population, as measured by standardised effect size, was good although the OHS was higher than WOMAC in the global score and the pain subscale (Dawson et al. 1996b, Ostendorf et al. 2004, Garbuz et al. 2006). The subtle differences between these two instruments reflects the more specific nature of the OHS when evaluating THA but further evaluation is needed to define the uses to which each is best suited. If a global score is to be used, the high levels of responsiveness of the OHS make it the instrument of choice providing that validity and reliability are satisfied for the study population. In addition, the OHS may be sufficiently responsive to allow comparison of prostheses within a patient population (Murray et al. 2007).

'Noise' is the confounding effect of other physical or psychological changes on an outcome score. It is thought that a joint and condition specific tool, such as the OHS, should reduce the problem of noise and, although noise has been found in both WOMAC and the OHS, the effect was smaller in the OHS (Dawson et al. 1996b, Daniel et al. 2004, Ostendorf et al. 2004, Wylde et al. 2005). This effect can be reduced by the inclusion of a second tool which enables the researcher to capture in advance information relevant to the condition being assessed e.g. a comorbidity index or a generic instrument such as the SF-12 or EuroQol (Ethgen et al. 2004,

Garbuz et al. 2006, Pollard et al. 2006, Murray et al. 2007). However, the balance between information retrieval and burden on the participant must always be considered.

The analysis of the scores reported in the papers was varied. Some reported the mean of the group scores pre-operatively and post-operatively and emphasised the improvement but did not relate the magnitude of this change to other cohorts for comparison (Kim et al. 2002, Eskelinen et al. 2006, Kim 2008, Petsatodis et al. 2010). Others adopted a similar approach but classified the final scores into groups such as excellent, good, fair and poor (Archibeck et al. 2001, Santori et al. 2010). A few papers interpreted the results with respect to normative values (Pospischill et al. 2005, Wangen et al. 2008, Ollivere et al. 2009) and a few related the results to clinical outcomes in a meaningful way (Kim 2005, Eskelinen et al. 2006, Rajaratnam et al. 2008). Some papers reported average outcome scores to compare study groups but without apparent significant differences (Keener et al. 2003, Kim 2005, Nixon et al. 2007, Incavo et al. 2008) and still others reported only a final follow-up score without any measure of change (Grübl et al. 2006, Yates et al. 2008).

Although a wide variety of outcome scores has been found in the literature, the choice of an appropriate tool for ongoing hip arthroplasty review should be with attention to the aims of surveillance. The outcome score must reflect the patient perspective, be simple to complete and must produce reliable data, all of which the OHS has been shown to achieve (Fitzpatrick et al. 2000, Ethgen et al.2004, Learmonth et al. 2005, Garbuz et al. 2006, Murray et al. 2007). It must be responsive to the population of THA and the confounding effect of comorbidities should be minimal. For surveillance purposes, it is also desirable that the outcome score is sufficiently sensitive to detect difference between sub groups within the population. The OHS satisfies all of these requirements and is intended to be used for hip arthroplasty review (Dawson et al. 1996b).

2.2.4.5 Radiographic assessment

The radiographic assessment of total hip arthroplasty utilised many different methods, some of which were from previous studies and some of which were the authors' own techniques. The extent and width of radiolucencies around a component were included in most assessments and in all of the methods of 'diagnosing' loosening quoted in the literature (Harris et al. 1982, Engh et al. 1987, Hodgkinson et al. 1988, Engh et al. 1990, Dall et al. 1992, Martell et al. 1993). The

methods of Dall (1992), Harris (1982) and Hodgkinson (1986) were all for cemented THA. The methods of Engh (1987 and 1990) were for cementless stems and that of Martell (1993) was for cementless THA. Although there were differences between these methods, assessment of radiolucencies was common to all as was any change in position of the component. Additional features for cementless stems were signs of osseointegration. More recently, methods have been developed with similar principles for hip resurfacing procedures (Amstutz et al. 1984, Pollard et al. 2006).

The wide range of radiographic methods quoted in the literature presents a challenge and the clinical relevance of some of the data gathered may be debated (Malchau et al. 2005, Pollard et al. 2006). For example, in one study, prostheses with criteria predictive of loosening at an early stage were not adversely affected when evaluated at 12 years (Ollivere et al. 2009). This shows how the evidence predictive of poor survival is continually developing. It has been suggested that radiographic assessment should include only the data indicative of changes which threaten the stability of the implant and that methods should be simple and easily reproducible (Malchau et al. 2005). The support for this principle requires further research.

The presence of osteolysis was recorded in 71% of the studies but the methods lacked consistency. Many of the measurements were made by estimation of area from the maximum length and width which does not take into account the irregularity of the shape of such lesions. In one study, any radiolucency greater than 1mm width was considered to be osteolysis (Hartofilakidis et al. 2008) and in another, width greater than 2mm was osteolysis (Ito et al. 2004). In a third, it was defined as any non-linear radiolucency at the bone-cement interface wider than 5mm (Altenburg et al. 2009). There were also differences between small and large lesion sizes (Archibeck et al. 2001, Ito et al. 2004). It appeared that there was considerable inconsistency in recording the size of osteolytic lesions.

Wear was measured by six different documented methods and another five methods which were author designed. Of the six recognised methods, four were for cemented cups, one for cementless and one could be used for both (see Appendix VIII). One technique was dependent on a computer software package whereas the others could be used with integral software on digital images or callipers on plain films. Some required templates to locate the centre of the femoral head. The method of Dorr and Wan (1996) was the only method which did not require

serial x-rays and which could be used for both types of cup. The disadvantage of this method was that it did not directly measure the point at which maximum penetration of the femoral head would normally occur.

The grading of the femoral cement mantle according to thickness, thought to be predictive for long term survival, was predominantly assessed by the method of Barrack et al (1992). However, there was a later modification (Mulroy et al. 1995) which was used only twice and another modification for use in hybrid THA (Schmalzried and Harris 1993) which was used twice. The same senior academic clinician was an author on all three papers (W. H. Harris). The frequency with which these methods were quoted indicates a dilemma: the oldest method is most familiar and tends to be used frequently whereas subsequent modifications, even though with the same prominent author, are often under utilized. For any reader unfamiliar with the methods, the choice of three can seem confusing. In addition, the value of assessing the quality of the cement grading has subsequently been questioned (Williams et al. 2002). The value of assessing cement grading at mid to long term follow up is debatable since the development of radiolucencies over the time since surgery will threaten the validity of such an assessment. There is a further challenge to the value of cement grading which is known as the 'French paradox'. This technique of THA involves inserting a femoral component which fills the medullary cavity and is consequently surrounded by a very thin (absent in some places) cement mantle. However, long term results indicate excellent survival of the prostheses and further challenge the widely held views on the quality of the cement mantle (Langlais et al. 2003, El Masri et al. 2010).

Approximately one third of the studies referred to some bony response – hypertrophy, resorption, pedestal formation or altered density of bony trabeculae (stress shielding). These signs are indicative of altered stresses on the bone and are interpreted with reference to the type of component in place. For example, the formation of a pedestal at the tip of an uncemented stem shows that the bone has responded to provide stability and is positive as long as it is in contact with the prosthesis and there are no new reactive lines around the stem (Engh et al. 1990). The stress shielding seen proximally in the femur may not be a threat to component stability if distal fixation is secure (Della Valle et al. 2002). Long term follow-up will continue to provide evidence of the changes which are most significant for evaluating the state of the prosthesis.

Of the 70 studies reviewed, 54% stated that acetabular loosening was a mechanism of failure whereas only 27% incorporated femoral loosening in their definition of failure. This suggests that failure of the acetabular component has continued to be of greater concern than femoral failure over the period of these publications despite the use of cementless cups which were introduced to reduce this problem (Kim et al. 2002). The risk of osteolysis continues to threaten the survival and many of the authors recommend continued follow-up to protect the patient from the unseen development of bone destruction (Iida et al. 2000, Norton et al. 2002, Malchau et al. 2005, McMinn et al. 2008, Ogino et al. 2008, Yates et al. 2008, Wroblewski et al. 2009). The process is known to be silent in the pelvis (Maloney et al. 1999, Utting et al. 2008) and absence of early osteolysis is no guarantee against later development (Ihle et al. 2008, Emms et al. 2010).

2.2.4.6 THA surveillance

There was overwhelming evidence that authors were convinced of the need for follow-up of THA. Many authors stated it explicitly in the conclusion (Maloney et al. 1999, Wroblewski et al. 2000, Norton et al. 2002, Nixon et al. 2007, Wroblewski et al. 2007, Ihle et al. 2008, Makela et al. 2008, Ollivere et al. 2009) and many more argued for it on the basis of surgeon and patient experience. Surgeons were convinced of the need for three reasons: because of the commonly asymptomatic nature of aseptic loosening (Haddad et al. 2007, Ghoz et al. 2008, Huddleston et al. 2010); because long term results are not known for newer prostheses (Walton et al. 2005, Ghoz et al. 2008, Emms et al. 2010); and because of the difficulties of revision surgery when extensive bone loss is present (Burns et al. 2006, Haddad et al. 2007, Howard 2009). For the patient, routine surveillance provides reassurance but more importantly, it allows early identification of impending failure. This allows the patient time to consider and plan for revision surgery before significant changes such as intolerable pain or fracture of the surrounding bone necessitate emergency surgery (Callaghan et al. 2004). Timely procedures also reduce the risk of comorbidity associated with periprosthetic fracture or complicated revision surgery due to extensive bone loss (Lavernia 1998, Maloney et al. 1999, Capello et al. 2003, Callaghan et al. 2004, Biring et al. 2007, Haddad et al. 2007, Wroblewski et al. 2007, Ihle et al. 2008).

2.3 Conclusion

The endpoint of an aseptically failing hip arthroplasty is revision surgery but important signs of impending failure are the appearance of extensive radiolucencies around or a change in position of a component, or expanding osteolysis. Other indicators are significant wear of the cup or cup liner or, in uncemented components, absence of osseointegration with a combination of reactive lines, calcar hypertrophy and particle shedding (Engh et al. 1990). All of these signs of failure are radiographic and require serial x-rays. The methods of quantifying the changes are varied which makes comparison between studies on the basis of change difficult although the rate of revision or intention to revise is a commonly quoted figure. There have been attempts in the past to standardise radiographic review (Johnston et al. 1990) but large amounts of data can prove cumbersome to obtain and are not always relevant (Malchau et al. 2005). Further research is needed to identify simple psychometrically sound methods of quantifying the indicators of failure across all types of THA. The information would be useful for THA surveillance and may encourage compliance with recommendations for routine follow-up. Similarly, it may be that further research can identify simple, reproducible radiographic tools to quantify specific changes, such as the expansion of osteolytic lesions.

The results from clinical assessment or outcome scores are used to supplement the radiographic information but are not used as separate indicators of failure. The use of clinical assessment methods is minimal but outcome scores are regularly used in the arthroplasty review process. Although the evidence to link scores with radiographic signs is currently limited (Nixon et al. 2007), the scores are used to capture some of the patient perspective. Further research to explore the relationship between validated outcome scores and radiographic changes could identify essential elements of the assessment process. The information obtained would be beneficial to surgeons and also to patients who may find it difficult to understand a recommendation for revision on the basis of x-ray changes when there is an absence of symptoms.

The choice of outcome tools used in THA surveillance shows some regional partiality as well as a historic tendency to continue a particular method. This is an understandable pattern but as the use of PROMs increases, it will be important to identify which is the best tool to capture

patient experience and measure change for defined categories of THA. Such research will contribute important information to the development of the surveillance process.

The highest quality studies reviewed were from registry data which provides information about the types of THA which have failed and the time at which this happened. The follow-up studies provided details of how and when assessment took place (Burns et al. 2006). However, at times of significant economic constraint, it is essential to identify the most important patient categories and optimal time intervals for review and this will require further research and definition.

The current health service environment requires that any service is both efficient and effective. Provision of THA surveillance relies on balancing the cost of the service and simpler revision procedures against the costs of complicated revision surgery for asymptomatic patients who have developed catastrophic failure. There is a need to keep surveillance costs as low as possible and yet maintain an effective service. The use of non-medically qualified health professionals within the orthopaedic team is one way in which this can be achieved but the effectiveness of this role substitution requires evaluation.

For the patient, the benefit of timely surveillance is undeniable. If revision is indicated, the patient can be involved in the decision making process, the procedure can be planned instead of an emergency and morbidity reduced.

In conclusion, the literature review provides insights into the methods commonly used in follow-up of THA and the definition of failing hip arthroplasty. Areas for further research are identified through an analysis of the components of the review process. Some of these will be addressed in subsequent chapters such as the use of non-medically qualified health practitioners in arthroplasty review, the development of a simple radiographic tool to quantify osteolytic lesions and the relationship between a validated outcome score and radiographic changes.

Chapter 3

Image Interpretation

This chapter describes the development of image interpretation skills by a non-medically trained health professional. The ability to assess x-ray images of joint replacements is an essential part of any arthroplasty surveillance service and is necessary for identifying a failing hip replacement. The review of current literature in the previous chapter demonstrated that there was no standard way to conduct this type of radiological assessment and consequently, training for this skill was also lacking in definition although clearly essential to the process. The academic requirements for gaining a doctoral degree provided an opportunity to describe and develop the acquisition of this skill through a formal master's level module. It demonstrated a method of achieving competency in an assessment skill which is then transferable and can be used clinically as well as in future research studies.

3.1 Background

The need for ongoing surveillance of hip arthroplasty was established in the previous chapter in which it was shown that a THA does not last indefinitely and that there are clearly recognised processes of aseptic degeneration which can lead to failure of the joint. The situation is not unique to hip arthroplasty and follow-up of total knee replacements is also advocated for the same reasons (British Orthopaedic Association et al. 1999). The provision of any type of arthroplasty surveillance service requires suitably trained health professionals with the ability to view and interpret x-ray images of the joint replacements.

The long term follow-up of joint replacements was traditionally carried out by the orthopaedic team responsible for the surgery but this is now impossible for many surgeons. The current environment in the National Health Service (NHS) in the United Kingdom is one of increasing economic pressure to deliver efficient, streamlined services. Alongside the increase in service demands, there is reduced availability of junior doctors as a result of the European directive (Department of Health 2004a). This has prompted the development of other health professionals to perform roles traditionally held by the medical profession (Department of Health 2000, Green et al. 2008, McPherson et al. 2006, Ruston 2008).

In response, the nursing and allied health professions have developed 'advanced practitioner' roles and those which are described as 'clinical specialist' or an 'extended scope practitioner' (ESP) (McPherson et al. 2006, Ruston 2008). A clinical specialist describes a health professional who has gained expertise in an area of medicine that enables them to work at an advanced clinical level and an ESP has added skills which are beyond the scope of their normal practice in order to work within their specific area of expertise (Ruston 2008). The physiotherapy profession, in particular, has encouraged individuals to extend their practice through suitable qualifications which support their advanced clinical roles (Gosling 1999, Green et al. 2008). This is seen as one way to support the provision of health services and increase the job satisfaction for the health professional (Ruston 2008). Physiotherapists elsewhere have specifically expressed a desire for further training (Li et al. 2009) and the need for investment in education for extended roles is recognised (Kersten et al. 2007).

From the patient's perspective, the delivery of care by non-medical health professionals has proved satisfactory and has contributed to a reduction in waiting times although more research is needed (Daker-White et al. 1999, Kersten et al. 2007, Walton et al. 2008). There are some questions about the extent of the cost savings but the service is as effective as that delivered by junior doctors (Walton et al. 2008). These findings suggest that there should be continued development of the roles and that suitable training should be provided to support the health professionals involved.

The long term follow-up of patients with joint replacements was traditionally the prerogative of medical staff but there has been support from within the orthopaedic community for other health professionals to provide this service (British Orthopaedic Association et al. 1999,

British Orthopaedic Association 2006). Described as an 'arthroplasty practitioner', the chosen health professional will be responsible for requesting and interpreting the relevant radiographic images for arthroplasty review as well as collecting outcome scores and performing a clinical examination (Walton et al. 2008). The additional training required to develop these skills is not a standardised procedure and has often been informally delivered by an enthusiastic medical colleague (Hardy et al. 2003, McPherson et al. 2006).

The first stage in developing a training programme for advanced health care professionals is to define the skills that are needed in order to perform the specific role. This has recently been completed for arthroplasty practitioners as a joint project under the direction of the Department of Health (DH) (see §1.3.2). Representatives of various professional groups were involved - orthopaedic surgeons, the nursing profession, the allied health professionals, a higher education institution (HEI) and ACPA. ACPA is an association of health professionals from varied disciplines who are working at an advanced level in close collaboration with orthopaedic consultants to provide care of arthroplasty patients throughout the patient pathway (ACPA 2011). It was formed to promote consensus of methodology and communication between all stakeholders involved in arthroplasty. The documents produced from the collaboration under the direction of the DH have provided a framework for assessing the level of skill of an arthroplasty practitioner and defining any training that might be required. They also form the basis for standardisation of such skills so that the role can become nationally transferable. It is thought to be the first such work in the UK to cross professional boundaries and it enables orthopaedic surgeons and ACPA to jointly control the future workforce development in arthroplasty.

The competencies described in the document cover all possible tasks of an arthroplasty practitioner - a non-medically qualified health professional involved in the care of patients undergoing elective joint replacement. The practitioner will be working at an advanced level, most commonly from a physiotherapy or nursing background, and the skill set that needs to be added will be dependent on the experience, training and qualifications held by the individual. The competencies provide a framework by which the required skills can be identified. One area in which training has commonly been required is the interpretation of x-ray images of joint replacements. In the past, there has been no defined route to achieve this goal, or method by

which an individual can demonstrate a competency in this skill. The new framework provides a method of demonstrating competency but as yet, there is no specific training method.

At a local level, the response to the lack of organised training for arthroplasty image interpretation has led to individuals utilising the resources available and adapting them appropriately. Similar situations have occurred in physiotherapy when there was a need for service delivery (Kersten et al. 2007). When the need for a specialist service becomes apparent, it is developed by senior medical colleagues in conjunction with a non-medical health professional who already has some knowledge of the specialism. The health professional will employ a number of methods to acquire additional skills using the expertise of colleagues, an HEI and self-directed learning.

In the particular case referred to in the following account, the health professional (the researcher) was a qualified chartered physiotherapist with 15 years experience in musculoskeletal conditions. She had been employed as an orthopaedic researcher in a district general hospital and had completed an outcomes study of patients with hip replacements. A need for ongoing follow-up of all local patients with joint replacements was identified at the completion of this study and she was asked to design and implement an arthroplasty review service. This required her to develop as an arthroplasty practitioner as well as to maintain her research role. The skill of image interpretation was required and, in view of the lack of formal training for this skill, a different model of learning was employed. 'Evidencing work based learning' (EWBL) is a method of combining intuitive and experiential learning with explicit learning in order to improve performance and demonstrate the learning achieved. It is assessed through an HEI in order to provide structure and to ratify the final product. The following account is an illustration of how this was done in order to acquire the skill of image interpretation in arthroplasty review.

3.2   Aims and objectives

The module of EWBL study was designed to show how the skill of image interpretation was acquired and how it had been an integral part of the development of a new service. The aims of the study were as follows:

- To describe a method of acquiring the skill of x-ray image interpretation in arthroplasty

- To demonstrate that a level of competency suitable for practice in arthroplasty review had been achieved

- To develop a local protocol for image interpretation in arthroplasty review

The aims were fulfilled through five specific learning objectives which were defined in the following way:

1. To develop the skill of image interpretation of x-rays of hip and knee replacements

2. To demonstrate competency in arthroplasty image interpretation

3. To demonstrate understanding of the contribution that x-ray images make to the management of complex situations of patients with joint replacements

4. To develop a local protocol for non-medical professionals requesting radiographic investigations in arthroplasty review

5. To critically reflect on the unique position in the local organisation: the responsibility for and management of learning the skill of image interpretation to support the provision of a new service

The development of a local protocol and reflection on service provision were included in recognition of the need for succession planning. There is an ongoing need for such services and the training of other professionals to undertake similar roles.

3.3   Methodology

The use of an EWBL module to capture the development of a new skill involves a different approach to learning than is traditionally adopted in medical training. The student is required to define the specific objectives and to identify the method of assessment for each of these objectives. The researcher chose to use a portfolio of evidence which was submitted for assessment in order to attain master's level academic requirements. This was accompanied by a written summary of the learning process, which is presented in the following account. The process of developing the skill is illustrated by a timeline constructed to map significant events (see Figure 3.2, page 68).

3.3.1   Objective One: To develop the skill of image interpretation of x-rays of hip and knee replacements

A health professional training in physiotherapy had provided knowledge of anatomy but no formal training in x-ray interpretation. The lack of an image interpretation course specific to arthroplasty presented a barrier to developing this skill but is not a unique problem (Kersten et al. 2007). It provided an opportunity to explore other methods of learning. The first response was to use critical thinking to analyse the situation. Critical thinking allows a person to analyse a situation and then determine a course of action. Brookfield (1987) states that a recognition of underlying assumptions and consideration of alternatives is central to critical thinking and will lead to active enquiry which is a further process of reflection and informed action.

A tool known as a SWOT analysis was used in the process of critical thinking, which is a documentation of all identified **S**trengths, **W**eaknesses, **O**pportunities and **T**hreats to the proposed action (see Table 3.1). This was instrumental in subsequently identifying different ways of learning using theoretical and reflective models.

Table 3.1 SWOT analysis of a skill in image interpretation at February 2007

Strengths • I have already looked at hundreds of films • My current practice provides a weekly assessment of x-rays • Recent local installation of a digital x-ray system has provided experience of digital films as well as plain films • There is a positive reinforcement to learn the skill as it directly enhances the work I am doing	**Weaknesses** • I have not looked at many x-rays over my years of physiotherapy practice • I have not completed any formal course on image interpretation • I have not yet been in the operating theatre to see the 3-dimensional reality of x-ray images of joint arthroplasty which has failed
Opportunities • The viewing of digital x-ray images allows frequent interaction with radiologists • Regular time with orthopaedic surgeons is provided through the review of arthroplasty clinic patients • Formalisation of this learning experience can provide a pathway of learning for future arthroplasty practitioners	**Threats** • Lack of time • Previous attempts to observe radiologists reporting on x-ray images have proved difficult to achieve • There is no standardised method of acquiring this skill • There is no recognised method of assessing this skill when acquired

3.3.1.1 Theoretical learning

In the absence of a didactic experience specific to arthroplasty, theoretical knowledge was gained through reading relevant literature and recording significant information, attending relevant courses or events and subsequently presenting to peers to demonstrate an understanding of the principles involved. This included, for example, a lower limb x-ray interpretation course, formal teaching from a consultant radiologist and a presentation to local radiographers (see Figure 3.2).

In addition, two assessment sheets were designed by the researcher to document all the relevant radiographic observations in the process of arthroplasty review, one for hip replacements and one for knee replacements. The design of the THA version was with reference to current orthopaedic literature and expert opinion and was subsequently used in a clinical study (see Appendix: V).

Peer discussion was another essential component in the early learning of this skill as it provided an opportunity to verbally test the understanding of the principles of image interpretation. This was achieved at local level with radiologist and radiographer colleagues who were also willing to explain concepts and illustrate principles using existing radiographic images. The transition from plain film to an electronic retrieval and storage system for radiographic images occurred during the early part of the skill development. It provided additional opportunities to increase the knowledge and understanding of x-ray review with support from consultant radiologists and radiographers, opportunities which were difficult to create prior to the adoption of the digitised system.

Interaction with other arthroplasty practitioners at national level through ACPA provided access to experienced colleagues in major orthopaedic centres where local systems of acquiring the skill of image interpretation had been developed. The use of ideas generated through this interaction was an invaluable source of learning and also stimulated visits to other hospitals where arthroplasty review was in place, thereby accelerating the process by which this skill was acquired.

3.3.1.2 Reflective learning

Reflective learning provides a different model from a rational approach. Reflection 'is a form of response of the learner to experience' (Boud et al. 1985). Experiences, provoked by internal or external agents, are thought about, reflected on, and meanings are derived. These then reinforce or change the course of future actions. This cyclical process relies on experience for initiation, and reflection to facilitate the subsequent learning. It was described by Kolb (1984) as a 'process whereby knowledge is created through the transformation of experience' (Kolb 1984). He developed the theory and identified four stages which are known as Kolb's Learning Cycle. An illustration of this is provided in Figure 3.1.

Figure 3.1 An illustration of the stages of Kolb's Learning Cycle (Kolb 1984)

Redacted due to copyright

The cyclical reflective process was fundamental to the learning gained through weekly sessions of x-ray review with senior orthopaedic colleagues, one of whom was a revision hip specialist and one was a revision knee specialist. These informal teaching sessions were initiated specifically to support the learning. In the early phase, learning was about the basic elements of assessment and a rich source of information for reflection was generated on each occasion. This was captured in reflective writing or notes summarising the principles derived in order to apply it at the next session. As understanding progressed, reflective learning was stimulated by more unusual findings which prompted further reflection and investigation in order to develop understanding of the concept and then, subsequently, to re-apply that knowledge to future reviews.

CHAPTER 3. IMAGE INTERPRETATION

A diverse range of experiences were used to stimulate learning using the model of reflective learning, as recommended in educational theory (Claxton et al. 1996). For instance, observation of surgery to revise hip or knee replacements provided an opportunity to relate the three-dimensional reality of degenerative changes to the two-dimensional observations made on the pre-operative x-ray images. That is, by viewing the pre-operative x-ray images and then attending the subsequent surgery, there was an increased understanding of the extent of the lesions seen on x-ray and the relevance for the surgeon. Reflective writing was used to capture this learning and to describe the potential impact on future practice.

Reflective writing is intended to help the writer identify what has taken place, what has been learnt and what is the future impact of that learning. An excerpt from a piece of reflective writing at the completion of the timeline illustrates some of the progress that was made:

> *Over the last three years of the pathway, in excess of 2000 x-rays have been viewed which has significantly increased my knowledge of normal and abnormal. I can confidently describe and interpret the images seen. However, as these images were mostly of the lower limb (hips and knees), I need to be aware of my limitations when assessing upper limb arthroplasty. I need to continue to take opportunities to view upper limb x-rays in order to develop the same level of expertise.*
>
> *The opportunity to discuss hundreds of cases with specialist orthopaedic surgeons has increased my understanding of the clinical presentation of problems as well as the images of arthroplasty. My discussions have been informed by attendance at conferences, reading of current literature and discussion with other members of the healthcare team both locally and nationally. While I have had some opportunities to use the acquired knowledge in case studies, I need to continue to use the skills in complex scenarios in order to consolidate my knowledge. This will be a further advancement in practice and the current assessment tools can be modified to reflect the level of skill.*
>
> *My ability to manipulate the software in the digital x-ray system has vastly increased through this period of learning and enables me to confidently obtain critical measurements of joint replacements. This transferable skill is useful for both*

arthroplasty clinics and research studies, and as such, directly contributes to the research training acquired through the PhD process.

3.3.1.3 Summary

This learning objective identified the methods by which the skill of image interpretation was acquired for use in arthroplasty review. The transferable nature of the skill meant that it would also be useful for the role of orthopaedic researcher. Future studies measuring the parameters indicative of the status of a joint replacement could confidently be undertaken. The use of a portfolio to capture the evidence reinforced the benefit of a variety of experiences to develop this skill. The lack of a formal method of learning had proved to be the catalyst for using a range of learning techniques. The benefit of this diversity was in its accessibility as it employed methods of learning which were available locally and nationally. It provides a template which has the potential to be applied elsewhere. A similar pattern could be employed by future arthroplasty practitioners needing to develop their competency in image interpretation.

3.3.2 Objective Two: To demonstrate competency in arthroplasty image interpretation

The development of new skills for health professionals who are working outside their traditional roles requires evidence of competency in order to ensure good quality of service for the patients, to provide a basis for audit and to protect the practitioner (McPherson et al. 2006). The newly developed framework for arthroplasty practitioners had not been completed at the time that the researcher was developing this skill. Consequently, in order to demonstrate competency in image interpretation, a variety of methods were used: the development of a tool for demonstrating competency, the application of that to the local arthroplasty service and the presentation of the acquired knowledge to professionals from a different discipline in a comprehensive but relevant way.

The challenge of demonstrating competency in a practical skill is well recognised and has been illustrated in the manual produced for musculoskeletal physiotherapists in the United Kingdom (The Chartered Society of Physiotherapy 2009). One of the recognised techniques used in medical training is the completion of a DOPS document (Direct Observation of Procedural Skills). This involves a senior member of a specialist medical team observing a junior member perform a practical skill. Peer discussion through ACPA led to the development of a simple proforma based on this medical model for use in arthroplasty review (see Appendix II). The document was subsequently used in the weekly x-ray review sessions with senior orthopaedic consultants to appraise the level of skill demonstrated by the researcher and to indicate where further study was required. It also provided a template for annual audit.

The presentation of acquired knowledge enables a learner to consolidate their understanding through articulation of the principles and to be challenged by alternative views (Weissman 2003). This principle was applied through delivery of an illustrated talk which was given to local radiographers, prompted by frequent interaction with them including the requests that had been made by the researcher for specific x-ray views. The senior radiographers regarded this as an opportunity to enhance their knowledge of what was needed for orthopaedic audit and research. It was also an opportunity to help the junior radiographers to improve their technique through understanding the intended outcome. The questions to the researcher from the radiographers

provided a stimulus for further reflection on the skill of image interpretation in arthroplasty review and orthopaedic research.

3.3.3 Objective Three: To demonstrate understanding of the contribution that x-ray images make to the management of complex situations of patients with joint replacements

Development of clinical reasoning to manage complex situations is an advanced skill which involves practical experience as well as underpinning theoretical knowledge (Atkinson et al. 2000, Jensen et al. 2000). The theory of integrated learning for adults suggests that a repertoire of learning activities is needed (Claxton et al. 1996). The methods of achieving this objective included the observation of surgical procedures, reflection on those experiences and some individual case studies. In addition, attendance at a specialist arthroplasty course significantly enhanced an understanding of the subject which was reinforced through reflection. The Edinburgh Summer University 2008 (see Figure 3.2) was a three day course for specialist orthopaedic registrars on the subject of hip replacement. Complex case studies were part of the interactive teaching on this course and they were illustrated with serial x-ray images which were presented and discussed.

3.3.3.1 Revision surgery

Revision surgery of the hip or knee is only undertaken after an orthopaedic consultant has completed a number of investigative tests and is convinced that the treatment is required. It is a more complex procedure than the primary surgery. By observing such operations and gaining familiarity with the patient's history, the complexity becomes apparent. This method of learning was used to improve the understanding of the contribution that x-ray images make in the management of these patients.

The reflective cycle was used to learn from observations of revision surgery for hip and knee replacements. This is illustrated by notes from a written reflection after observing the revision of a knee replacement following substantial osteolytic bone loss.

Learning points identified from observing revision knee surgery, May 2008:

1. *X-rays do not always reveal the extent of the underlying damage to the surrounding bone and pre-operative planning must take this into account. Use of appropriate x-ray views can eliminate some of the otherwise hidden damage* (Miura et al. 2005).

2. *A total knee replacement which has been in place for 16 years may be severely damaged by an osteolytic granuloma. Careful review of x-rays is required to identify early signs of these changes. A similar situation is seen in hip replacements* (Della Valle et al. 2009).

3. *Revision surgery is very demanding of time, resources and surgeons who need to be able to improvise and construct with what they have available and in the minimum of time. Providing them with the best available information pre-operatively will reduce the stress of the situation for the surgeon when the patient is on the operating table.*

4. *A decision to revise a joint must be very carefully considered because of the risks it involves.*

5. *The use of trabecular metal inserts is relatively new and the initial results suggest that they provide a useful option for orthopaedic surgeons* (Levine et al. 2006). *They are constructed with tantalum, an elemental metal, which is vaporised and then deposited in a way that gives a construction similar to trabecular bone. It is a biocompatible material and has been used in cardiac pacemakers for many years. The surgeon was able to use the trabecular metal components to reconstruct the damaged tibia.*

The problems with x-ray assessment of joint replacements and estimation of bone loss are well documented in the orthopaedic literature (Hozack et al. 1996, Saleh et al. 2001, Dumbleton et al. 2002). Although the researcher was aware of this concept, it was only through the observation of surgery on such patients that the implication became apparent. In particular, the experiences illustrated the use and limitations of x-ray images when assessing failing arthroplasty. It also increased the understanding of osteolytic damage and the difficulties that reconstructive surgery presents to the surgeon.

3.3.3.2 Case studies

Written case studies of six patients afforded further learning experiences. For each patient reviewed, the clinical reasoning process required an objective evaluation of the contribution of x-ray images. Relevant orthopaedic literature and colleagues were consulted to supplement the learning. The completion of the written records and the use of the reflective cycle highlighted concepts to be applied to future cases.

This is illustrated in the case of a male patient presenting with pain over the front of a hip which had been replaced five years previously. The pain was intermittent and associated with weight bearing activity. It had been present for three years and analgesia was required regularly by the patient. The researcher, in her role as arthroplasty practitioner, considered a number of possible diagnoses and used a process of clinical reasoning to determine the course of action (see Appendix II). The use of x-ray images was integral to this process. When the final outcome was known, the researcher used reflective writing to capture the learning about this complex situation and the contribution of image interpretation to the process.

3.3.3.3 Summary

At the same time as the development of image interpretation skills, another research study was initiated by the researcher to develop a simple clinical tool to be used in conjunction with x-rays for the assessment of osteolysis (see Chapter 4). The knowledge acquired in setting up that study, in combination with the methods described above, contributed to a greater understanding of the use of x-ray images in the management of patients with complex arthroplasty requirements.

3.3.4 Objective Four: To develop a local protocol for non-medical professionals requesting radiographic investigations in arthroplasty review

Further evidence of understanding the importance of this newly acquired skill was demonstrated by developing a local protocol. The use of a protocol provides clear boundaries within which non-medical health professionals can safely practice advanced skills. The process involved a search for relevant literature, contacting arthroplasty practitioners nationwide and accessing related discussions on interactive websites.

3.3.4.1 Literature review

A search of the English literature from 2000 to 2009 was undertaken using five databases: EMBASE, Medline, AMED, CINAHL and the British Nursing Index. The aim was to identify work on image interpretation by non-medical health professionals. The keywords used were: 'physiotherapist, nurse, non-medical professional, extended scope practitioner, allied health professional' and 'x-ray, image interpretation, image, radiograph, interpret, interpretation'.

Individual combinations of these terms were employed in order to explore the extent of the literature on this subject. Results were limited to use of these terms in only the title as preliminary searching indicated a large number of irrelevant documents. A search was also made on two authors known to have published on this subject.

Although 340 papers were identified, many were found to be irrelevant or of indirect interest only and only four papers were retrieved for full review. In addition, there were two papers which discussed the extended roles of health professionals and two reference documents of interest (see Appendix II).

There was one paper which referred to a specific protocol for imaging requests in ankle injuries but none for arthroplasty (Littlejohn et al. 2006). There were examples in the literature of the use of a protocol for requesting x-rays, but due to specialist roles being developed as a response to local situations, no universal proforma was available (Benger 2002, McPherson et al.

2006). Many of the papers referred instead to the use of a local protocol as recommended in the Royal College of Nursing (RCN) report (Royal College of Nursing et al. 2006).

The guidelines produced by the RCN and associated colleges set out the background and requirements for imaging requests from non-medical professionals. The emphasis was on local training and the development of a protocol in conjunction with local radiologists. The Royal College of Radiologists (RCR) permit non-radiologists to interpret images provided that the professional is part of a team with ready access to radiologists for advice (The Royal College of Radiologists 2006). The need for such an individual to demonstrate competency and advanced clinical reasoning is explicit in the RCN guidelines.

3.3.4.2 Protocol development

The combination of the literature review and use of the RCR and RCN reports gave some indication of the points to be included in a protocol. In addition, contact was made with other health professionals through ACPA and through an interactive physiotherapy website which enables clinical peers to interact and share knowledge (Chartered Society of Physiotherapy 2009). The ensuing discussion furnished examples of protocols used elsewhere although the continued development of the role of an arthroplasty practitioner meant that many of these were in need of review. However, the ability to interact with other health professionals electronically was integral to the development process and illustrates the contribution that technology can make in a situation of changing roles. The key learning point was the importance of dialogue on a national basis to supplement the literature in order to produce a relevant protocol.

The recommendations from the RCN guidelines were for the health professional to demonstrate adequate training and competency to refer patients for imaging and for a local protocol to be in place. Physiotherapists are eligible to refer for clinical images, as stated in the code of professional conduct, providing they have evidence of IR(ME)R (ionising radiation for medical exposure regulations) training and are aware of their boundaries (Royal College of Nursing et al. 2006). They should also have evidence of advanced communication and clinical reasoning skills.

3.3.4.3 Local protocol

Against this background, a protocol was produced for use locally in the district general hospital in which the arthroplasty service was being set up. It incorporated a description of the purpose and a list of the standards to which the requesting professional would conform (see Appendix: Development of Image Interpretation). The development through to the final version included a review by senior medical colleagues in orthopaedics, radiology and clinical governance within the local NHS Trust. It was also sent to a number of senor allied health professionals in major orthopaedic centres around the UK (Wrightington Orthopaedic Hospital, Royal Liverpool and Broadgreen University Hospitals, Royal Infirmary of Edinburgh).

The feedback from each of the reviewers was used to modify the protocol. The comments identified areas for clarification rather than any major changes and were beneficial in the development of a succinct but comprehensive document for local use. This protocol has now been accepted for local use and will be reviewed annually to reflect any changes in regulatory or professional requirements (see Appendix II).

3.3.5 Objective Five: To critically reflect on the unique position in the local organisation: the responsibility for and management of learning the skill of image interpretation to support the provision of a new service

The expansion of rôle from orthopaedic research assistant to arthroplasty review was locally unique. The only model existing in the local NHS Trust was that of specialist nurses in Oncology, advanced practice nurses who had developed their expertise 'on the job' but had not required external training. This unique position came with the responsibility to find ways to develop new skills with the support of local medical staff (Daker-White et al. 1999).

The formation of ACPA with subsequent opportunities to network with other arthroplasty practitioners around the UK was an important element in the development of image interpretation skills. For example, practitioners working for the Scottish Arthroplasty Project had an established method of in-house training and were willing to share the information. Contacts such as these were invaluable sources of ideas when deciding how to achieve the components of learning which were required.

When working in advanced roles, non-medical health professionals employ skills of clinical decision making in order to form provisional diagnoses (Daker-White et al. 1999). Communicating this information to the patient is an important part of the consultation. An arthroplasty patient will commonly experience significant pain relief and improvement of function following the primary surgery which allows re-engagement in social activities and relative independence from healthcare. For many patients, this is a positive move away from what Illich (1995) describes as 'social iatrogenesis' - disengaging from society and depending on healthcare due to a culturally induced image of what 'health' looks like (Illich 1995). When discussing with a patient the state of their artificial joint, there is a skill in emphasising positives and not heightening any apprehension about negative developments. As a health professional, the duty of care involves consideration of how the information is imparted in order to avoid social iatrogenesis. There is a need to weigh the risks and benefits for an individual with the intention of communicating an appropriate level of detail. This interactive reasoning is a

recognised part of advanced practice and was an additional part of the learning process (Jensen et al. 2000).

The opportunity to develop this new skill in response to the need for service delivery directly supported the care of local patients. The extended skills of the researcher have application in arthroplasty review and in research studies. There have been wider benefits from the contacts made with orthopaedic centres across the UK which continue to inform the local arthroplasty practice. It also had an unexpected local benefit in promoting inter-departmental relationships. The dialogue with and teaching from other local health professionals improved the mutual understanding of roles and subsequent interdepartmental communication.

3.4   Discussion

Acquiring the skill of image interpretation in arthroplasty was achieved through a multi-faceted approach which was firmly embedded in the principles of experiential learning (Kolb 1984). The need for this specific skill in a new arthroplasty follow-up service was seen as an opportunity despite a lack of clear training methods. Through a variety of learning techniques and the development of a tool to measure the new skill, competency in image interpretation of arthroplasty was established for a non-medical health professional working in orthopaedics. It required a strong personal commitment, a desire to learn and the frequent use of reflection. These are key elements for a practitioner working at an advanced level and have been shown to be important for developing new roles (Jensen et al. 2000).

The clinical implications of this process are multi-faceted. The practitioner benefited from learning a new skill and achieving a level of competence in the work environment through self-development (Honey 1997). The benefit for the employer, in this case an NHS Trust, was the addition of a supplementary service which enhanced the orthopaedic care already offered. For patients, the ongoing care has potential to prevent unwanted problems by timely screening and provides the advice and reassurance which are often needed in long term conditions. Although the threat of social iatrogenesis must be avoided, there are an increasing number of people with THA for whom this type of service is important (Dumbleton et al. 2002, Bozic et al. 2009).

In research terms, the ability to assess arthroplasty images provides an additional skill which is of use locally and is transferable to future work. The researcher can confidently use the acquired skill in measurement of parameters indicative of the status of a joint replacement. This will inform local data collection and the critical analysis of other research studies. In addition, documentation of the learning process increased the understanding of other possible learning methodologies. It demonstrated a pragmatic but consistent approach to training which is required for this level of advanced practice (Hardy et al. 2003, Kersten et al. 2007, Piper et al. 2009).

Nationally in the UK, the use of non-medical practitioners in orthopaedic pathways is likely to increase with added pressures to deliver surgical care in a limited time period (National Health Service 2009). The Department of Health has recognised the need for training of non-medical health professionals beyond traditional roles and has tasked organisations such as Skills for

Health (2009) with identifying the competencies required. A skill in image interpretation has been identified as one of the competencies that is required for an arthroplasty practitioner to conduct long term follow-up of joint replacements as part of an orthopaedic team. This account has described a method of achieving that competency.

When alternative methods of learning are successfully employed, it is important to share the model with a wider audience. The experiential learning described in the preceding account may provide a model of training in image interpretation to support the development of future arthroplasty practitioners. This model is more easily accessible than a formal training course and may be less costly. The disadvantages are that it requires significant commitment from the practitioner, medical colleagues and the HEI to complete the learning. This learning must then be captured in a formal competence assessment for it to be a transferable skill. A formal course has the advantage of providing a standardised assessment but does not allow for the experience gained through repetitive viewing of x-rays which is essential to develop the underlying knowledge base. Ultimately, a combination of both formal training and experiential learning may prove the best model.

In response to this situation, the Arthroplasty Care Practitioners Association set up an arthroplasty image interpretation course in conjunction with the University of Liverpool. The members of ACPA are offered a short course to provide basic training with suggestions of how to subsequently develop the acquired skill and demonstrate their competency to practice.

3.5   Conclusion

In conclusion, although there is limited evidence which demonstrates the effectiveness of non-medical health professionals undertaking image interpretation, there are many arthroplasty practitioners using this skill.   More high quality research is needed to evaluate this multidisciplinary professional development. Currently, the lack of formal education for the role has prompted individual practitioners to employ a variety of methods to achieve the advanced level of competency required.

This account has illustrated some of the learning techniques which have been employed to develop image interpretation skills in arthroplasty. Five learning objectives have been presented which show how the skill has been developed and how it has been incorporated into local service delivery. The learning has been captured in a master's level module of study undertaken as part of a doctoral degree which demonstrates academic rigour in the acquisition of an advanced assessment skill.

Further research is needed to test the protocols and the success of this method of training but the skill has immediate application in the arthroplasty review service in the local hospital. It is a transferable skill which is also applicable in orthopaedic research. Attaining competency in this skill was a precursor to obtaining research measurements used in long term follow-up of hip replacements. Further development of the skill and knowledge of the parameters used to assess the state of the hip replacement led to the production of a simple clinical tool. This tool was designed to be used in conjunction with x-ray images to monitor important changes of arthroplasty and is described in Chapter 4.

	DESCRIPTION OF EVENTS													
A C A D E M I C	Prepare for set up of arthro-plasty clinics ↑	ACPA conf-erence 2006 ↑	IR(ME)R training ↑			Formal teaching from radiologist and ortho-paedic consultant ↑	SWOT analysis ↑		Limb x-ray interpret-ation course ↑	ACPA conf-erence 2007 ↑	Edinburgh Summer University ↑		ACPA conf-erence 2009 ↑	Image Interpret-ation course ↑	Protocol develop-ment ↑
	JAN. 06	**MAR. 06**	**MAY 06**	**SEPT. 06**	**NOV. 06**	**DEC. 06**	**FEB. 07**	**MAR. 07**	**MAY 07**	**SEP. 07**	**AUG. 08**	**SEP. 08**	**MAR. 09**	**APR. 09**	**JUL. 09**
E X P E R I E N T I A L				↓ Weekly X-ray reviews begin	↓ Transfer to digital x-rays			↓ ACPA study day with case discus-sion				↓ Present to radiogra-phers			

Frequent discussions with radiographers and radiologists

ACPA committee: Image Interpretation course set up

← ← Observation of surgery and case studies → →

KEY:
IR(ME)R = Ionising Radiation (Medical Exposure) Regulations; ACPA = Arthroplasty Care Practitioners Association; SWOT = Strengths, Weaknesses, Opportunities, Threats analysis

Figure 3.2. Timeline of events in the development of Image Interpretation Skills

Chapter 4

Development of a radiographic assessment tool

The use of radiographic images is an essential part of assessing the state of a hip replacement. There are many features seen on an x-ray which are either classified or measured to indicate the condition of the components and the surrounding bone. This process is not standardised and there are a number of different methods currently employed to assess a hip replacement.

The information obtained from the radiographic image is used in conjunction with patient reported symptoms and other clinical signs to determine when further treatment might be needed. Although the decision to proceed with further treatment rests with the orthopaedic surgeon and the patient, the surveillance of patients is often conducted by another member of the orthopaedic team such as an arthroplasty practitioner. In cases where there has been a significant change in the signs or the symptoms since the last review, the type and quantity of these changes must be reliably conveyed to the surgeon for a decision on further treatment.

This chapter describes a simple and reliable tool to assist this process. A clinical tool was developed for use by the orthopaedic team in the radiographic review of hip replacements. It was designed to measure the osteolytic lesions that sometimes occur in the bone around a THA. It has potential to be included in the routine surveillance of hip arthroplasty and provides important information about the hip prosthesis.

4.1 Background

The survival of a hip replacement following the primary surgery is dependent on a number of factors. This includes the continued resistance to infection within the joint, the mechanical stability of the joint, durable properties of the component materials, the biomechanical stresses on the joint and the integrity of the surrounding bone into which the components are anchored (Duffy et al. 2005). A radiographic assessment of the components of a THA will include an estimation of the extent of any radiolucency, any migration of the acetabular cup or femoral stem and the presence of osteolysis because of the potential for reduced stability (Harris et al. 1982, Hodgkinson et al. 1988, Engh et al. 1990, Dall et al. 1992, Martell et al. 1993). Radiolucencies over 2mm in width are regarded as osteolysis in some studies and osteolytic lesions may be classified as linear or expansile (see §2.1.2) (Kobayashi et al. 1997, Paprosky et al. 2001, Incavo et al. 2008, Kim 2008, Utting et al. 2008, Altenburg et al. 2009, Santori et al. 2010). Osteolytic lesions are a particular threat to the stability of a THA if they become enlarged as the consequence is a reduction in the area of the THA component which is in contact with, and anchored to, the host bone (Mehin et al. 2004).

Surveillance of THA is important in order to identify any changes that may threaten the stability and survival of the hip joint and to facilitate timely intervention (Maloney et al. 1999, Wroblewski et al. 2000, Norton et al. 2002, Nixon et al. 2007, Wroblewski et al. 2007, Ihle et al. 2008, Makela et al. 2008, Ollivere et al. 2009). It is of benefit to the patient through increasing the chance of an improved outcome and a reduction of the associated comorbidity (Howard 2009). It is surgically and economically advantageous as treatment can be planned in advance, before the host bone is severely damaged, which minimizes the reconstructive work needed at surgery (Haddad et al. 2007). Radiographic assessment is an integral part of the surveillance of THA (British Orthopaedic Association 2006).

The development of osteolysis around a THA takes place over time and the patient is usually unaware of any symptoms until the underlying changes are advanced (Zicat et al. 1995, Maloney et al. 1999, Duffy et al. 2005). Osteolysis occurs as a biological response to wear debris particles produced from the articulation which subsequently alters the balance in bone turnover. Linear osteolysis, which extends along the prosthesis-bone interface, may lead to aseptic

loosening of the component. Expansile osteolysis, which extends away from the joint space, can destroy host bone leading to a loss of fixation of a THA component through loss of anchoring and carries a risk of periprosthetic fracture. Static osteolysis may not be a threat to the patient but progressive osteolysis is of concern and it is for this reason that monitoring is advised (Dumbleton et al. 2002).

The presence of osteolytic lesions was noted in over 70% of the studies included in the literature review (Chapter 2) but the methods employed for measurement varied between authors. They all identified location by zone but the quantification of size was inconsistent from one study to another. Some estimated the area by a calculation based on the length and width of a lesion but did not state if this was adjusted for magnification. Some categorized lesions into nominal groups based on a width measurement only. In view of the irregularities of many osteolytic lesions, these methods are imprecise and unable to provide reliable information from which to identify future changes in size.

Prior to revision surgery, there is often substantial bone loss from osteolysis and an assessment of this loss is made using a categorical scale to assist the planning for reconstructive work in surgery (Saleh et al. 2001, Della Valle and Paprosky 2004, Parry et al. 2010). However, it has been suggested that, in the earlier monitoring of osteolytic lesions, information should be obtained on the size, shape and location of individual lesions to identify progressive changes (Dumbleton et al. 2002). The size of a lesion has been measured in more precise ways than 'length x breadth', such as the use of geometric formulae and the manipulation of computer software, but these methods are not readily applied in routine practice (Han et al. 1999, Prevrhal et al. 2008, Hernigou et al. 2009).

Osteolytic lesions are a three-dimensional phenomenon and can be difficult to identify and measure on a two-dimensional image such as a radiograph. These difficulties are related to the size and location of the lesion and other imaging modalities may be more reliable in detecting small lesions (Kitamura et al. 2006b). The use of metal artefact reduction techniques in MRI and CT scans have improved the sensitivity of these techniques with CT being of particular use for the assessment of the bone quality (Cahir et al. 2009). The use of CT to estimate the volume of an osteolytic lesion is of particular value in pre-operative planning. However, it is associated

with a greater risk to the patient from significantly more radiographic exposure than for a plain x-ray and the higher cost is currently prohibitive in routine surveillance.

The sensitivity of a single antero-posterior (A-P) x-ray view for detecting osteolytic lesions in the pelvis has been quoted as 67% with a specificity of 72%, thus underestimating the presence of these lesions (Kitamura et al. 2006a). In order to improve the sensitivity, it has been recommended that a Judet oblique view of the affected hip be included together with an A-P view of the pelvis (Thomas et al. 2007). In addition, standardisation of the method for obtaining x-rays is recommended to identify the change in size of a lesion from the same perspective each time (ASTM International 2004, Toms et al. 2009). Standardisation requires consistent patient positioning and placement of the radiographic apparatus in order to produce images which are comparable when obtained over a period of time. By minimising the variation between serial images, an observer can more reliably attribute any changes in size of a lesion to progression of disease than to a difference in view. Adherence to these relatively simple measures contributes greatly to the radiographic assessment of THA.

In summary, the measurement of osteolysis is an important component of the radiographic review of THA. The measurement in follow-up studies appears to lack consistency and more complicated methods are unsuited to a clinical situation. The use of plain radiographs continues to be the basis of routine review for THA although other modalities, such as CT, can be used to provide additional information when required. The study that follows describes the development of a simple, clinical tool for the measurement of osteolytic lesions around a THA from plain radiographs.

4.2　Aims and objectives

The development of a new tool requires evidence of the validity and reliability for use as specified. The aim of this study was to complete the initial stages of this process for a clinical tool designed to quantify the area of osteolytic lesions which are seen on x-ray around a hip replacement.

The objectives of the study were as follows:

- To develop the concept of a morphometric grid to provide a simple clinical tool for measuring the area of osteolytic lesions that can develop after a total hip replacement
- To test the reliability of this clinical tool to measure osteolytic lesions on radiographic images when used by a representative range of healthcare professionals
- To conduct initial tests of validity of the clinical tool

By fulfilling these objectives, the tool would be suitable for clinical use and further studies could be conducted to establish other psychometric properties of the tool.

4.3   Methodology

4.3.1   Concept development

The measurement of aspects of the human musculoskeletal system, using surface or radiological techniques, has long been established in orthopaedics. The measurement of the shape or form is known as morphometry (Burwell 1997). There are a variety of techniques but the use of a two-dimensional grid has been shown, in biomedical research, to provide a method of quantifying the area of an irregular shape as long as the magnification and interval on the grid are known (Elias et al. 1971). When using such a grid, a cross intersect counting technique is applied. The grid is constructed with dimensions chosen according to the object to be measured, multiple grid lines are superimposed and the points at which these grid lines cross each other provide the 'points' to be counted. This concept was adapted to produce a morphometric grid suitable for the measuring of osteolytic lesions seen in the bone adjacent to the components of hip replacements on x-ray images.

The dimensions of the grid were determined with reference to the orthopaedic literature. In radiographic assessment of hip arthroplasty, 2mm is commonly used as a critical measurement of the width of a radiolucency (see §2.1.2). Radiolucencies less than this width are not considered to be a threat to the prosthesis but those of 2mm or more are thought to indicate a possible loosening although the length and progression of the radiolucency must be taken into account (Dall et al. 1992, Roder et al. 2003, Duffy et al. 2005, Hernigou et al. 2009). Consequently, a grid was developed, using a software programme (Photoshop, Adobe Systems Inc., USA), with 2mm between each small cross at the intersection of the invisible grid lines. This was saved as an electronic file on a personal computer which maintained the integrity of the information and was used to produce multiple grids on transparent film in conjunction with a laser printer.

The uniformity of the spacing from one grid to another was examined by measuring 15 fixed distances between crosses (three different lengths in five different places) on each of three grids using electronic callipers. These had been calibrated by the manufacturer to one hundredth of a millimetre (Absolute Digimatic, Mitutoyo, Japan). A satisfactory level of uniformity of the

distance was established for the intended purpose (see §4.4.1) and the grids were stored in a protective wallet for later use.

4.3.2 Validity and reliability

Any new assessment tool must be subject to testing for validity and reliability. The 'validity' is an estimate of the accuracy of the tool and is defined with reference to a specific population (Peat et al. 2002). It is dependent on reliability and is threatened by systematic error. For example, if the x-ray images were not produced in a standardised way, the measurements taken using the morphometric grid would not be comparable between images. Validity should be repeatedly tested on different samples of the population of interest (Bowling 2002). The 'reliability' refers to the reproducibility and consistency of the data produced and is independent of the purpose of the data collection (Sim and Wright 2000). It is threatened by random error so that, when the tool is used repeatedly, inconsistencies in the tool or the way it is used threaten the quality of the data produced.

To establish the external validity of a new tool, inclusion and exclusion criteria are required to describe the population to which it will be applied. In the case of the morphometric grid, it was intended to be used in conjunction with an x-ray for any patient who had received a THA and who had subsequently developed osteolytic lesions around the hip replacement.

The internal validity is described in a number of ways; it can be sub-divided into categories (see Figure 4.1) which establish different aspects of the total validity of the tool.

Content validity refers to the scope of a tool to measure a given concept. For a tool such as the morphometric grid, it is established as face validity by reference to experts in the field of knowledge (Sim and Wright 2000). The content validity of the morphometric grid was tested through use of a peer review process.

Criterion validity can be demonstrated by comparison of the tool with a 'gold-standard' using a concurrent method of testing (Bowling 2002, Peat et al. 2002). Further research will be needed to compare measurements obtained using the morphometric grid and CT scans, which are considered the gold-standard for osteolytic lesions. Research over a period of time will be needed to establish whether the measurements obtained using the morphometric grid are sensitive

to change but the inferences derived in this way may provide further confirmation of its validity (Pynsent 2001).

Figure 4.1. Types of internal validity
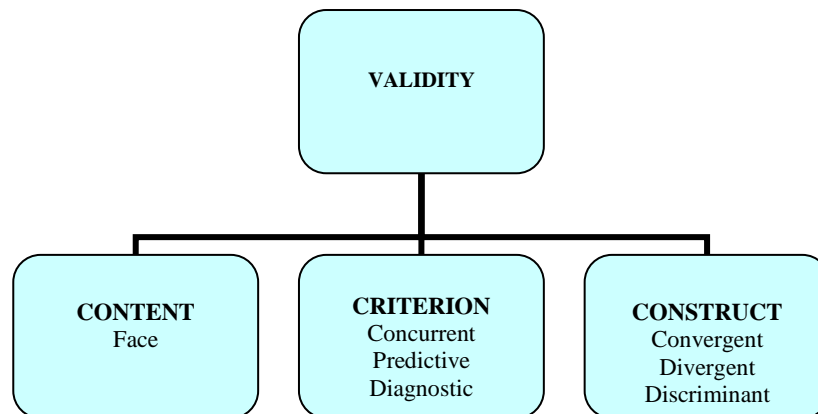
Construct validity is defined as the ability of a tool to measure a theoretical construct or trait (LoBiondo-Wood and Haber 1998). This can be established through correlation of the measurements obtained with another tool, either positive (convergent) or negative (divergent), or by the ability to distinguish between groups (discriminant). The 'contrasted groups' approach is one method of testing for discriminant validity. It involves using the tool to obtain measurements from sample groups, each of which represents a different part of the population of interest (Thomas and Nelson 1990). The results are then statistically tested to see if there is a difference between the groups. This was the method selected for testing the morphometric grid for construct validity.

Some authors consider that the sensitivity of an instrument to measure change is a form of construct validity (Streiner and Norman 2003) and others consider it to be a separate characteristic which should be included in the psychometric analysis (Pynsent et al 2004). The methods of measurement for sensitivity to change are less well-defined than for validity or reliability, and may involve statistics or relevant evidence (Pynsent et al 2004). If a statistical

method is selected, a dimensionless ratio is calculated although this is affected by variance in the sample and the nature of the change being measured (Streiner and Norman 2003).

A reliable instrument will demonstrate a minimum of error (Pynsent 2001). That is, when used by a range of observers, the scores produced will be consistent provided that the same set of circumstances exists. The testing is conducted with a group of observers who are representative of potential users. They obtain measurements using the tool and the data generated is tested statistically for inconsistencies. This may be in a test-retest situation in the hands of any one observer (intra-rater reliability), or the testing of inter-rater reliability (between observers) to assess the effect of numbers of observers using it to measure the same objects. For a new instrument, the testing of both is recommended by some (Morris 1997) while others suggest that, if inter-rater reliability is high, there is no need to explore test-retest reliability (Streiner and Norman 2003). Both inter-rater and test-retest reliability were assessed during the development of the morphometric grid to provide sufficient evidence for it to be used in a clinical situation.

In addition to estimating the reliability of an instrument, it has been recommended that the same data can be used to calculate a minimal detectable change (MDC) (Stratford 2004). This provides an estimation of the smallest change that can be detected by the measurement tool in the same units as the original measurement. Consequently, it can be used clinically to assess the likelihood of true change having occurred. This is in contrast to the dimensionless reliability coefficient which provides a clinician with an estimate of the consistency of the measurements obtained with the instrument.

4.3.3 Research design

The process of development of the morphometric grid was one of basic research from an empirical approach. Osteolysis is known to exist and a morphometric grid provided one way to measure the lesions occurring as a result of this phenomenon. The validity of this concept was proved through the peer review process; it was accepted as a poster presentation at the annual conference of the British Hip Society 2009 and the study was published in an international orthopaedic journal (see Appendix IX).

The testing of reliability and discriminant validity was conducted in three stages with a range of observers, as recommended in the development of a new clinical tool (see Figure 4.2)

Figure 4.2. Stages of testing in the development of the morphometric grid

```
┌─────────────────────┐
│ 1. Measurement of   │
│ simulated lesions   │
│                     │
│         ┌───────────┴──────────────────────────┐
│         │ a.) Produce simulated lesions in a   │
│         │ booklet                              │
│         │ b.) Four observers selected to record│
│         │ the measurements                     │
│         │ c.) Two observers repeat all the     │
│         │ measurements                         │
│         └──────────────────────────────────────┘
│
▼
┌─────────────────────┐
│ 2. Measurement of   │
│ actual lesions on   │
│ x-ray               │
│         ┌───────────┴──────────────────────────┐
│         │ a.) Identify and indicate osteolytic │
│         │ lesions on existing x-ray images of  │
│         │ THA                                  │
│         │ b.) Four observers record the area of│
│         │ each lesion                          │
│         │ c.) Four observers repeat the        │
│         │ measurements                         │
│         └──────────────────────────────────────┘
│
▼
┌─────────────────────┐
│ 3. Contrasted group │
│ testing             │
│         ┌───────────┴──────────────────────────┐
│         │ a.) Identify THA patients with       │
│         │ existing osteolytic lesions who have │
│         │ previously been categorised for the  │
│         │ state of the hip replacement         │
│         │ b.) Four observers record the area of│
│         │ each lesion                          │
│         │ c.) Calculate the total area of      │
│         │ osteolysis for each hip replacement  │
│         │ d.) Statistically test the data for  │
│         │ differences between contrasted groups│
│         └──────────────────────────────────────┘
```

(Wright and Feinstein 1992, Eliasziw et al. 1994, Morris 1997, Karanicolas et al. 2009). Currently, arthroplasty review is undertaken by medical members of the orthopaedic team or by non-medical health professionals who have received additional training specifically for this role. Consequently, four health professionals were selected to represent the range and levels of experience of those involved in arthroplasty review. They included the researcher (LKS,

Observer 1), a senior orthopaedic surgeon experienced in hip arthroplasty (Observer 2), a senior physiotherapist with more than eight years' experience of reviewing hip arthroplasty but no experience of quantifying osteolysis (Observer 3) and. an orthopaedic registrar (Observer 4). It was established that ethical approval was not required for this study (see Appendix IV).

4.3.3.1 Sample size

The sample size was determined by taking into account the number of observers and the type of statistical analysis to be used (see §4.3.5). A sample size of 20 to 50 observations is needed to obtain an intraclass correlation coefficient (ICC) with a standard error of less than 0.05 (Streiner and Norman 2003). Alternatively, four observers would each need to perform at least 30 observations for a hypothesis test against a "true" reliability coefficient of 0.90 at 80% power and $\alpha = 0.05$ (Eliasziw et al. 1994). A pragmatic decision was made to use 20 observations initially, as had been reported elsewhere, and 35 observations at the second stage (Bland and Altman 1986, Rankin and Stokes 1998). For test-retest reliability, the observers reviewed the set of lesions twice with the recommended interval of 14 days minimum between observations (Streiner and Norman 2003, Pynsent et al. 2004).

4.3.3.2 Stage One

Initially, the reliability of the morphometric grid to measure the area of osteolytic lesions was tested using 20 simulated lesions of varying size and of both recognised osteolytic shapes (expansile and linear). They were represented as line drawings on white paper (see illustration in Figure 4.3) and were reproduced in four booklets, one for each observer. The order was randomized so that there was a difference between each booklet. The intention of this was to reduce the bias which may have occurred due to a learning effect; that is, as the observers became more confident with using the grid, there was a chance of the measurements for any one lesion being more consistent. A description of how to apply the grid to a lesion and obtain a measurement was provided on the front page of each booklet, as illustrated in Figure 4.4. The crosses on the grid were referred to as 'points' because that is how they appeared to the observer. Instruction number four was specifically included to mimic a clinical situation: there are some

lesions that do not have clear edges and the individual observer must decide the extent of the lesion to be measured.

Figure 4.3. An example of a simulated osteolytic lesion with maximum dimensions of 2.0 x 0.9 cm. Observer ratings of the number of points on the morphometric grid were 29, 27, 29 and 32.

Additional booklets were produced with images in a new randomized order for repeat observations by two of the observers (Observers 1 & 2). The results from the measurements of the simulated lesions were analyzed before progressing to measurements on radiographic film.

Figure 4.4. Instructions for use of the morphometric grid with simulated osteolytic lesions

DEVELOPMENT OF CRITERIA FOR HIP ARTHROPLASTY REVIEW:
PHASE I – MORPHOMETRIC GRID TESTING

1. Place the grid over the shape.

2. Adjust the grid to maximise the number of points contained within the outline.

3. Count the total number of points inside the delineated area, including any on the line.

4. When the shape has no defined end to it, use an imaginary line that goes directly from one side to the other.

5. Directly adjacent to the shape which you have measured, please record the number of points you observed.

4.3.3.3 Stage Two

The second stage of testing was conducted on existing radiographs of hip joint arthroplasty with osteolytic lesions.   The x-ray images used had all been obtained by a standardised method in which the patient lies supine on the x-ray table and the x-ray tube is centred over the pubic symphysis (at the front of the pelvis) with a distance of 100cms from source to plate (ASTM International 2004).   The patient was instructed to 'turn the toes inwards' in order to internally rotate the femora and thus improve the imaging of the femoral component.   The lateral view was the iliac oblique view in which the patient lies on the affected side with the pelvis rotated 45° anteriorly (Thomas et al. 2007).   This allows better visualisation of the pelvic region posterior to the acetabular cup.   A request was made to the radiographers for the complete femoral component to be included in all views.

There were 35 lesions around 27 joint arthroplasties of which 11 were cemented and 16 were hybrid (cemented femoral stem, uncemented acetabular cup). The location of the osteolysis was in the pelvis on 8 images (23%) and in the femur on 27 images (77%); 28 (80%) of the lesions were expansile and 7 (20%) were linear; 20 (57%) of them were associated with an uncemented component, and the remainder with cemented components. The joint arthroplasties had been in situ for 4 to 16 years (average, 8 years) and the size of the lesions ranged from $0.04cm^2$ to $2.2cm^2$.

The method used to assess the inter-observer reliability of the morphometric grid was for each health professional to measure pre-identified lesions. The researcher (Observer 1) identified which lesion was to be measured but did not delineate the edges of the lesion precisely in order to allow the individual observer to decide on the size (Figure 4.5). When more than one lesion existed on a radiograph, the lesion of interest was clearly identified.   Not all lesions were included on any one radiograph as this study did not attempt to attribute clinical importance to the location or size of selected lesions, only to assess the reliability of the measurement tool.

Figure 4.5. Illustration of an application of the morphometric grid to one osteolytic lesion. The black line around the lesion did not define the extent of the lesion but identified the lesion to be measured. The ratings of the number of points for this lesion were 23, 21, 28 and 22.
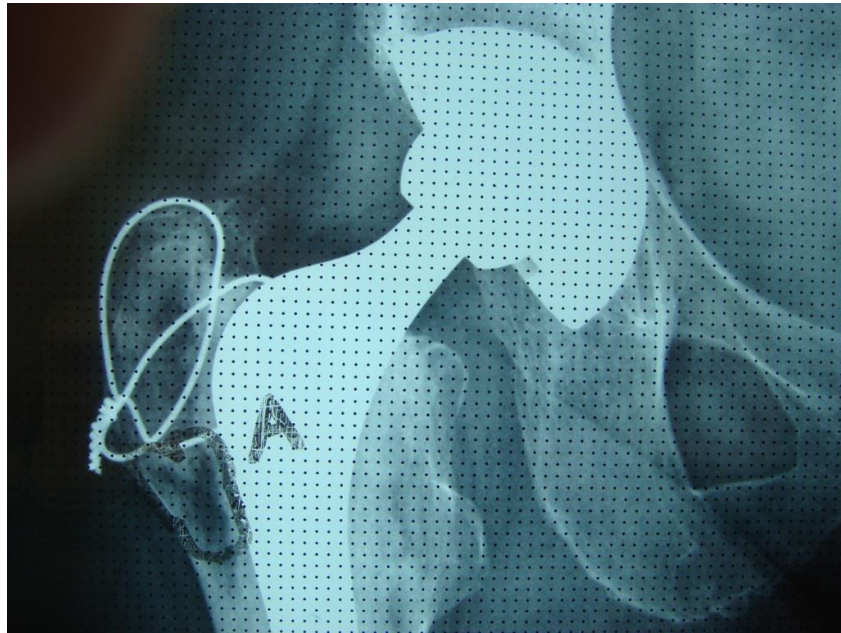
Instructions were provided for the placement of the grid and the measurement of the osteolytic lesions in a booklet. These were modified from the previous instructions to reflect the situations that occur when using x-ray images, as illustrated in Figure 4.6.

An osteolytic lesion was defined for the observers as a clearly darkened area in which no trabeculae were visible compared to the adjacent bone or prosthesis. The observers were blinded to the identity of the patient and had no access to older images of the hip replacement. Each observer recorded the number of points counted within the area of each lesion. The method of application of the morphometric grid is illustrated in Figure 4.5.

Figure 4.6. Instructions for the use of the morphometric grid on x-ray images

University of the
West of England
BRISTOL

Weston Area Health NHS
NHS Trust

DEVELOPMENT OF CRITERIA FOR HIP ARTHROPLASTY REVIEW:
PHASE II – USE OF MORPHOMETRIC GRID

1. Place the grid over the lesion which is defined as a clearly darkened area compared to the adjacent bone or prosthesis in which no trabeculae are visible.

2. If composite shadow complicates the picture, use the edge of the lesion as best estimated regardless of shading.

3. Adjust the grid to maximise the number of points contained within the lesion.

4. Count the total number of points inside the delineated area, including any on the edge of the lesion.

5. When the shape has no defined end to it, use an imaginary line that goes directly from one side to the other.

4.3.3.4 Stage Three

The construct validity of the morphometric grid was tested using a contrasted groups method. A series of x-rays had been categorised into groups prior to the commencement of this study by a senior orthopaedic surgeon in hip arthroplasty. The grading system had been in use locally for a year but had not been externally validated at this time.

There were three groups which were formed according to the severity of total radiographic changes around the THA. The first group (1) had been classified as 'poor' on the basis of extensive changes such as the progression of radiolucencies or the progression of osteolysis in the bone adjacent to the prosthesis. The second group (2) had been rated as 'moderate' with radiographic changes that were considered to require monitoring although they did not pose any immediate threat to the stability of the THA. The third group (3) had only 'mild' changes, such

as early stress shielding in the greater trochanter or a non-progressive radiolucency or a static osteolytic lesion, which were noted for comparison with future radiographic images. These groups were the basis for a comparison of the total area of osteolysis around each hip replacement measured using the morphometric grid.

Twenty-two patients, each with one THA, had one or more osteolytic lesions seen on x-ray. Eight measurements of each lesion were available from Stage Two. An average (mean) of the eight measurements was calculated to use as the size of each lesion for validity testing in order to eliminate bias attributable to any one observer. In cases where multiple lesions existed, the sizes of the individual lesions were summed to provide a single figure for each THA. This was considered to be a fair representation of the state of the hip replacement and surrounding bone since multiple lesions indicate greater bone loss and greater threat to the survival of the THA. The three categorical groups were then statistically compared using the final figure of osteolysis for each THA.

4.3.4 Bias

The measurements of osteolysis were made by four observers which included the researcher (Observer 1) who had developed the tool. This introduced a potential source of bias but was reduced by the use of inter-rater testing and review of the individual results to assess the effect.

4.3.5 Data management and analysis

The aim of the reliability testing was to estimate how much of the variability between measurements was due to the observers, how much was attributable to the subjects being measured and how much was due to random error. A coefficient of reliability expresses a relative proportion and is dimension-less. The result is interpreted with reference to the observers and the population from which they are drawn, and the number of measurements from which it has been derived. A value of ≥ 0.90 is considered to be high for clinical use (Streiner and Norman 2003). The same authors accept a lower value for research use (≥ 0.70) and other authors consider that any value between 0.75 and 0.90 demonstrates 'good' to 'excellent' reliability (Sim and Wright 2000).

Two methods of reliability testing were employed as this is recommended to provide a better understanding of the level of agreement between observations (Morris 1997, Rankin and Stokes 1998, Sim and Wright 2000, Pynsent et al. 2004). The first was the ICC which is a univariate test for relationship between multiple observations of the same variable using analysis of variance (ANOVA) for the calculations (Sim and Wright 2000, Streiner and Norman 2003). The second was the method of Bland and Altman (Altman and Bland 1983) which is independent of subject variability. It estimates both the error and bias included in the data and produces a graphical representation for ease of interpretation by clinicians (Bland and Altman 1986).

The assumptions for use of the ICC are that the various components are normally distributed and that there is no interaction effect between the subjects and the observers (Sim and Wright 2000). It is described in three different models, each of which has two forms (Shrout and Fleiss 1979). The models differ according to the assignation of observers to subjects and the type of observers. The two different forms within each model differentiate between a single measurement taken by each observer and a number of measurements from which an average is calculated (Table 4.1). The forms used in this study were the ICC (2,1) for inter-observer reliability with four observers and the ICC (1,1) for test-retest reliability of any one observer. A further specification was the use of 'absolute' agreement which tests the interchangeability of the observers in addition to the reproducibility of the measurements obtained. This is more stringent than a 'consistency' test because it requires agreement of absolute values (Eliasziw et al. 1994, Streiner and Norman 2003).

The formula for an inter-rater ICC, in which each observer makes only one measurement of each subject, is as follows (Rankin et al. 1998, Shrout et al. 1979):

$$\text{ICC (2.1)} = \frac{\text{Subject variability}}{\text{Subject variability + observer variability + random error variability}}$$

$$= \frac{\text{BMS} - \text{EMS}}{\text{BMS} + (k\text{-}1)\,\text{EMS} + k(\text{JMS-EMS})/n}$$

(Key: BMS = between subjects means square; EMS = residual mean square or mean square experimental error; JMS = between observers mean square; k = number of observers/judges; n = number of subjects/targets)

This has also been labelled as ICC 2 (A,1) when A = absolute agreement (Streiner and Norman 2003).

Table 4.1. Intraclass Correlation Coefficient models

Choice of raters	Assignment of raters to subjects	Measurements from each rater on a subject	
		Single	Mean of k
Randomly selected	Different set of raters assigned to each subject	Model 1, 1	Model 1, k
Randomly selected	Every rater assigned to each subject	Model 2, 1	Model 2, k
Not randomly selected (fixed raters)	Every rater assigned to each subject	Model 3, 1	Model 3, k

(Table used with permission of publishers. Sim and Wright 2000. ISBN 0 7487 3718 9)

The ICC (1,1), which was used for test-retest reliability, is a model with random people effects but the same sample set is measured on each occasion (Rankin and Stokes 1998). It can be expressed as:

$$\text{ICC (1,1)} = \frac{\text{BMS} - \text{WMS}}{\text{BMS} + (k-1)\,\text{WMS}}$$

(Key: WMS = within subjects means square; k = number of measurements)

The calculation of both test-retest and inter-observer reliability can be assessed with the ICC from a single set of data. A number of observers each take repeated measurements on a sample of subjects (Eliasziw et al. 1994, Sim and Wright 2000). It is referred to as a concurrent parallel-forms and test-retest study, and is advantageous in providing an increased sample size from which to calculate the reliability coefficients.

The Bland and Altman method was applied to two sets of results at a time and a graphic plot was produced to show the differences between each pair of measurements in relation to the mean of the pair. Limits of agreement (mean difference plus or minus two standard deviations from the mean) were then superimposed to provide a visual representation of the agreement between the

two sets of observations, as recommended (Bland and Altman 1986). The pairs of observers were chosen to reflect differences in profession and differences in experience.

The estimation of an MDC from a reliability study employs a standard error of measurement (SEM) calculation. The SEM is a value in the same units as the original measurements and indicates how much a score might vary from one time to the next when there is no true change in score. It can be calculated in four different ways depending on the information available although only three of these give values which are independent of sample size (Stratford 2004). The formula chosen for this study was as follows:

$$SEM = \frac{\text{Standard deviation of the difference in scores between observers}}{\sqrt{2}}$$

The MDC can be calculated for a given confidence level depending on the use to which the instrument will be put. It is suggested that an MDC_{95} be used in cases where the results will be used for significant life events such as surgical intervention (Donoghue and Stokes 2009). The formula used for this calculation is:

$$MDC_{95} = SEM \times \sqrt{2} \times 1.96$$

The Kruskal-Wallis test was used for the contrasted group analysis with non-parametric distributions. The level of significance for all the data analysis was set at $\alpha = 0.05$ and tests were conducted using SPSS® Version 15.0 software (SPSS Inc, Chicago, IL).

4.4   Results

4.4.1 Reproducibility of morphometric grid

The mean of each set of five measurements taken across grid points with digital calipers demonstrated <0.05mm variation (Table 4.2). This difference would not be visible to the naked eye and the method of reproduction of the grids was considered acceptable for them to be used as intended.

Table 4.2. Measurements taken across three morphometric grids (in mm)

Measurement length in mm	Grid 1 Mean (Range)	Grid 2 Mean (Range)	Grid 3 Mean (Range)
10 mm	9.97 (9.95 to 10.05)	9.97 (9.94.to 9.97)	10.00 (9.91 to 10.05)
20 mm	20.00 (19.90 to 20.09)	20.02 (19.93 to 20.09)	20.02 (20.00 to 20.04)
30 mm	29.99 (29.94 to 30.08)	30.00 (29.93 to 30.09)	29.99 (29.90 to 30.06)

4.4.2 Content validity

The use of a peer review process for a poster and a paper presentation of the morphometric grid, and subsequent correspondence, provided independent confirmation that the tool was appropriate for measurement of osteolytic lesions (see Appendix IX) and supported the content validity.

4.4.3 Stage one: results

There were no missing data and the raw data for observers 1 and 2, which were used for test-retest reliability, are shown in Table 4.3.

Table 4.3. Raw data measurements for Observer 1 and Observer 2 on simulated lesions

Lesion	Observer 1 Time 1	Observer 1 Time 2	Lesion	Observer 2 Time 1	Observer 2 Time 2
1	14	14	1	15	13
2	16	18	2	17	18
3	20	20	3	21	23
4	27	31	4	28	25
5	27	24	5	23	28
6	46	55	6	42	46
7	27	29	7	23	26
8	6	6	8	5	5
9	16	15	9	14	16
10	39	44	10	41	42
11	9	9	11	8	7
12	22	22	12	20	19
13	25	23	13	25	23
14	31	34	14	32	29
15	25	25	15	20	22
16	3	3	16	3	4
17	21	21	17	19	20
18	19	20	18	19	18
19	9	10	19	8	9
20	29	29	20	29	28
Mean	21.55	22.6	Mean	20.6	21.05

The results obtained from the reliability testing (inter-observer and test-retest) of the measurements from simulated lesions are shown in Table 4.4 (SD = Standard deviation).

Table 4.4. Results for inter-observer and test-retest reliability on simulated lesions

Reliability	Intraclass correlation coefficient		Bland-Altman				
	Coefficient	95% confidence interval	Range of points recorded	Mean difference	95% confidence interval for mean difference	SD	95% limits of agreement
Interobserver	0.93	0.75 to 0.98	3 to 55	0.95	0 to 1.9	2.04	−3.13 to 5.03
Test-retest: Observer 1	0.97	0.93 to 0.99	3 to 55	−1.05	−2.29 to 0.19	2.65	−6.35 to 4.25
Test-retest: Observer 2	0.98	0.95 to 0.99	3 to 46	−0.45	−1.49 to 0.59	2.21	−4.87 to 3.97

The distribution plots from the Bland and Altman method are shown below (Figures 4.7 & 4.8).

Figure 4.7. A distribution plot from a Bland and Altman test of the inter-observer reliability of the morphometric grid when used with simulated lesions. The mean of the differences is shown as a solid black line; the 95% limits of agreement are shown as dotted lines. (There are only 18 points shown on the graph as there were two sets of identical results).
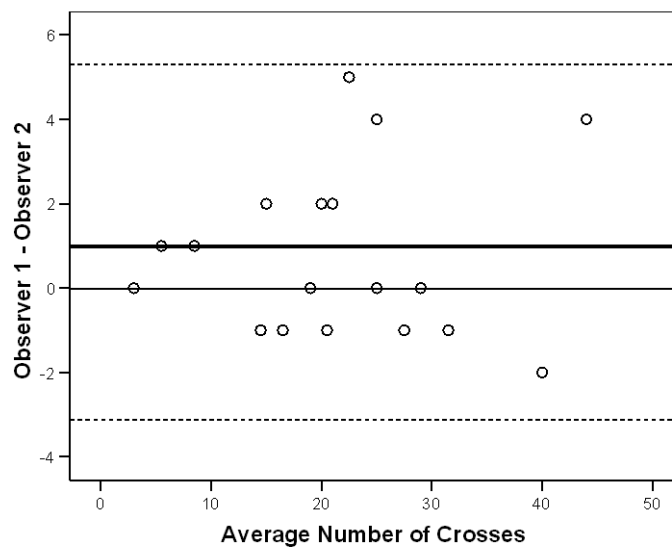
In Figure 4.7, the limits of agreement for observers 1 and 2 are -3.13 to 5.03, the range of which would be of concern for smaller lesions. However, it can be seen in the distribution plot that there was little variation for the smaller lesions but a greater variation between the observers on measurements of larger lesions. The potential variation with larger lesions contributed to the range of the limits of agreement and the effect was recognised and discussed (see §4.5).
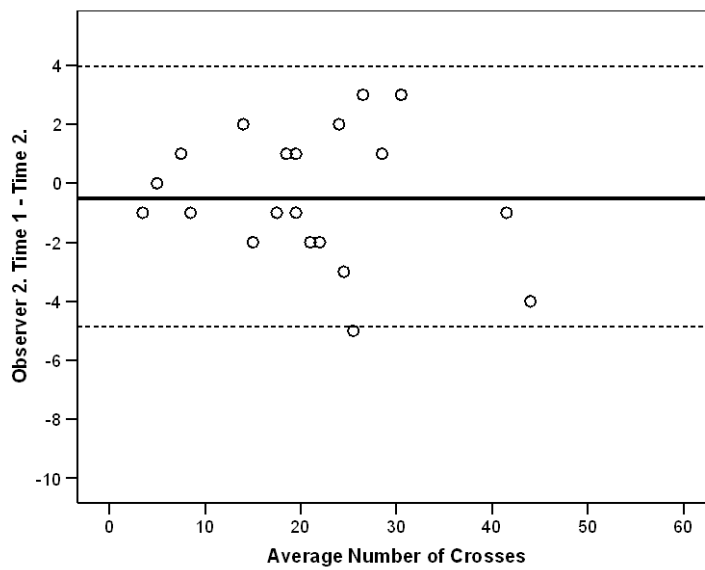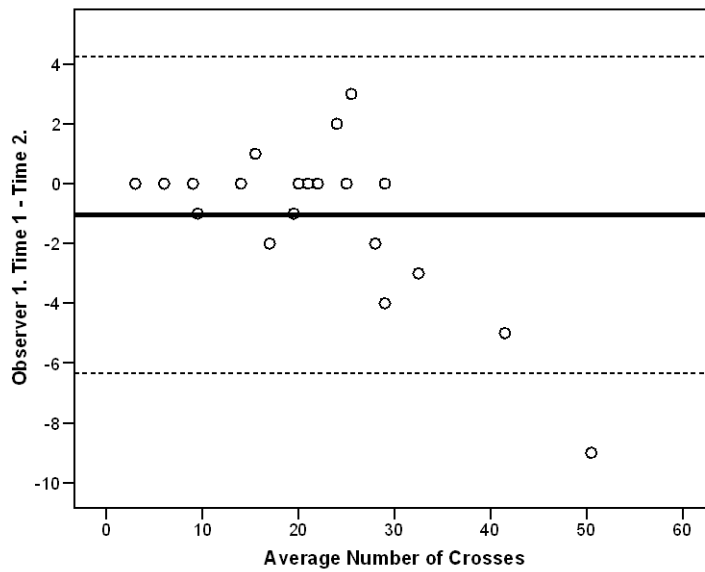
Figure 4.8. The Bland and Altman distribution plots of test-retest (intra-observer) reliability for Observers 1 and 2 using simulated lesions. The mean of the differences is shown as a solid black line; 95% limits of agreement are shown as dotted lines.

The limits of agreement for the test-retest reliability, seen in Figure 4.8, are of a similar range to those for the inter-observer testing (Figure 4.7). Also, the distribution plots shown in Figure 4.8 display a similar pattern to that in Figure 4.7 with less variation of agreement for smaller lesions.

4.4.4 Stage two: results

4.4.4.1 Inter-rater reliability (stage two)

There were no missing data when four observers used the morphometric grid to measure 35 osteolytic lesions on plain x-ray films. A summary of the raw data for the four observers is presented in Table 4.5.

Table 4.5. Data of the measurements from four observers of osteolytic lesions on plain x-ray films

	Number of points recorded on 35 observations		
	Minimum	Maximum	Mean
Observer 1, time 1	2	61	17.3
Observer 1, time 2	3	73	20.2
Observer 2, time 1	3	64	16.9
Observer 2, time 2	3	51	16.4
Observer 3, time 1	2	63	18.6
Observer 3, time 2	3	79	20.0
Observer 4, time 1	1	71	18.0
Observer 4, time 2	2	73	15.9

The results for inter-observer reliability of the tool when used on x-ray films, using the ICC (2,1) test, are shown in Table 4.6.

Table 4.6. Intraclass correlation coefficient results for inter-observer reliability on x-ray films

Interobserver reliability	Intraclass correlation coefficient	95% confidence interval
Time 1	0.96	0.93 to 0.98
Time 2	0.90	0.83 to 0.95

If the researcher's results (Observer 1) were removed, the ICC at time 1 was 0.95 and at time 2 was 0.89.

The Bland and Altman method was used to test inter-rater reliability between four pairs of observers. An illustration of one of the plots is provided (Figure 4.9) and the others are available in Appendix VI. The Bland and Altman plots for all four pairs of observers showed good agreement with most points lying within the limits of agreement and close to the mean which indicates that no significant bias existed between observers. The results are presented below in Table 4.7.

Figure 4.9. A Bland-Altman distribution plot for inter-observer reliability of the morphometric grid when used to obtain measurements of osteolytic lesions on 35 radiographs (Observers 1 and 2). The mean of the differences is shown as a solid black line; 95% limits of agreement are shown as dotted lines
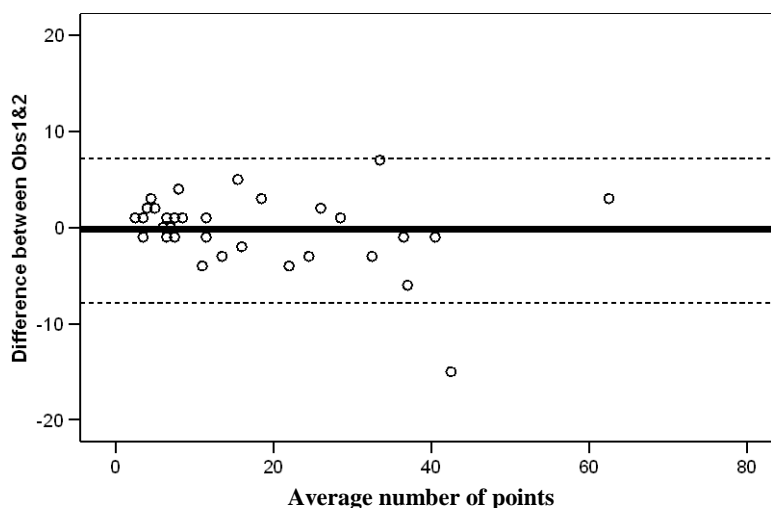
Table 4.7. Bland-Altman results for inter-observer reliability on radiographs at Time 1

Observers	Bland-Altman				
	Range of points recorded	Mean difference	95% confidence interval for mean difference	SD	95% limits of agreement
1 and 2	2 to 64	−0.31	−1.60 to 0.98	3.76	−7.83 to 7.21
3 and 4	1 to 71	−0.60	−2.30 to 1.10	4.95	−10.50 to 9.30
1 and 3 (physiotherapists)	2 to 63	1.31	0.02 to 2.60	3.76	−6.21 to 8.83
2 and 4 (surgeons)	1 to 71	1.03	−0.83 to 2.89	5.40	−9.77 to 11.83

SD = Standard deviation

The limits of agreement between observers 1 and 2 (-7.83 to 7.21) were slightly larger than on the simulated lesions but the range of points recorded was also larger. The distribution plot (Figure 4.9) shows a similar trend to previous plots with greater variation in agreement for larger osteolytic lesions.

The values from these inter-observer reliability tests were used to calculate an SEM and an MDC_{95} for two pairs of observers. The SEM for observers 1 and 2 was 3 points and the MDC_{95} was 4 points. For observers 2 and 4, the SEM was 4 points and the MDC_{95} was 11 points.

4.4.4.2 Test-retest reliability (stage two)

Table 4.8 presents the ICC (1,1) results for test-retest reliability for each of the four observers when measured with a minimum 14 day interval between observations.

Table 4.8. Intraclass correlation coefficient results for test-retest reliability using the x-ray films

Test-retest reliability	Intraclass correlation coefficient	95% Confidence interval
Observer 1	0.95	0.90 to 0.97
Observer 2	0.93	0.87 to 0.97
Observer 3	0.93	0.87 to 0.97
Observer 4	0.96	0.91 to 0.98

An illustration of one of the Bland-Altman distribution plots for test-retest reliability is shown in Figure 4.10 and further plots are available in Appendix VI.

Figure 4.10. A Bland-Altman distribution plot for the test-retest reliability of the morphometric grid when used to obtain measurements of osteolytic lesions on 35 radiographs (Observer 3). The mean of the differences is shown as a solid black line; 95% limits of agreement are shown as dotted lines
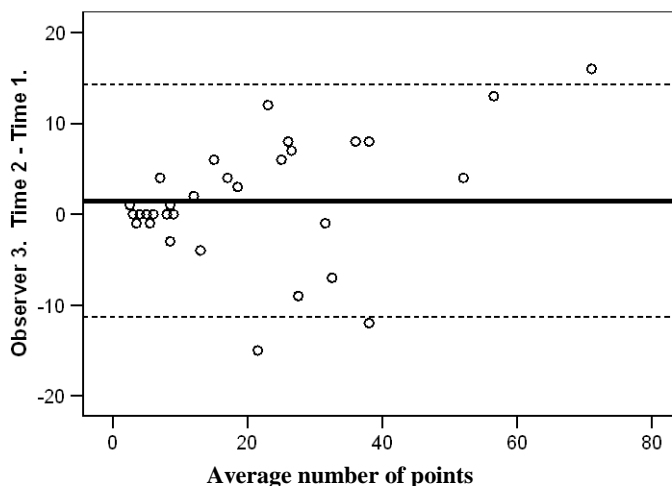
The Bland and Altman results for test-retest reliability on x-ray films (four observers) are presented in Table 4.9. The range of points recorded is similar in each case although the limits of agreement were wider for Observer 3.   The distribution plot for this observer (Figure 4.10) shows the same tendency for greater variation in measurement of larger lesions.

Table 4.9. Bland-Altman results for test-retest reliability on radiographs

Observer	Bland-Altman				
	Range	Mean difference	95% confidence interval for mean difference	SD	95% limits of agreement
Observer 1	2 to 73	2.9	1.6 to 4.3	4.0	-5.1 to 10.9
Observer 2	3 to 64	-0.5	-2.2 to 1.1	4.8	-10.2 to 9.1
Observer 3	2 to 79	1.5	-0.8 to 3.7	6.4	-11.46 to 14.3
Observer 4	1 to 73	-2.1	-3.6 to -0.6	4.4	-10.8 to 6.7

SD = Standard deviation

4.4.5 Stage three: results

There were 22 THA x-rays available for inclusion in the assessment of construct validity. Each hip replacement had previously been assigned a severity index following radiographic assessment

with four at grade one (worst), seven at grade two and eleven at grade three (best). The range in size of the total area of osteolytic lesions for each THA, as measured by the morphometric grid, was from 2 to 97 points (mean 25, median 24) as shown in Table 4.10.

Table 4.10. Summary of data for contrasted group analysis

Reference number of x-ray	Grade of severity of THA	Total area of osteolytic lesions measured with morphometric grid (number of points)
1	1	39
2	1	34
3	1	97
4	1	25
5	2	21
6	2	30
7	2	24
8	2	54
9	2	26
10	2	5
11	2	24
12	3	16
13	3	4
14	3	2
15	3	20
16	3	5
17	3	30
18	3	25
19	3	36
20	3	8
21	3	3
22	3	11

The Kruskal-Wallis test demonstrated a statistically significant difference in size of lesion between the three groups: X^2 (2, n=22) = 7.5, p= 0.02. The median score of cumulative lesion size for grade one (poor) was 18 points, for grade two (moderate changes) was 13 points and for grade three (mild changes) was 8 points. In view of the possibility that the outlier (no. 3, 97 points) affected these results, the analysis was repeated with this one value replaced at one unit more than the maximum value in group 1 (as recommended in Tabachnick and Fiddell 2001, see §5.3.8.3), that is, at 40 points. The median score of lesion size for each group was unchanged and the Kruskal-Wallis result was still statistically significant: X^2 (2, n=22) = 7.3, p= 0.03.

4.5   Discussion

This study has described the development of a concept into a simple clinical tool and the testing of the reliability and validity of the tool. This tool can be used to quantify osteolysis observed on radiographic images of hip arthroplasty. The testing of the reliability and validity was informed by the relevant literature (Karanicolas et al. 2009, Thomas et al. 1990).

Reliability of the morphometric grid to measure areas of osteolysis was initially tested with simulated lesions and four observers. This resulted in producing high coefficients for inter-observer and test-retest reliability ($\geq$0.90) which suggested that it was a good tool for clinical and research application (Streiner and Norman 2003). The agreement of inter-observer and of test-retest reliability was confirmed by observation of the Bland and Altman distribution plots which showed that there was no significant bias between observers. The limits of agreement were better for inter-observer reliability than for the test-retest reliability which may have been due to the learning effect for individual observers (see §4.3.3.2). The high values for inter-observer agreement allowed for testing to proceed to the next stage.

Testing continued with measurement of osteolytic lesions on radiographic films of THA and the coefficients remained high, indicating that it was a highly reliable tool when used to measure osteolytic lesions on x-ray films. The Bland and Altman plots for each of the pairs of observers showed good agreement. There was a slight bias in each pair, shown by the magnitude and sign of the mean, but in all cases the 95% confidence interval for the mean included zero which confirmed that the bias was not significant. The osteolytic lesions measured were found in both femoral and pelvic locations and represented the sizes and shapes (linear and expansile) of these lesions which are commonly seen. These results support the interchangeability of the morphometric grid when used by any health professional that is part of an orthopaedic team conducting THA surveillance.

Initial testing of construct validity indicates that the measurement of the total area of osteolytic lesions seen around a THA, using the morphometric grid, can differentiate between contrasted groups of patients. The patients were classified into three groups on the basis of the severity of their radiographic changes. There was a statistically significant difference between the groups when they were compared using only the measurements of osteolytic lesions obtained

with the grid. This supports the validity of the morphometric grid when used as a tool in the assessment of radiographic changes around a THA although further testing of the validity using an externally validated categorisation would be recommended.

A coefficient of reliability is a dimensionless property of an instrument and the value must be context based. The results show that this simple clinical tool can be reliably used to measure the area of osteolytic lesions in hip arthroplasty on a radiographic image. The ICCs were all 0.90 or more for a representative range of observers which indicates high reliability (see §4.3.5). This is an acceptable level of agreement for the information obtained to be used as part of the decision making process about the state of a hip replacement. If this tool was the sole source of information on which such a decision was based, a higher coefficient of reliability would be required (Streiner and Norman 2003).

The magnitude of the Bland and Altman limits of agreement initially appeared to be too large for small lesions but observation of the Bland and Altman distribution plots indicated that the greater difference was occurring between observers and between repeat observations for larger osteolytic lesions. This may be explained by the greater number of points to be counted which increases the chance of error. It could be improved by providing a grid with increased spacing (> 2mm) for larger lesions to optimize measurement by reducing the number of points to be included. Repetition of the reliability testing would be required to establish if this improved the agreement between observations.

The MDC_{95} for observers one and two was four points based on a group of 35 osteolytic lesions which ranged from two to 64 points when measured with the morphometric grid. This indicates that with a similar range of lesions and similar observers, one could be 95% confident that a change in size of lesion of greater than four points was a true change and not attributable to observer error. However, when a different pairing of observers (two and four) was used for the calculation from the same lesions, the MDC_{95} was 11 points. This may be due to the attention to detail of individual observers affecting the consistency of the data obtained with the morphometric grid or to random error. A change of 11 points represents a change in area of approximately $0.25cm^2$ which is of clinical significance but is unlikely to indicate a critical change in the condition of the THA such that surgical intervention is required immediately. This means that if the change in size of an osteolytic lesion was less than 11 points and the clinician

was unsure whether it was due to true change, it would be possible to delay a decision about treatment until a subsequent review when total change in size could be measured. Consequently, the values obtained for the MDC_{95} support the usefulness of the grid in a clinical situation.

There are limitations in using a two-dimensional tool to measure a three-dimensional object because the relationship between the area and the volume is not known. In studies of the relationship between area and volume of osteolytic lesions, one study was unable to find a direct correlation despite their data showing that the volume was 2 to 3 times the size of the area (Claus et al. 2003). Another study reported a positive correlation, especially when the lesion was greater than 10ml in volume (Kitamura et al. 2006a). The reason why lesions less than 10ml in volume have not been shown to correlate with volume calculations may be due to the difficulty of accurately defining the lesion. The dimensions (length and width) would be relatively small, around 2mm, and, although a computer software programme can be used to measure the area, it still requires an observer to manually delineate the lesion. Small changes in this drawing process and the computer interpretation of it will affect the calculation of the area. The morphometric grid was seen to be reliable with small lesions, and may provide a better estimation of area to correlate with the volume as it does not rely on computer interpretation of a hand-drawn image. However, this requires further research to investigate the relationship between numbers of points measured with the grid and corresponding volume measurements obtained from CT scans in a study of criterion validity. Further work is also needed to relate the number of points on the grid to the actual area of the lesion being measured.

This developmental study of the morphometric grid did not relate the size and location of osteolytic lesions to the clinical importance in the diagnosis of a failing hip replacement. A useful future study would be to measure the area and location of lesions seen around hip replacements in a cohort of patients that included some who were due to have a revision for aseptic loosening due to osteolysis. The relationship between the area, as measured with the grid, and the location of the lesion could be related to the state of the hip replacement. The results would provide useful information in the clinical diagnosis of THA patients and validation of the morphometric grid.

The role of the researcher (Observer 1) as an observer was a potential source of bias. However, when her set of observations was removed, there was no significant change in the

inter-observer reliability coefficient (see §4.4.4.1). In addition, the Bland and Altman distribution plots for Observer 1 were similar to other observers, suggesting that the effect of bias was minimal.

In the development process, some observers commented that it was occasionally difficult to see the points of the grid when superimposed on a black-and-white x-ray film. The researcher experimented with the production of coloured grids but no improvement was obtained. However, the use of supplemental light sources or manipulation of the images using the integral digital software was found to provide a means of overcoming the problem. The observers also noted that it was sometimes difficult to distinguish the outer limits of a lesion to estimate the size of an area using the morphometric grid. This has long been an issue with this type of tool and has not detrimentally affected the results produced (Elias et al. 1971). In the evaluation of THA radiographic images, it was noted that the observers became more confident with identifying the precise area as they became more familiar with the process. By testing both the inter-observer and the test-retest reliability, a statistical evaluation of the error was provided, and the high value of coefficients produced suggested that there had been no adverse effect from this particular problem.

An essential component of the surveillance of THA is the measurement of change over time with use of serial radiographs of the hip replacement. Differences in the magnification between radiographic films could potentially distort the comparison of lesions on serial images. However, the known diameter of the spherical femoral head provides a fixed measurement for calibration of radiographic films. If a retrospective comparison with plain radiographs is required, a formula based on the size of the femoral head can be used to compare the number of points observed (see Appendix III). Within the widely used digitized system for radiographic images, the integrated software allows resizing of an image and measurements can be calibrated to the diameter of the femoral head. Transfer of the morphometric grid to a digital platform is currently being explored and it is recognized that this must include the consideration of measurement integrity if the images are manipulated. Testing of the sensitivity of the grid to measure change over time (Peat et al. 2002, Pynsent et al. 2004) was not possible in the initial development of the morphometric grid but was explored further in Chapter Five (see §5.3.7.7).

The aim of THA surveillance is to identify changes which threaten the survival of the hip joint in order to initiate further investigations and/or treatment if required. Expansile osteolytic lesions in the pelvis are often asymptomatic and pose a significant threat if they remain undetected (Needham et al. 2008). Monitoring of such lesions is widely recommended but has often been conducted using visual estimation of the area with a length and breadth measurement. This lacks reliability due to the limitation of the naked eye to spot small changes in area and due to the irregularity of the shape of lesions such that they cannot be defined by two linear measurements. The morphometric grid is specifically designed to record the area of irregular shapes using points spaced at 2mm, a distance that can be easily seen with the naked eye. Consequently, it has potential to be incorporated into the THA assessment process to improve the monitoring of patients. It is a clinical tool which is readily applied without the need for mathematical or software expertise, rendering it accessible to any healthcare professional involved in arthroplasty review. From the patients' perspective, the morphometric grid provides a simple method of demonstrating the change in size of a potentially damaging asymptomatic osteolytic lesion. This enhances their understanding of the condition and its implications.

The use of imaging modalities such as a CT scan or an MRI will provide a three-dimensional assessment of osteolytic lesions when needed. However, the cost associated with these investigative procedures, the risk to the patient from the radiographic exposure in CT and local availability are all barriers to their use in routine surveillance of hip arthroplasty. Consequently, plain x-ray images continue to form the basis of long term follow-up and the morphometric grid is designed to be used in conjunction with these x-ray images (Claus et al. 2003, Stamenkov et al. 2003, Kitamura et al. 2006a, Thomas et al. 2007).

The use of the morphometric grid to measure periacetabular osteolytic lesions may provide evidence of the prevalence of osteolysis for specific acetabular components. This information would be important to a surgeon in his or her selection of a prosthesis (Stulberg et al. 2002). It has been suggested that, for periacetabular lesions, the proportion of the circumference of the cup occupied by the lesion should also be measured (Mehin et al. 2004). This information may be of use to a surgeon in a decision about the need for revision surgery. Mehin (2011) has suggested that the area of a lesion measured with the morphometric grid should be related to the

circumferential involvement to predict acetabular instability. This would improve the quality of the information available to the surgeon in the monitoring and decision process (Smith 2011).

4.6   Conclusion

This study developed a simple, radiographic tool to quantify the area of osteolytic lesions seen on x-ray images following hip arthroplasty. These important radiographic changes potentially threaten the stability of the implant but their commonly asymptomatic nature necessitates routine follow-up to identify their presence and to monitor their size. The concept of a morphometric grid has been expanded to produce a tool specific to the surveillance of hip arthroplasty. It has the potential to support the THA assessment process and reliability for interchangeable use by any qualified health professional involved in arthroplasty review. The surveillance of THA is a multifaceted process in which radiographic assessment is only one aspect but the use of a simple, reliable, clinical tool could improve the monitoring process and may be useful for future research.

Chapter 5

Clinical study

Previous chapters in this thesis have illustrated aspects of the role of a non-medical practitioner in reviewing hip arthroplasty. In particular, the acquisition of the skill to interpret x-ray images was described in Chapter 3 and the development of a tool to measure significant radiographic changes was described in Chapter 4. In practice, the surveillance of hip arthroplasty includes clinical assessment in addition to radiographic review and it is a combination of the data gathered that informs the decision about the state of the joint replacement and the host bone.

Although the clinical assessment of patients with THA is recommended by the British Orthopaedic Association (2006),   there is no uniform procedure and the methods employed vary from unit to unit, as does the frequency with which patients are reviewed. Evidence from the orthopaedic literature provides some guidance on the post-operative period during which signs of deterioration may begin to occur but this can differ between patients and prostheses.

This chapter describes a clinical study involving a cohort of patients for whom previous details of arthroplasty review were available. The purpose was to review these patients again at a time when signs and symptoms of deterioration in the hip arthroplasty were beginning to appear. New data from validated outcome measures and radiographic review were collected and compared with the previous details of each participant. Associations between the data were explored using appropriate statistical techniques. The results are used to further develop the assessment processes used in hip arthroplasty review.

5.1   Background

THA was the first joint replacement to become widely used as long ago as the 1970's and long term survival of the cemented stems and cups has been documented (Williams et al. 2002, Wroblewski et al. 2007). The later addition of the hybrid hip replacement with an uncemented metal acetabular cup and polyethylene insert and a cemented femoral stem has had mixed success (Barrack et al. 1997, Maloney et al. 1999, Malchau et al. 2002, McCombe et al. 2004, Utting et al. 2008, Hernigou et al. 2009,). Further developments include the completely uncemented THA (both cup and stem) or the reverse hybrid with a cemented cup and an uncemented stem.

Joint replacements do not last indefinitely and some will require revision during the life time of the patient. Joint registers such as the National Joint Registry in the UK and the Swedish registry show that approximately 10% of the total number of THA are revisions (Malchau et al. 2002, National Joint Registry 2010). In the United States, a registry has only been recently been established (American Association of Orthopedic Surgeons 2010) but estimates of revision are similar to those in the UK (Bozic et al. 2009).

Failing hip replacements can be broadly categorized into two groups – septic and aseptic. Septic hip arthroplasties are those with an acute or chronic infection in the joint which may be present from the immediate post-operative period or may occur later as a result of introduction of an organism from another source. Patients with an infected joint replacement usually experience pain and pro-actively seek a medical opinion.

Aseptic failure of THA excludes infection as an underlying cause and can occur for a number of different reasons, as described earlier in Chapter 2. The signs and symptoms of an apparently aseptically failing THA are varied and will be related to the underlying cause. Patients presenting with a painful THA are questioned about their pain and medical history, are given a thorough physical examination of the hip and associated joints, and a current x-ray of the hip is closely analysed. If necessary, further investigations are ordered to support or refute the preliminary diagnosis. These include blood tests, joint aspiration, anaesthetic injection, CT and isotope scanning (Duffy et al. 2005). In some centres, a specialised technique of MRI, known as metal artefact reduction sequencing (MARS), is also available (Cahir and Toms 2009). The

focus of these investigations is to exclude infection and other medical conditions as a cause of the pain, and to identify what mechanism of failure has taken place.

For a painless THA, there are rarely any symptoms to assist the diagnosis although some patients notice functional changes. The changes seen on x-ray are the primary method of diagnosis of failure and in units where monitoring of THA is well established, serial x-rays traditionally form the basis of arthroplasty surveillance (Wroblewski et al. 2002).

The monitoring of all patients with a hip replacement has been recommended in order to identify any adverse changes and intervene in a timely fashion to improve the chances of a good outcome for the patient (British Orthopaedic Association 2006, Wroblewski et al. 2007, Utting et al. 2008). However, it is not mandatory and practice varies from one hospital to another. The tools used in the assessment process are also varied including surgeon estimation of hip function, patient completed outcome scores, physical examination, history taking and evaluation of x-rays. The use of PROMs is increasing as they are shown to be effective and sensitive, and can be administered by post (Dawson et al. 1996a, Ostendorf et al. 2004, Department of Health 2008).

The frequency with which these assessments take place is mixed although it is currently recommended that they are at five year intervals up to ten years post-operatively and then more frequently thereafter (British Orthopaedic Association 2006). Traditionally, most of them have taken place in the acute hospital in which the original surgery was performed but there are suggestions that this type of service should be offered in primary care (Haddad et al. 2007). With the shift over time from surgeons to non-medical health professionals carrying out the review, there are implications in terms of responsibility and the methodology of the review process. While many of these health professionals are trained for the job by their local surgeons, nationally designed guidance and simple, reliable tools would support training and competency of future professionals in this role.

The economic constraints of healthcare delivery remain a significant challenge and annually increasing numbers of primary THA are likely to result in an increasing number of revision surgeries (Burns et al. 2006, Bozic et al. 2009, National Joint Registry 2010). In this environment, the cost of a surveillance service is more difficult to justify and any financial support must be used to best effect. Some units have attempted to set up a 'virtual' surveillance service in order to reduce the cost of clinical time (Hugill et al. 2010). In other centres, a lack of

funding has led to suggestions for the use of outcome scores alone (Price 2010). This is in contrast to the suggestion that the combination of a simple radiological review and a subjective measure of outcome are the essential elements of long term surveillance (Malchau et al. 2005, Ollivere et al. 2009).

In view of these issues, arthroplasty review must be cost-effective. This requires evidence to identify the most reliable methods of review, the patient groups most susceptible to adverse change and the ideal frequency of review for individual types of hip replacement. The methods chosen should allow any suitably trained member of the orthopaedic team to perform the review with appropriate support.

Consequently, a clinical study was designed to investigate some of these issues in a group of patients who were beyond five years and before ten years after the primary surgery, a period when the major signs of deterioration of a hip replacement often first appear (Malchau et al. 2005, Wroblewski et al. 2007, Hallan et al. 2010). The intention was to explore associations between the changes seen on x-ray with changes in a hip specific outcome score in order to provide evidence about the methods employed. In addition, the data were explored for sub-groups of patients at greater risk of adverse change.

Although one previous study had suggested that clinical results were not linked to the mechanical state of the joint as indicated on x-ray, the statement was not supported by statistical evidence in the results (Wroblewski et al. 2002). Other studies have used an outcome score at one time only to relate to the presence of radiographic symptoms and conclusions are mixed (Utting et al. 2008, Hernigou et al. 2009, Ollivere et al. 2009). The present study was designed to use the change in score on a joint specific outcome tool, as recommended (Murray et al. 2007), and to relate this to x-ray changes over the same period of time.

It has been suggested that a measure of general health and a comorbidity index should be added when assessing arthroplasty in order to capture other changes in health (Ostendorf et al. 2004, Bjorgul et al. 2010). One of each type of measure was incorporated into this study and the age of the patient was also included to assess the possible effect on radiographic changes. Information from the physical examination of patients was not utilised because of the difficulty of ensuring reproducibility of measurements such as the range of motion, and because it would necessitate a clinical interface in future arthroplasty review which increases the cost of

surveillance (Holm et al. 2000, Soderman and Malchau 2001). In addition, the chosen hip outcome score was designed to include an indication of the range of movement from the patient's answers to questions about function (Dawson et al. 1996a).

5.2   Aims and objectives

5.2.1 Research question

There was evidence in the orthopaedic literature and from knowledge of current practice that the use of outcome scores and radiographic review were normal practice. However, there did not appear to be any statistical evidence to support the need for both or to explore the associations between them at the mid-term review. Consequently, the research question was formulated as follows:

> *For patients with total hip arthroplasty undergoing a mid term review, is an x-ray needed in addition to a specific hip outcome score which measures their pain and function?*

The assumption was that, if a hip outcome score were sufficiently sensitive to indicate the state of the underlying THA, then an x-ray would not be needed at the mid-term review. This would be a cost saving and would reduce the risk to the patient of radiographic exposure.

5.2.2 Research hypothesis

The research question was formulated into a research hypothesis centred on the Oxford Hip Score (OHS) which is an outcome score specific to hip arthroplasty. It stated that:

> The magnitude of change in an individual's OHS between early and mid-term review will be associated with the number of radiographic changes around the total hip replacement over the same period of time.

The null hypothesis was that there was no association between changes in the OHS and radiographic changes at mid-term review of a total hip replacement. In the event that the null hypothesis was retained, the radiographic changes may be explained by other patient factors included in the exploratory analysis.

5.2.3 Aim

The aim of the study was to explore an association between change over time in the OHS and radiographic changes. The use of an exploratory technique allowed the inclusion of other measures which might affect the THA. The other measures were a general health index, a comorbidity score and the age of the patient.

5.2.4 Objectives

The study objectives were defined as follows:

- To obtain an OHS, a general health measure and a comorbidity score from each patient in a cohort with total hip arthroplasty completed six to nine years previously
- To assess the radiographic status of the hip replacement at six to nine years and to measure the changes from the early review of each of these patients
- To compute the change in OHS from those obtained at the early review
- To explore associations in the data between the radiographic changes and the change in OHS, the general health, existing comorbidities and age
- To assess the need for an x-ray to be included at mid-term review

5.3 Methodology

5.3.1 Research design

This clinical study involved exploratory research using longitudinal, observational methods. A cohort of patients from a district general hospital had previously been recruited to an observational study three years after their primary total hip replacements. This second study was designed to review the same cohort in the mid-term period, six to nine years after their initial surgery.

The participants were originally identified as consecutive patients for total hip replacement receiving one of two acetabular components available at that time. The choice of component was determined by age and activity level but was subject to the orthopaedic surgeon's discretion in discussion with the patient.

5.3.2 Setting

The primary surgery took place between the years 2000 and 2003 at Weston General Hospital (WGH), Weston-super-Mare, Somerset, UK. The operations were performed by six orthopaedic surgeons who were all permanent staff in the Department of Trauma and Orthopaedic Surgery at that time. One of these surgeons was a specialist hip surgeon but none of the other five had a special interest in hip arthroplasty.

The primary surgery was carried out using either an anterolateral or a posterior approach to the operative site depending on the surgeon's preference (both are standard surgical approaches). All surgery took place in a laminar flow theatre which is specifically designed to reduce the possibility of intra-operative infection. Post-operative care was the same for all patients with mobilization commencing the day after surgery. Discharge from hospital was generally four to five days after surgery when the patient was considered, by a physiotherapist, to be safe to mobilize at home.

In common with many units around the UK at that time, it was standard practice to use a cemented acetabular component with a cemented femoral stem (cemented THA) in an older patient, and an uncemented metal cup with a polyethylene liner and a cemented femoral stem

(hybrid THA) in younger patients (Emms et al. 2010). The break in age was approximately 70 years but the decision for an individual patient was made in conjunction with the surgeon and with consideration of all relevant health factors. A cemented THA was the treatment of choice for patients when the bone stock was of poorer quality, which requires fixation with cement, and was considered to have good long term results (Berry et al. 2002). The hybrid THA was designed to allow osseointegration of the cup to promote longevity of fixation in order to counteract aseptic loosening and was, therefore, usually selected for younger patients (Gonzalez Della Valle et al. 2004).

In the previous study, all participants were seen at three years (mean 3.1 years) after the primary surgery, by the researcher (LKS), to assess the survival of the acetabular component (Smith et al. 2008). An OHS was obtained and a radiological assessment of the acetabular cup was completed by a senior orthopaedic surgeon in hip arthroplasty. The present study took place between the years 2008 and 2010 which provided an interval of six to nine years for each participant from the original surgery.

Information brought forward from the previous study included the results of the outcome score and the data from the radiographic assessment – the angle of inclination of the acetabular cup, the number and location of radiolucencies or other changes in the surrounding bone and a linear wear rate for the polyethylene component (the cup in cemented prostheses and the liner in hybrid prostheses). The state of the femoral prostheses had not been recorded as part of the study but was available retrospectively via analysis of the radiographic films (see §5.3.6.1).

5.3.3 Arthroplasty components

The cemented acetabular component used for these participants was a Cenator cup (Corin Medical, Cirencester, UK) which is a high density polyethylene flanged device with the option for an extended posterior wall. It was manufactured from GUR 1050 resin by a ram extrusion process and the final product was gamma sterilized in an inert gas atmosphere. This reduces deterioration of the product on the shelf and improves the resistance to wear in vitro by moderate cross-linking of the polyethylene.

The uncemented cup was the EPF Plus (Smith & Nephew UK Ltd., London) which is an uncemented pure titanium cup coated by plasma spray with open-pored pure titanium. This

produces a base layer of 50μm depth and a coating of 250μm. An hydroxyapatite coating of 50μm is applied simultaneously with the top coat to promote bony ingrowth by the roughness of the surface created. The cup is equatorially expanded with screw options and a polished interior surface. The liner is ultra-high molecular weight polyethylene produced from GUR 1020 resin by a compression moulding process, packed in nitrogen and sterilized by gamma irradiation.

There were three femoral components used, all of which were cemented (see Table 5.2). The Exeter V40 stem (Stryker (UK) Ltd., Newbury, Berkshire) is a stainless steel modular component used with a 28mm metal femoral head; the CPS plus stem (Smith & Nephew UK Ltd., London) is also modular, manufactured from a metal alloy and used with a 28mm head; the Charnley stem (DePuy International, Leeds, UK) is a monoblock stem of metal alloy with a 22.22mm head.

5.3.4 Outcome measures and scores

5.3.4.1 Oxford Hip Score

In the previous study of this cohort, the OHS was used as a PROM (Smith et al. 2008). This tool had been specifically selected at that time because of its qualities, which will be discussed further in this section. It was re-applied in the present study to construct a change score for each patient as recommended in non-randomised studies (Murray et al. 2007). A change score eliminates some of the confounding factors introduced by participant interpretation as it is the magnitude of change that is compared.

The OHS is a 12-item questionnaire completed by the patient. It asks a patient about pain and function around the hip replacement (see Appendix X). It is specific to hip arthroplasty and has been shown to demonstrate a relatively large effect size in comparison with less specific outcome measures (Dawson et al. 1996a, Fitzpatrick et al. 2000, Ostendorf et al. 2004, Garbuz et al. 2006). An effect size indicates the magnitude of change in an outcome measure. It has been suggested that the OHS is sufficiently responsive to be used when only subtle changes are expected between groups of patients (Dawson et al. 1996b).

The OHS was originally shown to be valid and responsive to the change from before hip replacement to one year afterwards and it has now been adopted nationally for outcome assessment of hip prostheses in England (Dawson et al. 1996a, Department of Health 2008). The

score remains relatively static between one and five years after surgery (Field et al. 2005, Pynsent et al. 2005) but has been shown to be sensitive to failure of at least one arthroplasty component (cemented femoral stem) at mid-term (Ollivere et al. 2009). In addition, its validity for use before and after revision hip replacement has been demonstrated (Dawson et al. 2001). The mid-term review period of this study was beyond the period in which the OHS is known to be static and coincides with the time at which radiographic changes around the arthroplasty are known to appear (Malchau et al. 2005, Wroblewski et al. 2007, Hallan et al. 2010). Consequently, the use of the OHS to explore associations in the data was considered appropriate.

5.3.4.2 EQ-5D questionnaire

EuroQol have developed measures of general health, one of which is the EuroQol 5-dimension questionnaire (EQ-5D) (EuroQol Group 2010). It is recommended that a participant complete a general health score prior to a joint specific score in order to capture other health information and thus, reduce the potentially confounding effect of physical or psychological changes not associated with the joint on the joint specific score (Ethgen et al. 2004, Garbuz et al. 2006, Pollard et al. 2006, Murray et al. 2007). The EQ-5D is a patient reported outcome measure which consists of five descriptive questions with three possible responses to each. It has been developed by the EuroQol Group to provide an international measure of health and the scores constructed from responses can be compared to normal populations by country and by region (Szende et al. 2007). It is designed for self-completion, has been recommended in conjunction with the OHS and is now being used in a similar way in THA follow up elsewhere (Malchau et al. 2002, Ostendorf et al. 2004, Murray et al. 2007). It is quick to complete and does not place undue response burden on the participant.

5.3.4.3 Charlson comorbidity index

A comorbidity score provides an indication of the level of co-existing illness and is completed by the researcher rather than the patient. The Charlson comorbidity index was chosen for this study as it is a weighted index of comorbidity with age (see Table 5.1) (Charlson et al. 1987).

Table 5.1. Charlson weighted index of comorbidity

Assigned weights for diseases	Conditions
1	Myocardial infarct
	Congestive heart failure
	Peripheral vascular disease
	Cerebrovascular disease
	Dementia
	Chronic pulmonary disease
	Connective tissue disease
	Ulcer disease
	Mild liver disease
	Diabetes
2	Hemiplegia
	Moderate or severe renal disease
	Diabetes with end stage organ damage
	Any tumour (initially treated in the last 5 yrs)
	Leukaemia
	Lymphoma
3	Moderate or severe liver disease
6	Metastatic solid tumour
	AIDS

The age of the participant is assigned a value as follows:

- <50 years = 0
- 50-59 = 1
- 60-69 = 2
- 70-79 = 3
- 80-89 = 4
- 90-99 = 5

The value assigned to age is added to the summed value of all existing comorbidities from the weighted list to give the final index value for each participant. For example, an eighty year old with cerebrovascular disease, diabetes and hemiplegia would be assigned a value as follows:

Comorbidity index = 4+1+1+2 = 7.

5.3.4.4 Charnley classification

The Charnley classification (Charnley 1972) was designed to indicate the level of disability of a patient and consists of four categories describing the extent of lower limb problems, as follows:

- A – one hip involved
- B – two hips involved
- BB – two hips involved, one already replaced
- C – other factors affecting mobility

This simple system of classification is commonly used in arthroplasty studies and was included in this study to allow sub group analysis of the results according the extent of lower limb morbidity.

5.3.5 Participants

5.3.5.1 Eligibility and selection

Potential participants for the current study were identified from the previous study and their mortality status was ascertained. Each was invited by letter to attend a routine follow up appointment related to their joint replacement and details of the clinical study were included for their consideration; on attendance, their participation was discussed and informed consent obtained if they wished to proceed (see Appendix V). All potential participants whose age or comorbidities prevented their attendance at the clinics were contacted by telephone to establish if the hip replacement was problematic in any way or if any further surgery had been required.

5.3.5.2 Method of follow up

On attendance, each participant was asked to complete an OHS and an EQ-5D questionnaire, and to note any complications with the joint replacement. These questionnaires were part of the routine arthroplasty surveillance in the unit and were completed without assistance from the researcher. If any help was required to read or understand the questions, it was provided by an accompanying friend or member of the nursing staff in the department. The participant would then proceed for an x-ray of the study hip.

The final stage of the assessment was a face-to-face interview and it was at this stage that written informed consent was obtained. A history was taken, including current medical problems, and was followed by a brief physical assessment of the affected hip(s) and other areas

if indicated. The participant was given an explanation of the x-ray images and the condition of their hip replacement. If required, advice was offered on how the participant could best manage any related problems. The Charlson comorbidity index and Charnley classification were completed following the patient interview.

In a few cases, further tests were ordered to confirm or refute the findings of the examination prior to onward referral to an orthopaedic surgeon or to other services such as physiotherapy or podiatry. For example, one participant had developed muscle weakness in the leg and another had difficulty with walking due to some problems with the foot. All relevant details of the arthroplasty review were sent in a letter to the participant's general practitioner.

5.3.5.3 Ethical approval

All potential participants were contacted by letter in advance of the clinic allowing time for consideration of their participation in the research. This included a telephone number to enable them to contact the researcher and ask any questions that they may have relating to the study. Reassurance was given that their decision would not affect any further treatment. For those who chose not to participate, routine data would be collected in the follow up clinic but not used for the study. Informed consent was obtained on their attendance at the clinic and after addressing any questions from the participant about the study.

In this study, no participant was exposed to additional risk from radiation as the x-ray images obtained were part of routine follow up care. Ethical approvals were obtained from the North Somerset and South Bristol NHS Research Ethics Committee and from the University of the West of England Faculty Ethics Sub-Committee, School of Health and Social Care.

5.3.5.4 Regulatory requirements

The WGH Research and Development department approved the study prior to commencement (see Appendix IV). All participants attended one of the weekly arthroplasty clinics held in the Orthopaedic Department of WGH and the demographic information was recorded on routinely used case report forms although it was anonymised before statistical analysis. The arthroplasty clinics are subject to the Weston Area Health NHS Trust regulations and all health and safety

issues are addressed through regular training for staff which included the researcher who manages these clinics.

5.3.6 Radiographic assessment

The method for obtaining x-rays and conducting the radiographic assessment are described in the following paragraphs.

5.3.6.1 Production of x-ray images

Two x-rays were taken of each hip replacement - a standardised A-P view and an iliac (posterior) oblique view. The A-P projection was centred on the pubic symphysis (low centred) with a source to plate distance of 100cm (ASTM International 2004). The patient was placed supine on the x-ray table with internally rotated femora (referred to as a 'toes in' position) in order to better visualise the femoral component (see §4.3.3.3). The iliac oblique view was obtained with the patient in side lying with the affected hip on the table and the pelvis turned 45° anteriorly in order to better visualise the pelvic region behind the acetabular component (Thomas et al. 2007). For both views, specific requests were made for the radiographer to capture the full extent of the femoral stem in addition to the acetabular component.

Radiographic assessment of the THA images obtained at mid-term review was completed with digitised x-rays displayed on a high definition computer screen, as used by radiologists for reporting. The integral software provided options to re-size the images, magnify or highlight areas of interest, change the grey scale, measure angles and measure distances after calibration using a known length, for which the diameter of the femoral head was used. Assessment of the earlier plain films (taken three years after surgery) had been made using a standard light box, electronic callipers (Absolute Digimatic, Mitutoyo, Japan), a 360 degree geometric protractor and an additional light source when needed. The acetabular data were brought forwards and remeasured by the researcher (LKS). The femoral component data were collected at the completion of radiographic data collection in the present study. Potential bias was reduced by ensuring that there was an interval of two months or more after the mid-term review of each participant before collecting the femoral data from the earlier x-ray images.

5.3.6.2 Radiographic review process

All data collected as part of the radiographic review were recorded by the researcher (LKS) on a clinical record form specifically designed for arthroplasty review (see Appendix V). Initially, an overview was taken of the image quality and soft tissue to identify any problems, followed by a detailed examination of the prosthesis and surrounding bone. On the acetabular side, measurements were taken of the angle of inclination of the cup and linear wear of the polyethylene liner or cup. The acetabular cup was also assessed for signs of migration when compared with earlier films and the opposite hip.

On the femoral side, the inclination and subsidence of the component were noted as were any rounding or resorption (convex shaping or reduction in apparent density) of the calcar which is the medial femoral neck (Williams et al. 2002, Gonzalez Della Valle et al. 2004). Any excess bone which had developed in the soft tissues - heterotopic ossification - was graded using the system described by Brooker et al. (1973):

- Grade I: Islands of bone within the soft tissues about the hip
- Grade II: Bone spurs from the pelvis or the proximal end of the femur, leaving at least one centimetre between opposing bone surfaces
- Grade III: Bone spurs from the pelvis or the proximal end of the femur, reducing the space between opposing bone surfaces to less than one centimetre
- Grade IV: Apparent bone ankylosis of the hip

The detail of how the measurements were obtained is included in §5.3.7.

5.3.6.3 Radiographic zones

Changes around the components were recorded by zone. On the A-P view of the femur, the zones of Gruen et al (1979) were used which are numbered from one to seven (see Fig. 5.1a). The zones of Johnston et al (1990) were used for the oblique view, numbered from eight to fourteen. The acetabular zones I to III of DeLee and Charnley (1976) were used for the A-P acetabular view (see Fig, 5.1b) and zones IV (anterior) to VI (posterior) for the oblique view (DeLee et al. 1976, Pollard et al. 2006).

Figure 5.1. Antero-posterior radiographic zones around a total hip replacement (A = femoral zones, B = acetabular zones) (Used with permission of Orthoteers http://www.orthoteers.com)
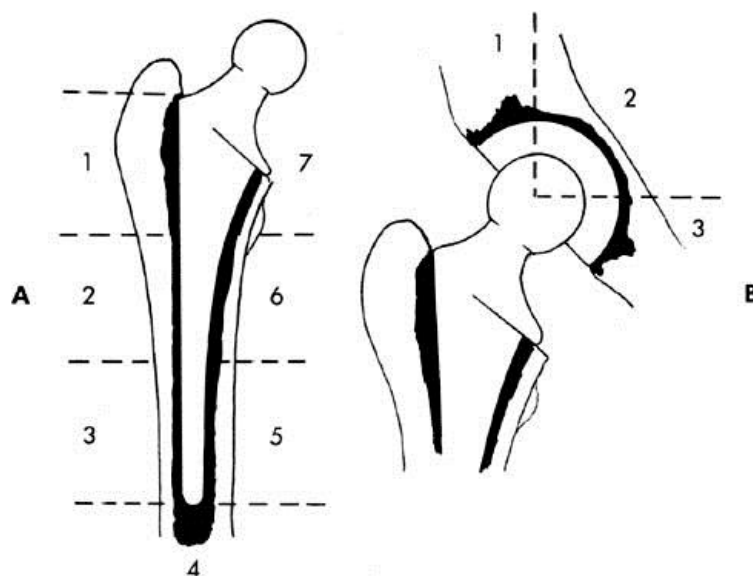
5.3.6.4 Radiographic changes

The changes noted on the x-ray as part of the radiographic assessment were as follows:

- Radiolucencies: a darkened lucent line, usually between the cement and the bone. If it was greater than 2mm in width, this was noted as osteolysis (Utting et al. 2008)

- Cortical hypertrophy: a thickening of an area of bone cortex evident when compared with adjacent cortex

- Osteolysis: a new or expanding radiolucent area (darkened area) adjacent to either the cup or the femur and in which no trabeculae were visible compared to the adjacent bone or prosthesis (Gonzalez Della Valle et al. 2004, Hernigou et al. 2009). Comparison was made with previous radiographs in order to exclude pre-existing cysts or osteopoenia.

- Granuloma: a new, expansile, darkened area less dense than surrounding bone in which some trabeculae were visible and commonly seen behind the screw holes of uncemented acetabular components

- Cement mantle deficiency: an area adjacent to a cemented prosthesis with cement width of less than 2mm width (Yates et al. 2008)

5.3.7 Measurement and data management

5.3.7.1 Oxford Hip Score

The OHS was completed by each participant at the mid-term review and scored from 0 to 4 for each of the twelve questions, giving a range of zero to 48 (best to worst). This method had been used in the earlier study of this same cohort. The current recommendation is to use a 4 to 0 format but as long as the method is stated, results are valid and comparable (Murray et al. 2007). The score from the three year review was subtracted from the score obtained at mid-term to produce an OHS change score which was independent of the method of scoring.

It has been recommended that when using a change score, it should be independent of the baseline data (Kaiser 1989). The OHS change score was visually assessed using two scatter plots − one of the absolute changes against the three year score (baseline data) and one of the percentage change against the three year score (baseline data). The object is to select the least dependent of the two by interpretation of the plots; independence is shown when points are scattered such that there is no obvious linear or curvilinear relationship.

An effect score is a measure of the magnitude and direction of change in an outcome tool and is useful to interpret the results obtained. A standardized effect size of the change in OHS was calculated using the following formula (Sim and Wright 2000):

$$\text{Standardized effect} = \frac{\text{Mean of OHS at mid term less mean of OHS at 3 years}}{\text{Standard Deviation of OHS change score}}$$

A size of 0.2 is considered small, 0.5 is a medium effect and 0.8 is a large effect size (Cohen 1988).

5.3.7.2 Age

The age of the patient was entered in years but in some of the supplementary analysis, the participants were grouped by decade in order to compare these sub-groups on a non-parametric variable. The decades were defined from 45-54, 55-64, 65-74, 75-84, and over 85 years which is comparable with the EQ-5D population norms (Kind et al. 1999). This allowed capture of all the

ages of the participants with sufficient numbers in each group for the statistical analysis (see Appendix VI, Figure A.8).

5.3.7.3 EQ-5D

The EQ-5D questionnaire produces five answers, each of which is scored from one to three and then assigned a value from a national set constructed by the EuroQol group. The sum of the five individual values less a constant gives the general health score for the participant at that time. The national sets were constructed from sample populations who were asked to relatively value different health states.

There are two value sets available for each country, one computed using the time trade off (TTO) method and one using a visual analogue scale (VAS) method (this should not be confused with the separate EQ-VAS score for perceived health used in some studies which is a mark on a vertical line). Both methods produce a score from 1.0, equivalent to full health, through zero (death) to -1.0, a state assumed to be worse than death. The TTO method required participants to rate ten years in a number of health states in comparison to full health and to death. The VAS method required participants to relatively rate health states on a VAS. There is no clear recommendation about which method to use in a given situation but in this study, the TTO method was chosen as it has been shown to be valid for hip revision surgery and is recommended by the National Institute for Health and Clinical Excellence in the UK (Szende et al. 2007).

5.3.7.4 Angle of inclination of acetabular component

The angle of inclination of the acetabular component was measured on the x-ray from the A-P view by calculating the angle between a line drawn through the base of both teardrops (the inter-teardrop line) and a line drawn across the longest ellipse of the cup opening (McCombe and Williams 2004). The teardrop is a reference point seen on x-ray at the base of the acetabulum on the ilio-ischial line.

Any potential migration of the acetabular component was assessed visually with reference to the teardrop. If there appeared to be a change in position on initial viewing, precise measurements were taken and compared with the opposite hip or previous radiographs. Vertical migration was calculated as the perpendicular distance from the centre of the femoral head to the

inter-teardrop line. Horizontal migration was calculated from the teardrop to a line perpendicular to the inter-teardrop line and running through the centre of the femoral head (Maloney et al. 1999). A change in distance of greater than 3mm in any direction was considered to represent migration (Geerdink et al. 2009).

5.3.7.5 Inclination of femoral component

The position of the prosthetic stem in relation to the femur was categorised as neutral, varus or valgus from the A-P x-ray view. The tip of a stem that is in varus is angled away from the mid-line of the body and a valgus stem is angled in the opposite direction (towards mid-line). The angle between the shaft of the component and the mid-line of the femur was measured and any angle greater than 4° was noted using a similar method to previous studies (Yates et al. 2008).

5.3.7.6 Subsidence of femur

Subsidence of the femoral component within the cement mantle was measured on the A-P x-ray view with reference to two lines perpendicular to the mid-line of the femur, one at the tip of the greater trochanter and one at the proximal point of the shoulder of the prosthesis (Lusty et al. 2007). Subsidence greater than 3mm was considered to be of significance (Kim et al. 2002, Williams et al. 2002).

5.3.7.7 Osteolysis

Measurements of osteolysis were made using the morphometric grid described in the previous chapter. There were no osteolytic lesions present on the plain films taken at three years post-operatively, only at mid-term review. The x-ray images (all digital at mid-term) were re-sized to life size with reference to the femoral head. The morphometric grid was then superimposed and the number of points within the area of the osteolytic lesion was recorded. Each lesion was classified as either expansive or linear; expansive lesions were scalloped and extending away from the prosthesis and linear lesions were more uniform in shape along the bone-implant interface (Paprosky et al. 2001).

     The researcher (LKS) obtained data on two occasions with an interval of two months between in order to reduce bias. These measurements were tested for reliability (see §5.3.8.4)

and the mean of the two was calculated for use in the statistical analysis. It was not known in advance how many osteolytic lesions would be present at mid-term and whether this would affect the variable constructed from the morphometric grid data.

The data obtained in this way was also used to calculate the sensitivity to change which provides a standardised index that can be used to compare tools measuring a similar phenomenon (see §5.3.8.7).

5.3.7.8 Radiographic changes

The number of radiographic changes present at mid-term review was constructed by summing the number of changes observed in each zone. For example, the development of radiolucencies in two zones around the acetabular cup plus osteolysis and radiolucency in one femoral zone would be valued as '4'. This was then compared with the number of changes observed at three years and the difference between the two was described as the total number of radiographic changes. It included the development or extension of radiolucencies, osteolysis, granulomata and cortical hypertrophy but not wear measurement which was treated as a separate variable.

The use of radiographic changes, instead of one-time observations, is thought to provide better information about the state of the joint (Hodgkinson et al. 1988). A similar approach was adopted by the Swedish Registry when they included changes in their assessment. They only evaluated radiolucencies, osteolysis and wear as these were viewed as the most important changes with clinical relevance (Malchau et al. 2005).

5.3.7.9 Wear rate

A measurement of linear wear of the acetabular cup or liner was made using the method of Dorr and Wan (1996) as it can be used with both uncemented and cemented components, on plain film radiographs as well as digitised x-rays and has been applied in other clinical settings (Dorr and Wan 1996, Barrack et al. 1997, Pollard et al. 2006). The integral software or digital callipers were used in conjunction with the standardised A-P pelvic x-ray to obtain measurements of the depth of the cup (cemented components) or cup and liner (uncemented) at the superior and inferior edges. Linear wear was calculated as half the difference between the two with all

measurements in millimetres. All measurements were corrected for magnification by using the known diameter of the femoral head for reference.

A steady state wear rate is calculated after initial 'bedding in' of components which occurs in the early post-operative stage, usually by two years (McCalden et al. 2005, Jacobs et al. 2007). As the early review of these participants had taken place three years post-operatively, the films obtained at that time allowed for completion of this process. The steady state wear rate (mm/yr) for each participant was calculated as follows (Hamilton et al. 2005):

$$\text{Steady state wear rate} = \frac{\text{Linear wear at mid-term review} - \text{linear wear at initial review}}{\text{Number of years between reviews}}$$

There were a number of patients who appeared to have negative rates of wear but this is a recognised phenomenon due to the small size of the linear measurements and differences between serial x-rays (McCalden et al. 2005). It is recommended that the number with negative wear is reported in the results.

5.3.8 Data analysis

This section describes the statistical tests conducted using the Statistical Package for Social Sciences version 15.0 (SPSS, Chicago, Illinois, USA). No attempt was made to include data for participants lost to follow up as any methods used to account for radiographic changes would be unreliable and inappropriate.

5.3.8.1 Tests of normality

The Kolmogorov-Smirnov statistical test of distribution was used to assess all outcome variables. A significant result implies non-normal distribution but the outcome of each test was considered with reference to the cumulative evidence as some tests for normality will appear to be negative when dealing with larger samples (Pallant 2007). It has been suggested that in such samples, data can be accepted as normally distributed based on the visual appearance because deviation due to skewness has little effect on the analysis (Tabachnick and Fidell 2001).

All of the variables to be entered into the analysis underwent preliminary univariate testing for normality. This included an examination of the descriptive data for each variable with comparison of mean and median values, trimmed mean, skewness and kurtosis. Visual assessment was made of the frequency plots (histograms), box plots and normality pots for shape and possible outliers in the data sets.

5.3.8.2 Transformation of data

In some cases, transformations of the variable were applied with subsequent re-checking of data for normality. The decision whether to include a transformed variable or to retain the original values was based on the shape and the possible effect of transformation on the coefficients in the final statistical analysis (Tabachnick and Fidell 2001).

5.3.8.3 Outliers

Where initial explorations identified outliers, raw data was re-examined to check for accuracy of data entry followed by a visual assessment of the corrected data. If this showed that the tails of the histogram were comparatively even in shape, the data were retained (Pallant 2007). In cases where the outliers were clearly separated from the tail of the histogram, the standardized score of the outlier was examined. For those with a standardized score of >3.29, a new value was assigned to the outlier which was one unit larger or smaller than the extreme of the existing distribution (Tabachnick and Fidell 2001).

5.3.8.4 Reliability tests

The ICC was used to test the reliability of radiographic measurements. Reliability within observers (test-retest reliability) was tested with a one way-random effects model (1, 2). Reliability between observers (inter-observer reliability) was tested using a two-way mixed effects absolute model (2, 1). In addition, Bland and Altman plots were produced in every case to visually represent the data and further assess the reliability. Further details on these methods can be found in Chapter 4.

5.3.8.5 Hierarchical regression analysis

Multiple regression analysis produces a coefficient of determination (R^2) which resolves how much of the variance in a dependent measure is explained by the independent variables. The particular method selected was hierarchical regression analysis as it is designed to explore the relationship between multiple independent variables and one dependent variable while controlling for covariates that may be affecting the model (Tabachnick and Fidell 2001). In this way, the type of hip replacement could be controlled before assessing the individual and cumulative effect of each of the other variables on the change in OHS score.

The correlations between each of the variables were examined for multicollinearity (high level of correlation) before entry into the hierarchical regression analysis. It is recommended that if two variables are highly correlated (correlation coefficient >0.90), one of them is deleted (Tabachnick and Fidell 2001).

5.3.8.6 Non-parametric tests

Further statistical exploration was conducted to identify sub-groups with evidence of greater change over time as measured by radiographic features. It was decided that, in the event of a small data set for osteolytic measurements, this data was to be treated as a dichotomous variable. Data were grouped by presence or absence of osteolysis and the type of THA.

The Kruskal-Wallis test was used to compare groups with non-parametric distributions and the Mann-Whitney U with a Bonferroni correction was used for post hoc tests.

5.3.8.7 Sensitivity to change

The measurement of sensitivity to change of the morphometric grid was calculated from the measurements of osteolytic lesions at mid-term review. None of the lesions were present at the three year review and this change was used as the basis for the calculation of an effect size. For parametric data, the following formula is one of those recommended (Sim and Wright 2000):

$$\text{Effect size} = \frac{\text{Mean of outcome at time 2} - \text{Mean of outcome at time 1}}{\text{Standard deviation of outcome measure at time 1}}$$

For non-parametric data, the Wilcoxon Signed Rank test is used to obtain a *z* value and then the following formula is applied (Pallant 2007):

Effect size = z / √N   (N = total number of observations)

The results from either method of calculation are evaluated with reference to a change in the mean or median when a size of 0.2 is considered small, 0.5 is a medium effect and 0.8 is a large effect (Sim and Wright 2000).

5.3.9 Study size

The sample size was calculated retrospectively using a recommended formula for multiple regression analysis (Tabachnick and Fidell 2001). The size is related to the number of independent variables to be entered, as follows:

Size of sample = 50 + (8 x number of independent variables)

This would require a minimum of 90 participants for a study of five variables (50 + 8x5). The final study sample was determined by the number of participants still available from the previous cohort. It exceeded the minimum number which is a strategy recommended if the anticipated effect size is small (Tabachnick and Fidell 2001).

5.3.10 Bias

The researcher was also the arthroplasty practitioner in this study which introduced potential bias in collection of outcomes and interpretation of x-rays. In recognition of this problem, patient completed outcome measures were finished before the participant was seen by the researcher. On the few occasions where a participant required help to complete questionnaires, verification of the written answers was clearly obtained from the participant.

The potential bias in image interpretation was addressed in a number of ways. All data that had been collected were reviewed at a later date by a senior orthopaedic surgeon in hip arthroplasty to obtain agreement on the changes that had taken place. In any instance of difference of opinion, the surgeon's results were retained. All area measurements of osteolysis were repeated at a two week interval and the data were tested for reliability. A sample (10%) of all wear measurements and angles of inclination were independently measured by an orthopaedic fellow (an experienced trainee) with an interest in radiographic measurement in order to assess the reliability of the data collected. The blinding of all x-rays prior to interpretation was another possible method of reducing bias introduced through knowledge of previous results, but the additional work involved to achieve this proved a barrier to implementation within the confines of this study. A pragmatic decision was made to include a time interval between assessments of x-rays to minimise this possible bias.

The dual role of the researcher had positive benefits as attention to detail was necessarily high which was valuable for participants. In addition, the potential outcomes of the study were continuously reviewed with respect to pragmatic application in a busy clinical setting.

5.4 Results

5.4.1 Participants

The flow of participants in this study is illustrated in Figure 5.2. Of the 13 people unable to attend due to age and/or comorbidities, none reported that they were experiencing any problems with their hip replacement and none had required any further surgery. All individuals who were contacted agreed to participate and provided written consent.

Figure 5.2. Flow chart showing the number of participants entered into the study
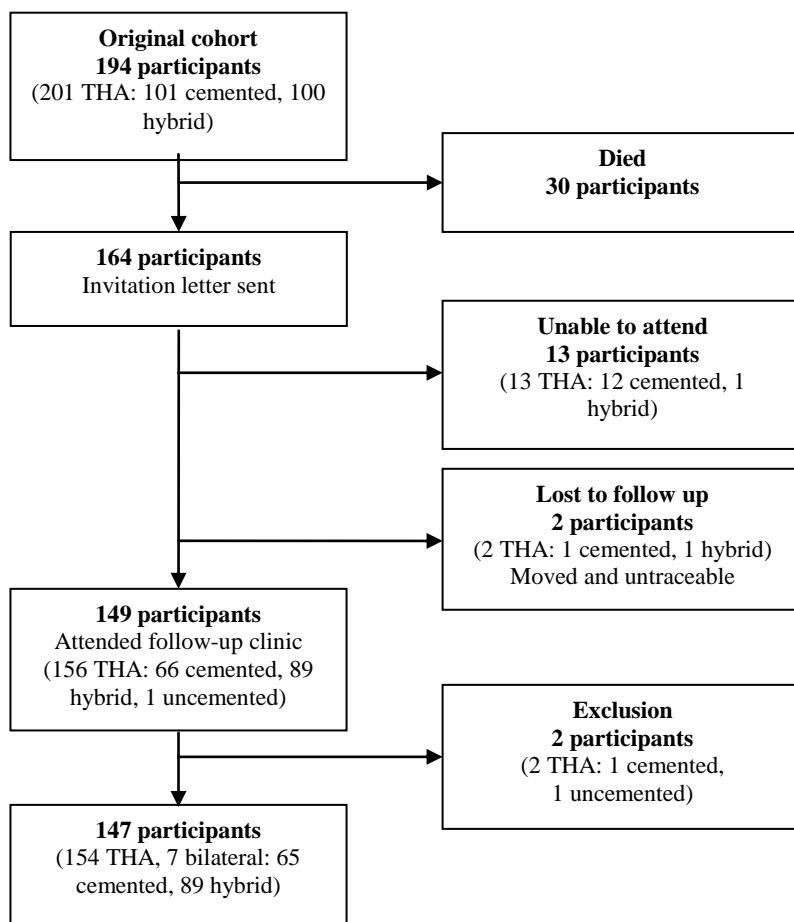
There were two participants excluded from the final analysis. One of these was due to the type of prosthesis which was an uncemented THA and was the only case of this type in the cohort. He was excluded as this sub group would be too small for analysis. The other participant was excluded due to missing x-ray films from the three year review which prevented comparative analysis.

5.4.2 Descriptive data

5.4.2.1 Demographic and diagnostic

The demographic and diagnostic data from the study cohort are shown in Table 5.2. The collection of data from the OHS was not affected by missing information as staff were available to help with any problems and all questionnaires were checked for completeness before the participant departed.

At three years, two of the hips had been revised for deep sepsis (one cemented, one hybrid) and three hybrid THA had required a polyethylene liner and femoral head exchange (two for repeated dislocation, one for soft tissue interposition). All five of these hips were excluded before commencing the current study and, at mid-term review, there were no further revisions for any reason. One case of deep sepsis was receiving conservative treatment with regular monitoring and was included in the study cohort.

Table 5.2. Demographic and diagnostic data of study cohort

Description		Total (%)	Cemented THA (%)	Hybrid THA (%)
Number of THA		154	65 (42)	89 (58)
Age in years: Mean		74.5	81.3	69.4
Range		42 to 94	64 to 94	42 to 90
Laterality:	Right	83 (54)		
	Left	71 (46)		
Gender:	Male	61 (40)	18 (28)	43 (48)
	Female	93 (60)	47 (72)	46 (52)
BMI		27.8	26.2	28.9
Range		18 to 47	19 to 40	18 to 47
Years since surgery		7.5	7.6	7.5
Range		6.2 to 9.1	6.4 to 8.5	6.2 to 9.1
Diagnosis:	Osteoarthritis	142 (92)	60 (92)	82 (92)
	Rheumatoid arthritis			1 (1)
	Septic arthritis		1 (2)	
	Post-traumatic arthritis			1 (1)
	Femoral neck fracture		3 (5)	3 (3)
	Perthes' disease			1 (1)
	Slipped upper femoral epiphysis			1 (1)
	Avascular necrosis		1 (2)	
Femoral component	Exeter (28mm head)	128 (83)	48 (74)	80 (90)
	CPS Plus (28mm head)	13 (8.5)	4 (6)	9 (10)
	Charnley (22.2mm head)	13 (8.5)	13 (20)	
OHS score at 7.5 year review (0 best to 48 worst): Mean		8.6	11.0	6.8
Range		0 to 36	0 to 36	0 to 28
OHS score at 3 year review: Mean		7.5	8.7	6.6
Range		0 to 31	0 to 31	0 to 29
EQ-5D score: Mean		0.76	0.69	0.80
Range		-0.07 to 1.0	-0.07 to 1.0	0.19 to 1.0
Charlson comorbidity score: Mean		3.6	4.5	3.0
Range		0 to 11	2 to 11	0 to 7

(Key: THA= Total hip arthroplasty, BMI = Body mass index, OHS = Oxford Hip Score, EQ-5D = EuroQol 5-dimension questionnaire)

5.4.2.2 Radiographic assessment

A summary of the descriptive data from the radiographic assessment is presented in Table 5.3.
The early review took place at a mean of 3.1 years after surgery.

Table 5.3. Descriptive data from radiographic assessment

Description		Total (154) (%)	Cemented THA (65) (%)	Hybrid THA (89) (%)
Acetabulum				
Inclination of acetabular cup in degrees: mean [SD], Range		46.7 [6.0] 30 to 65	46.6 [5.2] 35 to 60	46.8 [6.6] 30 to 65
No. of cups with radiolucencies at 3.1 yrs		51 (33)	51 (79)	0
No. of cups with radiolucencies at 7.5 yrs		54 (35)	54 (83)	0
No. of granuloma behind cup at 7.5 yrs		55 (36)	0	55 (62)
No. of cement mantle deficiencies at 7.5 yrs		23 (15)	23 (35)	0
No. of cups with changes at 7.5 yrs		95 (62)	40 (62)	55 (62)
Femur				
Alignment of stem of prosthesis	Neutral	100 (65)	45 (69)	55 (62)
	VR $< 4°$	31 (20)	11 (17)	20 (23)
	VL $< 4°$	19 (12)	8 (12)	11 (12)
	VR $\geq 4°$	3 (2)	1 (2)	2 (2)
	VL $\geq 4°$	1 (1)	0	1 (1)
Stem subsidence at 7.5 yrs	0 to 1.5 mm	31 (20)	14 (22)	17 (19)
	1.5 to 3.0 mm	123 (80)	51 (78)	72 (81)
Calcar	Rounding	49 (32)	19 (29)	30 (34)
	Resorption	78 (51)	39 (60)	39 (44)
No. of stems with radiolucencies at 3.1 yrs		62 (40)	27 (42)	35 (39)
No. of stems with radiolucencies at 7.5 yrs		146 (95)	62 (95)	84 (94)
No. of stems with cortical hypertrophy at 7.5 yrs		6 (4)	0	6 (7)
No. of cement mantle deficiencies at 7.5 yrs		24 (16)	11 (17)	13 (25)
No. of stems with changes at 7.5 yrs		141 (92)	61 (94)	80 (90)
Heterotopic bone formation (Brooker grades)	Grade 0	91 (59)	39 (60)	52 (58)
	I	49 (32)	20 (31)	29 (33)
	II	11 (7)	4 (6)	7 (8)
	III	3 (2)	2 (3)	1 (1)
	IV	0	0	0

(Key: SD = Standard deviation, VR = varus, VL = valgus)

There was no acetabular component migration. There were 42 cemented cups with radiolucencies in zone I at 3.1 years and only eight of these (12%) had progressed from zone I to zone II at the mid-term review. The cement mantle deficiencies seen around the acetabular cups were predominantly in zones I and II.

On the femoral side, cement mantle deficiencies were predominantly seen in Gruen zones one and eight. There was no participant with subsidence greater than 3.0mm (Williams et al. 2002). Three femoral stems (2%) were in varus greater than four degrees but none more than five degrees (Kim et al. 2002, Yates et al. 2008). There was one stem in seven degrees valgus which belonged to a participant with sequelae of a slipped upper femoral epiphysis including abnormality of the femoral anatomy.

Following the radiographic review, all participants x-rays were categorised based on the severity of the radiographic changes, as in §4.3.3.4. In addition to the three groups described previously, there was a fourth group in which there were no significant changes. As a result of this categorisation, 80% (120) of the participants had no changes and were scheduled for routine review in a further five years; 16% (25) had mild changes and were scheduled for earlier review in 3 to 4 years time; and 6% (9) had evidence of moderate changes and were given appointments for a review in 12 to 18 months. There were no patients requiring immediate referral pending revision surgery.

5.4.2.3 Osteolysis

There were 15 participants (10%) with osteolysis and details are shown in Table 5.4. One of those with a cemented THA had known bony metastases in the pelvis and one participant with a hybrid THA had a chronic infection with sequelae but was unfit to undergo revision surgery. The relatively small proportion of participants with osteolytic lesions prevented use of the data as an independent variable in the regression analysis as it was highly positively skewed.

Table 5.4. Descriptive data for osteolytic lesions

Description		Total (%)	Cemented THA (%)	Hybrid THA (%)
Number of participants		154	65	89
No. participants with osteolysis		15 (10)	3 (5)	12 (14)
No. participants with more than one lesion		4 (3)	0	4 (5)
No. of lesions	Total	25	3	22
	Femoral	19	2	17
	Pelvic	6	1	5
Area of lesions (no. of points on morphometric grid)	Median	7	8	7
	Range	2 to 34	5 to 11	2 to 34
No. of lesions greater than MDC_{95}	4 points	19	3	16
	11 points	9	0	9
Shape of lesion:	Expansile	21	3	18
	Linear	4	0	4

5.4.3 Reliability and sensitivity

5.4.3.1 Angle of inclination

The ICC for interobserver reliability between two observers each measuring the angle of inclination on a sample of 20 x-rays was 0.97 (95% CI: 0.93 to 0.99).

5.4.3.2 Wear

The ICC for interobserver reliability between two observers with a randomly selected sample of 20 x-rays was 0.87 (95% CI: 0.69 to 0.95). A Bland-Altman plot was used to visualise the data and showed acceptable agreement (see Appendix VI, Figure A.3).

5.4.3.3 Osteolysis

The test-retest reliability of the measurements of osteolysis had an ICC value of 0.91 (95% CI: 0.80 to 0.96). A Bland-Altman plot was constructed to represent the agreement graphically (see Appendix VI, Figure A.4).

5.4.3.4 Sensitivity to change

The measurements of osteolysis obtained from 15 patients (25 lesions) were evaluated with a Wilcoxon Signed Rank test and showed a statistically significant change from non-existence at three years to the mid-term review, $z = -4.376$, $p < 0.001$ with a moderate effect size ($r = 0.62$). The median size of lesion measured with the morphometric grid at mid-term was 7 points.

5.4.4 Analysis of variables

5.4.4.1 Change in Oxford Hip Score

When the OHS change score against baseline (the three year OHS) was compared with percentage change of OHS against baseline, no dependency was noted in either so the actual change score was retained as the variable to enter the regression analysis (see Appendix VI, Figures A.5 & A.6).

The Kolmogorov-Smirnov test for the OHS change score indicated non-normality but an inspection of the histogram showed that it was approximately normal with some outliers (see Appendix VI, Figure A.7). Consequently, a decision was made to treat the data as normally distributed and the outliers were retained following closer examination of the descriptive data of this variable (see Appendix VI).

The Mann-Whitney U test was used to explore the difference between the change in OHS in the two groups of THA, cemented and hybrid. There was no statistically significant difference between the groups, $U = 2526$, $z = -1.35$, $p = 0.178$.

The standardized effect size of the change in OHS was calculated as 0.2. The difference between the scores obtained at three years and those at mid-term review had a mean value of 1.07 (95% CI -0.01 to 2.15) using the paired sample t test ($t = 1.963$, $df = 153$, $p = 0.05$).

5.4.4.2 Age

The data on age of participant satisfied the requirements for normal distribution but there was one outlier at age 42 years ($z = -3.47$). It was assigned a value at the lower extreme of the distribution and re-entered at 47 years (see Appendix VI, Figure A.8).

5.4.4.3 EQ-5D

The distribution of the EQ-5D scores was negatively skewed (see Appendix VI, Figure A.9). A number of transformations were attempted but there was no significant improvement over the raw data. In view of the large sample size and the effect of transformation on interpretation of coefficients in regression analysis, the untransformed EQ-5D score were retained.

The EQ-5D scores were compared by age with the value set for the South West (SW) of England (Kind et al. 1999). In the younger participants, the values were comparable but, in the higher age groups which constituted the majority (86%) of the study participants, the general health was poorer than population norms (Table 5.5). The overall median score was similar to a patient cohort of THA reviewed at six years in another study (Malchau et al. 2005).

Table 5.5. EQ-5D scores by age and SW England population norms (Kind et al. 1999)

Age	Study results (SD)	South West Region England norms (SD)
Age 45-54	0.88 (0.15)	0.90 (0.15)
Age 55-64	0.87 (0.19)	0.80 (0.27)
Age 65-74	0.80 (0.20)	0.90 (0.17)
Age 75+	0.70 (0.24)	0.82 (0.17)

5.4.4.4 Charlson comorbidity index

The distribution of the Charlson comorbidity data satisfied the requirements for normality apart from one outlier, a participant with bony metastases from a solid tumour who scored 11. This score was replaced with a value of eight to place it at the extreme of the range of the existing data which was from zero to seven (see Appendix VI, Figure A.10).

5.4.4.5 Radiographic changes

The variable created to represent the number of radiographic changes was normally distributed (see Appendix VI, Figure A.11).

5.4.4.6 Wear rate

The data for steady state wear rate (mm/yr) included 12 negative values (-0.04 to -0.01) and five zero values which were replaced with a negligible value of 0.01 for handling in the data analysis. The variable was positively skewed (see Appendix VI, Figure A.12).

5.4.4.7 Summary statistics for test variables

The summary statistics of the test variables are shown in Table 5.6.

Table 5.6. Summary statistics for test variables of 154 THA

Variable	Mean (SD)	95% CI Mean	Median	Range
OHS change score	1.07 (6.77)	-0.01 to 2.15	0	-13 to 23
Age	74.5 (9.23)	73 to 75.9	75	47 to 94
EQ-5D score	0.76 (0.23)	0.79 to 0.87	0.74	-0.07 to 1.0
Charlson comorbidity	3.61 (1.47)	3.38 to 3.85	3	0 to 8
No. of radiographic changes	3.71 (2.29)	3.34 to 4.07	4	0 to 14
Wear rate cemented THA (mm/yr)	0.07 (0.07)	0.05 to 0.09	0.06	0.01 to 0.26
Wear rate hybrid THA (mm/yr)	0.12 (0.11)	0.10 to 0.14	0.10	0.01 to 0.57

(Key: SD = Standard deviation, CI = Confidence interval, OHS = Oxford Hip Score, EQ-5D = EuroQol 5-dimension questionnaire, THA = Total hip arthroplasty)

5.4.5 Multiple regression analysis

5.4.5.1 Univariate correlations

The correlations between variables entered into the multiple regression analysis are shown in Table 5.7. Correlations with the radiographic changes (dependent variable) were very low in all cases. Significant but low correlations were seen between the EQ-5D score with age and with Charlson comorbidity, and a significant moderate correlation with the OHS change score. There was a moderate strength correlation between age and Charlson comorbidity due to the inclusion of age in the construction of the Charlson index. However, there were no instances of high

multicollinearity (correlation coefficient >0.90) and consequently, all variables were retained for the regression analysis (see §5.3.8.5).

Table 5.7. Univariate correlations of test variables

	1	2	3	4	5
1 Radiographic changes	-				
2 OHS change score	-0.07	-			
3 Age	0.13*	0.10	-		
4 EQ-5D score	0.03	-0.39**	-0.27**	-	
5 Charlson comorbidity	0.02	0.09	0.75**	-0.28**	-

*P=0.05   **P<0.001

5.4.5.2 Hierarchical multiple regression

Table 5.8 shows the final regression model with radiographic changes as the dependent variable. In Step 1, the type of THA was entered as a dichotomous variable and explained 2% of the radiographic changes but this was not statistically significant (*P =0.06).*

In Step 2, the four variables of interest were entered after controlling for the type of THA and explained an additional 3% of the changes but without statistical significance (*P =0.41).* Examination of the ß coefficients showed that none of the variables contributed significantly to the model.

Table 5.8. Results of regression analysis

Step and variable	R^2	R^2 *change*	*F* change	Standardised *ß*	*t*
1 Type of THR	0.02	0.02	3.50		
2 Predictive variables	0.05	0.03	1.01		
OHS change score				-0.07	-0.84
Age				0.21	1.47
EQ-5D score				0.04	0.42
Charlson comorbidity				-0.18	-1.43

5.4.6 Sub-group analysis

5.4.6.1 Age and radiographic changes

As the correlation between age and radiographic changes achieved statistical significance of $p = 0.05$, the Kruskal-Wallis test was used to further explore the effect of age using five groups defined as shown (see Table 5.9). The test showed no statistically significant difference between the groups for the number of zones with changes, X^2 (4, n=154) = 4.6, $p = 0.33$. This is illustrated in a box plot of the results (see Appendix VI, Figure A.13).

Table 5.9. Group sizes by age

Group	Age in years	Number of participants
1	45 to 54	4
2	55 to 64	18
3	65 to 74	50
4	75 to 84	63
5	85 to 99	19

5.4.6.2 Age and wear rate

The Kruskal-Wallis test was also used to explore the difference in wear rate between the same five age groups. There was no statistically significant difference between the groups, X^2 (4, n=154) = 4.0, $p = 0.41$.

5.4.6.3 Charnley score and radiographic changes

The Kruskal-Wallis test was used to explore the difference between the four Charnley groups and the radiographic changes but there was no significant difference: X^2 (3, n=154) = 1.04, $p = 0.79$.

5.4.6.4 Osteolysis and wear rate

The Kruskal-Wallis test was used to explore the difference in steady state wear rate between sub-groups formed from prosthesis type and presence or absence of osteolysis (see Table 5.9). There was a statistically significant difference in steady state wear rate between the four groups, X^2 (3,

n=154) = 15.72, p <0.001. The median values of steady state wear for each group are shown in Table 5.10.

Post hoc analysis with a Mann-Whitney U test and Bonferroni correction showed a significant difference between the hybrid groups ($U = 245$, $z = -2.6$, p <0.025) with an effect size of $r = 0.3$ ($r = z$ / square root of N where N = number of cases). The difference between the cemented THA groups was not statistically significant at the adjusted p level ($U = 30$, $z = -2.0$, $p = 0.05$).

Table 5.10. Median wear rate for groups with and without osteolysis

Group	Wear rate in mm/yr (median)
Cemented prosthesis with osteolysis (n = 3)	0.11
Cemented prosthesis without osteolysis (n = 62)	0.05
Hybrid prosthesis with osteolysis (n=12)	0.16
Hybrid without osteolysis (n = 77)	0.08

5.4.6.5 Osteolysis and OHS change score

The Kruskal-Wallis test was used to explore associations between the groups of participants with and without osteolysis and the change in OHS score. There was no statistically significant difference between the four groups, X^2 (3, n=154) = 2.02, $p = 0.57$.

5.5 Discussion

5.5.1 Summary

The aim of this study was to explore associations between changes in the OHS over time and radiographic changes during the same period. In addition, the data generated were explored to identify any sub-groups of participants who were at greater risk of degenerative change in the hip replacement.

There were 147 participants (154 THA) available for review at a mean of seven and a half years after the primary surgery from a cohort of 194 participants (201 THA) who had previously been reviewed three years post-operatively. Additional data were gathered on the co-morbidities and general health of this cohort to inform the analysis of the changes.

A hierarchical multiple regression analysis was used to explore the ability of four measures to predict radiographic changes after controlling for the type of total hip replacement. The four measures were:

- OHS change score
- Age of participant
- EuroQol general health index
- Charlson comorbidity index

The regression model summary showed that none of the entered variables were able to significantly predict the radiographic changes at mid-term review. The type of THA accounted for 2% of the radiographic changes and the other four variables for a further 3%, but these were not statistically significant.

Further analysis showed that age was not significantly associated with the number of radiographic changes or with the steady state wear rate of the polyethylene component. The wear rate was found to be significantly higher in those participants with osteolysis compared to those without it, but this difference was not reflected in a significant difference between the OHS change scores of these groups.

5.5.2 Limitations

5.5.2.1 Role of researcher and practitioner

The potential bias in image interpretation introduced by the dual role of the researcher in collecting and analysing data had been recognised and so, reliability tests of data were undertaken. The values of the intraclass correlation coefficients for reliability were all acceptable, indicating that no undue bias was introduced (Streiner et al. 2003). This was confirmed visually by the Bland-Altman plots showing agreement between observers. Blinding of all the x-rays may have improved the study (see §5.3.10) but the results for reliability provide some reassurance of the consistency of the data produced.

5.5.2.2 Wear rate

Measurement of wear rate is not a standardised procedure and the method should be clearly stated. The method used in this study provides a relatively simple way of comparing sequential x-rays and different types of acetabular component (Dorr et al. 1996). However, it does not measure the area of maximum penetration directly as in the Livermore technique (Livermore et al. 1990) and interpretation of results and comparison with other studies must be with attention to the methods employed.

5.5.2.3 EuroQol score

The EuroQol score used in this study was obtained at one point in time whereas the OHS was a change score. It is possible that a change score for the EQ-5D questionnaire would have produced a different regression coefficient. However, the EuroQol is designed to be used as a single index value at a given point, as in this study, and the comparison with established values sets and other studies provided a method of validation to support the data included in the analysis.

5.5.2.4 Radiographic images

The x-rays taken at three years were printed on plain film whereas those at mid-term review were digitally recorded. Differences in the viewing medium and tools might have affected the results

obtained. This was minimised by the use on plain films of electronic callipers for linear measurement, supplementary light sources to view areas of osteolysis and radiolucency and re-sizing of all measurements using the known diameter of the spherical femoral head as the reference point.

In addition, the researcher reviewed all films obtained at three years after the mid-term review had been completed. This provided two forms of validation; first, the original reviewer was a surgeon and so a second observer validated the information obtained. Secondly, it provided the opportunity to re-measure any data about which there was uncertainty.

5.5.2.5 Radiographic changes

The radiographic changes were represented by the total number of changes seen on the x-rays. The number of changes present at mid-term was compared with the number at three years: the numerical difference between the two was calculated and entered as the total radiographic changes for an individual patient. Included in this would be any development of radiolucency, prolongation of radiolucency, cortical hypertrophy, breakage of components and development of granuloma or osteolysis (Johnston et al. 1990, Duffy et al. 2005). This is a simplistic representation but captures important changes and is easily calculated and verified by any member of the orthopaedic team. More detailed methods of assessment used elsewhere, such as minor changes in the width of a radiolucency, are difficult to measure reliably and may not be as important as the change in length (Hodgkinson et al. 1988). Any major change in the width of a radiolucency in this study would have been recorded as 'osteolysis'. Similar simplified systems have been used elsewhere (Geerdink et al. 2009) and the Swedish Hip Registry trialled a concise radiographic assessment which was specifically designed to capture clinically important changes without the need for complicated measurement techniques (Malchau et al. 2005).

In this cohort, there was no migration of acetabular component or subsidence of the femoral component outside accepted parameters and no significant change in alignment of the femoral stem (Utting et al. 2008, Geerdink et al. 2009). If such a change was observed in future studies, it could easily be added to the number of radiographic changes. By using a method which records the total number of adverse changes instead of subdividing into precise changes, there is flexibility to include any change relevant to the type of hip replacement that is being evaluated. It

may be necessary to conduct future research to explore the need for a weighted system, but importantly, the present model captures any progression of changes seen on x-ray which are an essential component of hip arthroplasty review (Hodgkinson et al. 1988, Wroblewski et al. 2002).

5.5.2.6 Osteolysis

The measurement of osteolysis using the morphometric grid was limited in this study to a few patients, none of whom had evidence of osteolysis at three years after surgery. Consequently, the amount of data collected was not sufficiently large to include in the multiple regression analysis which limited exploration of data produced from using the morphometric grid. An estimate of the sensitivity to change was calculated from the available data with a moderate effect size (r =0.62) which supports the use of the tool to measure these lesions. However, the sensitivity to change was dependent on both the sample and the tool. The sample in this study was a group of patients with osteolytic lesions evident at mid-term but not at the three year review. In a clinical situation, the observer is more likely to be interested in a change in size after a lesion had been identified and the first measurement of such a lesion would form the baseline for sensitivity to change rather than the non-existence of the lesion. Further work is needed to establish the sensitivity to change from such a sample and would provide evidence of the responsiveness of the tool to identify clinically important change (Streiner and Norman 2003).

In this sample of 15 patients with osteolytic lesions, the proportion of lesions greater than the MDC_{95} of four points (see §4.4.4.1) was 76% (19 out of 25), suggesting confidence in the measurement obtained. If the MDC_{95} of 11 points was used, an observer could be 95% confident that 36% of the lesions (9 out of 25) exhibited a definite change. The difference in these proportions illustrates the need for further work to assess the application of an MDC_{95} calculated from a range of observers and its application to a representative sample of patients with known osteolytic lesions that had changed in size over time.

5.5.3 Evaluation of participants

There were 194 patients (201 THA) in the original study at three years post-operatively and 149 patients (156 THA) were available for mid-term review at a mean of seven and a half years. The final cohort of 147 participants (154 THA), following exclusions, included 89 hybrid THA and 65 all cemented THA.

This cohort, with a mean age of 75 years and 58% hybrid THA, represents a typical patient group which may present for mid-term review in a district general hospital. The six orthopaedic surgeons under whose care these participants received the joint replacement represent a mix of skills with only one being a specialist hip surgeon. As such, this sample provides a realistic representation of patient groups in non-specialist centres across the UK at this time (National Joint Registry 2010).

No attempt was made to predict data for participants lost to follow up as radiographic changes would be difficult to estimate. The loss to follow up consisted of 30 deaths, 13 participants unable to attend due to age and infirmity and two who were untraceable. All those who were unable to attend were contacted and were satisfied with the hip replacement, had not required further treatment and did not report any ongoing difficulties with pain or function as a direct result of the arthroplasty. The two participants who were untraceable (one of each type of THA) had left the area with no forwarding addresses. In the worst case scenario, they may have gone elsewhere for further treatment but neither of these participants reported problems or dissatisfaction at the three year assessment.

5.5.4 Evaluation of Oxford Hip Scores

The variable constructed from the OHS used the actual change in score rather than the percentage change as the requirements for independence from baseline (the three year OHS) were satisfied by the preliminary checks (Kaiser 1989). This means that this variable was representative of the individuals' symptomatic changes over the intervening period of time and eliminates the bias introduced by their interpretation of the questionnaire (Murray et al. 2007). It does not account directly for any 'response shift' that may have occurred with age or changes in morbidity. However, by including an EQ-5D score and the Charlson comorbidity index, it was possible,

when necessary, to explain the effect of other conditions on the change in OHS. For example, one eighty year old lady had an OHS change score of 22 points which was at the upper end of the scale. However, she also had an EQ-5D of -0.07 and a Charlson score of 6 points. This indicated a very low level of general health and co-existing illness. When investigated further, she was found to have a heart condition, high blood pressure, diabetes, chronic kidney disease, gout and osteoarthritis in one of her knees. These additional problems may have affected her level of pain and functional activities so that she was unable to differentiate the cause of specific problems when answering the OHS questions.

The slight increase in mean value of the OHS from the three year review to the mid-term review is of interest. The OHS was originally designed to measure the change from 'before' to 'after' a total hip replacement when there is a major change in pain and function and effect sizes are large. However, it has been recognised that the OHS will also be used to measure change in cohort studies and long term audit (Dawson et al. 1996a, Murray et al. 2007). The differences between the scores obtained at three years and those at mid-term was just statistically significant for this cohort ($p = 0.05$). However, the statistical significance may be of less importance than the magnitude of the change as measured by the standardized effect size (Neill 2008). The effect size was 0.2 which, although small (Cohen 1988), was unsurprising as it had been anticipated that there would be some changes in the OHS but that these would not be large as the mid-term review took place soon after the static period of the OHS (Pynsent et al. 2005). Consequently, the effect size was encouraging, indicating that the OHS is responsive enough to detect changes in pain and function when used at the mid-term review.

The age of the participants was not correlated with the change in OHS although the cemented THA group (the group with higher mean age) recorded a higher mean OHS. This same tendency has been noted elsewhere - older patients or those with greater musculoskeletal disability have a tendency to record a higher OHS (Field et al. 2005). However, the present study used a *change* score rather than a one time score for the analysis. Consequently, the effect of the higher scores was reduced as it was present at three years and at mid-term review. The result was that the age of the participants was not related to the change in OHS.

5.5.5 Evaluation of regression results

The regression model allowed exploration of variables that might predict the radiographic changes. The results suggest that radiographic changes at mid-term review cannot be predicted by an OHS change score even when controlling for the type of hip replacement. There was no correlation between these two variables and no significant contribution from any of the other variables in the regression model. In a study by Malchau et al (2005), radiographic changes at mid-term were not associated with pain, satisfaction or EQ-5D. However, no specific, validated hip score was used in that study and so the present study adds to the existing information.

In the preliminary analysis of the study variables, a significant correlation of moderate strength was shown between the EQ-5D and the OHS changes score ($r = -0.39$, $p < 0.001$) (Cohen 1998). This negative correlation signifies that participants with a good general health score generally experienced a smaller change in their OHS. This result suggests that, although the OHS is a useful hip specific tool to indicate the patient's view of pain and function related to the THA, the addition of a general health index may allow better interpretation of the score obtained. This same suggestion has been described by other authors (Ostendorf et al. 2004, Garbuz et al. 2006). By way of contrast, the EQ-5D was only weakly correlated with age and comorbidity, indicating that the participant's view of their general health may have been affected by other factors. This study did not include any psychosocial analysis but it has been suggested elsewhere that psychosocial function does not significantly influence the outcome in THA (Learmonth and Cavendish 2005). Further exploration of these factors would be necessary to understand the implications.

The changes around a hip replacement as seen on x-ray are an important indication of its stability and endurance, and have traditionally formed the basis for clinical decisions made by orthopaedic surgeons. The OHS measures pain and function related to the hip replacement from the patient's perspective but the results of this analysis suggest that it cannot be used as an indicator of the underlying condition of the hip joint as seen on x-ray at mid-term review. The use of the OHS as a supplementary measure, rather than as a replacement for conventional methods of assessment, is stated as the main purpose of the questionnaire by the authors (Dawson et al. 1996a) and is reinforced by these results. The lack of association between the OHS and the

radiographic changes is important information to add to discussion about hip arthroplasty review in the current health environment. In particular, it contradicts the expressed view that a THA can be assessed solely with an outcome score (Price 2010).

5.5.6 Evaluation of sub-groups

The results of the radiographic review (see Table 5.3) showed that the proportion of participants with radiographic changes was similar in both types of THA. From the correlation values obtained in the regression analysis, it was clear that age was not strongly correlated with the number of radiographic changes and this was reinforced by the sub-group analysis. Some studies have suggested that younger patients experience more changes or a higher wear rate (Han et al. 1999, Maloney et al. 1999). It may be that the type of radiographic change which is being recorded will affect the results obtained. In the present study, all adverse changes were included in a simple numerical summary instead of specific criteria. Similarly, in the present study, no association was found between age and the steady state wear rate. The study by Han et al (1999) was based on a gross wear rate and no attempt was made to calculate a steady state wear rate, which may explain the difference in results. The results from the present study suggest that the patient groups for assessment cannot be selected on the basis of age. However, when analysing the results obtained, a patient's age in conjunction with their general health should be considered in relation to the length of time that the prosthesis must remain functional.

Further analysis was conducted on the data to ascertain if the commonly used Charnley classification was associated with patients at higher risk of developing adverse radiographic changes. It was found that it was not associated with increased changes despite the fact that it is focussed on the lower limbs. This suggests that the Charnley grouping cannot be used to identify those patients with greater radiographic changes at mid-term review.

5.5.7 Wear rate and osteolysis

In this study, the steady state wear rate data was positively skewed and was not included in the hierarchical regression analysis. The average value (median) for the cemented THA group was 0.06mm/yr and for the hybrid THA group was 0.10mm/yr. These rates and skew of the data are

comparable with other studies (Hamilton et al. 2005, Geerdink et al. 2009). In a study by McCombe and Williams (2004), using the Livermore technique for an eight year follow up, a wear rate of 0.07mm/yr was recorded in cemented cups and 0.15mm/yr in uncemented cups (all Exeter femoral stems) and Livermore et al (1990) recorded an average rate of 0.08mm/yr in a cemented THA with a 28mm femoral head.

There were 15 (10%) participants in this study with evidence of osteolysis in one or more locations, 12 of whom had a hybrid THA. However, one participant had known neoplastic metastases in the pelvis and one patient had a chronic infection of the THA with sequelae but was unfit to undergo revision surgery. Of the remaining 13 participants (8%), all but one were under 70 years of age at the time of primary surgery (average age 61 years, range 34 to 72 years). Patients in this age group are usually living independently, involved in activities of daily living and possibly sports, and this may increase their chances of developing osteolysis. This has been suggested due to a higher level of activity and subsequently higher wear rate resulting in osteolysis (Maloney et al. 1999). The one patient who was over 70 years at time of operation had chronic foot problems which significantly affected his gait pattern although he remained mobile. He had a high wear rate of polyethylene (0.49mm/yr), the cause for which was unknown, but his abnormal gait pattern may have altered the biomechanics of the THA. The overall percentage of those with osteolysis in the hybrid group (12%) was similar to other studies (Kim et al. 2002, Kitamura et al. 2006, Utting et al. 2008).

Further analysis was conducted to explore the link between wear rate and osteolysis in this cohort in view of previous studies which have suggested that a higher wear rate leads to more osteolysis (Amstutz et al. 1992, Hozack et al. 1996, Dumbleton et al. 2002). A comparison of participants with and without osteolysis (sorted by type of THA) showed a statistically significant difference in wear rate. Those in the hybrid group with osteolysis had a median wear rate of 0.16mm/yr compared to a rate of 0.08mm/yr in those without. The magnitude of this was of a moderate size ($r = 0.3$) and the results seem to concur with the previous studies. The difference between the cemented THA groups with and without osteolysis did not achieve statistical significance although the values were 0.11mm/yr for those with osteolysis and 0.05mm/yr for those without. This may have been due to the adjustment of level of significance made for multiple analyses or due to the sample size.

The rate of wear and its relationship to presence of osteolysis in the present study was directly comparable with the literature review completed by Dumbleton et al (2002) and shown in other studies (Barrack et al. 1997, Orishimo et al. 2003, Gonzalez Della Valle et al. 2004, Geerdink et al. 2009). Dumbleton et al (2002) concluded that rates of wear greater than 0.1mm/yr are most likely to produce osteolysis and rates less than 0.05mm/yr are unlikely to produce any osteolysis. However, a recent study has questioned the value of such wear thresholds and suggested that the association between wear rate and osteolysis might be a dose-response relationship (Emms et al. 2010). The implication of this is that even patients with a low wear rate may develop osteolysis over time in response to the volume of wear particles produced. It may be that the study cohort will exhibit this tendency beyond the ten year post-operative period. The newer polyethylenes, such as highly cross-linked or those with exposure to vitamin E (α-tocopherol), have been developed to reduce wear but as yet, long term results are not available. Mid term results suggest that highly cross-linked polyethylene may substantially reduce the wear rate (McCalden et al. 2009, Huo et al. 2010) but there are concerns about the biological effect of sub-micron particles generated (Holt et al. 2007, Calvert et al. 2009, Mu et al. 2009). In view of these changes, recommendations for continued monitoring of hip arthroplasty were included in all these scientific papers.

A comparison of changes in the OHS between groups with and without osteolysis found that there was no significant difference. This suggests that the OHS score alone is not sufficient to identify the presence of osteolysis at mid-term review and reinforces the silent nature of it as the participants were not signifying associated changes in pain or function through the OHS. It also reinforces the need for an x-ray to monitor the state of the hip replacement at mid-term in order to identify these potentially important changes.

The data on the area of osteolysis, as measured with the morphometric grid, were not entered into the regression analysis due to the small proportion of participants with such lesions at this mid-term review. However, the presence of osteolysis was included in the variable constructed to record radiographic changes as it had developed since the three year review in every case. In longer term reviews, it would be necessary to include changes in the area of osteolytic lesions as part of the radiographic tool. The minimal detectable change of the morphometric grid has been shown to be 11 points (maximum) with the participant sample in this

study. This information plus additional information regarding sensitivity to change would need to be considered when constructing a radiographic tool for long term THA assessment.

The positive association of increased wear rate with the presence of osteolysis in this study and others (Amstutz et al. 1992, Hozack et al. 1996, Dumbleton et al. 2002) indicates that the measurement of wear rate may be an important part of the assessment. As discussed earlier, there is no simple standardised method of measurement of wear rate and developing a simple technique for use in a busy clinical setting is a challenge. However, it may be that the simpler approach is to monitor the area of osteolysis from an early stage, when it first becomes apparent, and the morphometric grid provides a tool for this purpose.

5.5.8 Generalisability

This research study has employed an exploratory approach to look for associations in the data. The cohort of participants from whom data were obtained was representative of two types of hip replacement commonly used in the UK in the last ten years although other combinations exist. The primary surgical procedures were all completed in one hospital by a range of surgeons typical of those seen in non-specialist centres and the sample size was large enough to satisfy all statistical requirements. The EQ-5D scores for the cohort were comparable with UK norms and other THA patients (Malchau et al. 2005); the average value of the OHS at 6-9 years for participants with a hybrid THA was comparable with a similar local cohort (median 4.0, average age 69.4 years) (Pollard et al. 2006), indicating that the participants were a representative group.

The two types of hip replacement included in this cohort reflected common practice across the UK at the time of the primary surgery. The difference in average age of each group confirmed the tendency to select younger patients for hybrid THA but this difference was allowed for in the statistical analysis by the hierarchical regression process. The results from this study have external validity when applied to the mid-term review process for cemented or hybrid THA patients in a district general hospital setting in the UK.

5.6　Conclusion

This study has shown that the radiographic changes around a THA at mid-term review are not associated with the changes in a joint specific PROM. The radiographic changes are an important indication of the state of the prosthesis and were not reflected in the score from the outcome measure, even though it was joint specific and sensitive to change.

The results of this study of both cemented and hybrid total hip replacements identified that those participants with a higher steady state wear rate were more susceptible to the development of osteolysis but did not identify any other group with significantly more change on the x-ray.

The measurement of radiographic changes by number of zones is a simple method of summarising important information which captures any type of change adjacent to the component and the development of any new expansile lesions. However, the measurement of pre-existing osteolytic lesions requires additional information to describe any change in area that has taken place. The wear rate may be used as an indication of susceptibility to the development of osteolysis but is not always possible to measure, especially in a routine clinical setting. This reinforces the need for a simple tool, such as the morphometric grid previously described, to measure the area of osteolysis in addition to recording change by radiographic zones.

Chapter 6

General discussion

Each of the three preceding chapters has described an aspect of hip arthroplasty review set against the background of current orthopaedic literature. This chapter will bring together the principal findings and will discuss the implications and identify areas for future research.

6.1 Principal findings

The review of current orthopaedic literature supported the need for long term surveillance of hip arthroplasty in order to identify patients with a failing THA. A hip replacement that has functioned well for the early post-operative years is likely to continue to provide the patient with relatively pain-free function for some time but that period cannot be precisely defined. The aseptic processes of failure are often silent until substantial damage has occurred and without periodic review, the patient will be unaware of the changes. The benefit to the patient of timely revision is substantial - an improved chance of a good outcome and a reduction in the risks involved. It is also of benefit to the surgeon if the procedure is not complicated by extensive bone loss which requires significant reconstruction.

Although there is a need for long term surveillance of THA patients, the methods by which this has been achieved are varied. The service was originally provided by the orthopaedic medical team, under whose care the surgery was performed, but increasing pressures of work and a reduction in resources have affected this provision. This has led to the inclusion of non-medical health professionals in the orthopaedic team in order to undertake some of the work

traditionally performed by doctors. The training of these professionals in their extended roles has presented some challenges and has required the use of alternative models of learning to acquire the additional skills.

The interpretation of x-ray images of arthroplasty is an essential skill for any health professional involved in the long-term review of THA. An account was provided to demonstrate how this skill could be developed to a level of proficiency that was deemed competent for the purposes of routine arthroplasty review. An existing academic structure, which ratified evidence of work based learning, was used to show what had been achieved and to document the process so that it could provide a model for future training. The acquisition of this skill had immediate application in the establishment of a local arthroplasty review service and was transferable for use in research, as shown in subsequent chapters.

An important aspect of the review of radiographic images of THA is the estimation of any bone loss through the process of osteolysis. This can occur in the femur or in the pelvis and has been a major cause of aseptic loosening of THA (Harris 2004). There has been an emphasis on the development of prosthetic materials to reduce this phenomenon but it is still a significant issue in the surveillance of THA. The methods of measurement of osteolytic lesions on plain x-ray images have been varied and often involve a degree of visual estimation. The use of digital images still requires delineation of the area of interest and the irregularity of osteolytic lesions makes them difficult to define. The concept of a morphometric grid was developed to provide a clinical tool for measuring osteolytic lesions on x-ray images and was shown to be reliable for clinical or research purposes. It was interchangeable between the health professionals that were part of an orthopaedic team which supports its use in a busy clinical environment. It was also shown to be sufficiently sensitive to differentiate between groups of patients with a THA that had signs of deterioration – there was a statistically significant difference between groups who had been classified according to the severity of the changes. This morphometric grid provides a new clinical tool that can be used in routine clinical surveillance to assess the threat to the stability of a THA from osteolytic lesions.

In the clinical study that was reported in Chapter Five, a cohort of patients was reviewed in the mid-term period following THA. Two groups were represented in this cohort, one with an all cemented hip replacement and one group with hybrid THA. They were broadly representative of

UK practice at the time of their primary surgery which was completed in the years 2000 to 2003. The technique of hierarchical statistical regression analysis was used to explore the data for predictors of adverse radiographic change over the time from early review at three years to mid-term review.

It was found that the number of adverse radiographic changes was not predicted by the change in the OHS over the same period of time. Similarly, the age of the patient, the general health score of each and the Charlson comorbidity index were not predictive of the number of adverse radiographic changes. This is important in deciding what information is required when reviewing patients with THA.

If there are several radiographic changes evident on the x-ray of a THA, it is suggestive of aseptic loosening and the situation requires monitoring for further deterioration. Similarly, if the rate of wear of the polyethylene appears to be high, there is an increased likelihood of osteolytic lesions produced from the wear debris generated. The sub-group analysis of the clinical study confirmed that the age of a patient was not statistically related to the number of radiographic changes or the rate of wear of polyethylene.

The average wear rate for the patients with cemented THA and for those with a hybrid THA was comparable with other studies. There was a statistically significant increase in wear rate in the patients with evidence of osteolysis, as has been suggested in the orthopaedic literature.

6.2   Meanings and implications

The purpose of this study was to develop some aspects of assessment in hip arthroplasty review. The improvement of patient outcomes is an important aspect of this work and has been achieved through supporting the need for surveillance of THA patients. The literature review showed the overwhelming weight of orthopaedic opinion for long-term surveillance. The results from the clinical study showed that over 90% of patients had radiographic changes at mid-term, regardless of their age. The results also clearly showed that a joint specific PROM, despite being sensitive to change, is not predictive of the radiological state of the underlying hip replacement at this point in time and cannot be used as a surrogate measure of the THA. This information is important to healthcare providers in planning arthroplasty services. Cost saving cannot be achieved by monitoring a THA using a PROM without the addition of an x-ray. Also, patients cannot be selected purely on the basis of age as the number of radiographic changes and the steady state wear rate of polyethylene, both of which are predictors of possible deterioration of the joint, were not affected by age.

The frequently asymptomatic nature of osteolysis is well known in the orthopaedic community (Harris 2004), but as patients are often unaware of this 'silent disease', explaining its presence can be difficult as the patient associates the lack of symptoms with a successful outcome. The development of a clinical tool to measure the area of osteolytic lesions has an impact on patient care as it provides an objective measure on which to base a discussion about lesion size and to demonstrate a change in size. This allows a patient to understand the problem and to have increased involvement in any decision making about further treatment, as recommended in recent NHS policy (Department of Health 2004b).

Previous studies have addressed some aspects of the link between outcome scores and radiographic changes such as the predictive nature of a one time score for radiographic signs of loosening or revision, or the link between pain and radiographic changes (Malchau et al. 2005, Utting et al. 2008, Ollivere et al. 2009). However, no study was identified which specifically explored the link between an individual's change on a validated PROM, which is a recommended use of the OHS (Murray et al. 2007), and radiographic changes over the same period of time. It had been shown that the OHS changes significantly in patients with a THA that has deteriorated

to the point of requiring revision surgery (Dawson et al. 2001, Field et al. 2005). It was not previously known if a change score constructed from the OHS could replace the need for an x-ray at mid-term review before the THA has deteriorated to the point of needing revision. This study contributes new information to the scientific base for THA surveillance by demonstrating with robust statistical techniques that there is a need for an x-ray in addition to a widely used PROM at mid-term review.

The development of a morphometric grid into a reliable clinical tool is a further addition to the scientific base. Much of the assessment of osteolytic lesions in clinical practice is completed with an estimation of area and even the use of a digital imaging system does not provide a readily available and accurate solution. The morphometric grid is a clinical tool which is quick to apply and is reliable when used to assess the size of these irregular lesions, regardless of which member of the orthopaedic team is assessing the THA. It can be applied in a busy clinical situation or in a virtual surveillance system, and provides a reliable objective measure to add information to the assessment of the state of a hip replacement. The objectivity is of particular advantage when training non-medical health professionals to conduct arthroplasty review.

The increasing use of arthroplasty practitioners, both nationally and internationally, requires that consideration is given to all aspects of a service provided by such health professionals. The staff potentially suitable for this work will have a recognised health professional qualification (usually physiotherapy or orthopaedic nursing) and some evidence of study at a postgraduate level, indicating an ability to pursue further study. They are likely to need the support of a local orthopaedic surgeon or an arthroplasty mentor through an organisation such as ACPA. It is important that they are based in an orthopaedic team, even if the service is delivered away from a hospital site, due to the need for frequent communication with other members of the team.

The training of an arthroplasty practitioner could be achieved in a number of different ways depending on the individual situation. The potential practitioner would need to identify the local orthopaedic requirements in conjunction with the surgeons and the fund holders. The documents produced by the DH in conjunction with the BOA (British Orthopaedic Association 2010) provide well defined skill sets for each area of the patient pathway in hip replacement and can be used as the basis for describing the type and level of work required. The practitioner, having

identified areas of competency and areas for further study, would need to decide which methods of training were best suited to their situation.

The variety of training opportunities to develop arthroplasty practitioner skills includes access to HEIs for short courses or one or more master's level modules on relevant topics if available. If not, the use of an EWBL module allows skill development to be demonstrated through academic processes, as illustrated in Chapter Three. In addition, the developing practitioner could make use of ACPA resources (online as well as courses and conferences) or those offered by other professional bodies such as the Extended Scope Practitioners network of the CSP (Chartered Society of Physiotherapy 2009). The shadowing of an experienced arthroplasty practitioner would provide further valuable support and such evidence could be captured through a system of reflective writing. The supplementation of any or all of these methods with personal study of relevant orthopaedic literature would enhance the practitioners understanding of the concepts involved. Although these methods require personal application and time, any practitioner desiring to work at this advanced level will recognise the need for further work and the similarity of this work with other specialty training. As stated in Chapter Three, the opportunity to work at an advanced level often enhances job satisfaction for experienced health personnel (Ruston 2008). Further evaluation is needed of the methods by which these skills can be acquired and this work will be continued through ACPA and the international links that it has now established.

The external costs associated with training arthroplasty practitioners should be provided by the fund holders in order that individuals are not excluded by lack of available funds. This may include attendance at a short course or conference, or financial support for a master's module through an HEI. In addition, the health professional would require some protected time for personal study within working hours.

Arthroplasty practitioners, when trained, are working as advanced musculoskeletal practitioners and the associated salary is similar to that of a doctor in specialist training. As stated earlier (Chapter Three), there is an increasing tendency for non-medical health professionals to undertake work traditionally delivered by doctors and, in the case of arthroplasty, there are some benefits to this change. For example, an arthroplasty practitioner will provide stability in such a role as they are not required to move on after six months, as with trainee

doctors. This facilitates continuity for the patient in the extended pathway of care and may be further enhanced by the same practitioner working in other areas of the orthopaedic service which interface with the same patients. The experience gained as a health professional will provide the practitioner with good communication skills and allows a problem solving approach to any clinical interface which adds to the surgical and pharmaceutical options provided by the doctors. A well established arthroplasty practitioner will have knowledge of the preferences of the local orthopaedic team and is, consequently, in a good position to communicate this to the patient and to interpret the signs and symptoms described by the patient in light of their knowledge.

In addition, arthroplasty practitioners provide an excellent resource for collection of accurate data for audit or research purposes, and the distancing of the patient from the surgeon in assessing outcomes reduces bias and improves the quality of the information obtained (Learmonth and Cavendish 2005). This subsequently informs local service evaluation through knowledge of the local patient population and their typical scores, which allows managers and clinicians to meaningfully interpret and explain nationally produced data such as the PROMs reports (The NHS Information Centre 2011). Although national joint registers are currently unable to collect data on the period between primary surgery and revision, a well-established arthroplasty review service with knowledgeable arthroplasty practitioners has potential to supply important data on arthroplasty outcomes.

There is a recognised effect on medical trainees from those arthroplasty practitioners already in situ and this would increase with the employment of more of these health professionals. As the follow up of these patients is often carried out in designated arthroplasty clinics, the junior doctors no longer have the opportunity to assess the patients in routine orthopaedic clinics which undermines the learning experience through lack of pattern recognition. This problem could be addressed by allowing junior doctors to attend arthroplasty clinics as part of their training which also promotes communication between the medical and non-medical health professionals. In some centres, this is supplemented by a presentation about arthroplasty review given by a senior arthroplasty practitioner to the junior doctors at the start of their placement in an orthopaedic department.

There is also an effect on senior orthopaedic staff when an arthroplasty review service is established. For a newly appointed orthopaedic consultant, the feedback obtained by reviewing

their own patients over a period of time is important in shaping their ongoing practice and lack of follow up in their own clinics may have a detrimental effect. However, if there is good communication with the arthroplasty practitioner and follow up is assessed with appropriate scoring systems, regular feedback can be established as required. For senior orthopaedic consultants, the same system may be necessary if they introduce a new type of prosthesis or procedure to their practice. In addition, a joint review between the consultant and the arthroplasty practitioner may be advisable for an initial group of patients until the learning effect has reached a plateau and both are satisfied that the ongoing follow up can be continued by the arthroplasty service.

In all these potential situations, it would seem that it is important for the arthroplasty practitioner to be a part of the orthopaedic team to facilitate frequent communication. Although an arthroplasty service could be delivered in the community, the practitioner needs frequent interaction with the orthopaedic surgeons for exchange of information for the benefit of the patients. Changes in practice or use of newer prostheses will not reach the orthopaedic literature until long after the practitioner encounters them in clinics, and it is only by regular communication within the orthopaedic team that this exchange can take place.

6.3 General strengths and limitations

In general, the strength of this research was its representation of current surgical practice and follow up in the UK which supports the external validity of the results obtained. The number of participants included was sufficiently large for the statistical techniques employed although a larger sample from more than one hospital site and a wider variety of THA in the sample would improve the validity of the results obtained.

The PROM chosen for use in this research was widely available and its use in this study adds to the literature supporting its applicability in the long term follow up of THA. A possible limitation was the lack of analysis of the potential effect of psychosocial factors on the OHS. However, it is thought that in THA, the changes in psychosocial function that are experienced by the patient are relatively small and do not influence the outcome in a significant way (Learmonth and Cavendish 2005).

It is possible that personal bias has influenced this research through the researcher's own position as president of ACPA during the course of the study. The view expressed may be biased towards the positive contribution made by arthroplasty practitioners in THA surveillance and the desire to improve the educational opportunities for such health professionals. A barrier to implementation of the educational model described might be local unavailability of an EWBL module from an HEI. However, with increasing opportunities for electronic access to higher education, this research provides evidence of how an individual practitioner can use such a course to support professional development when it is available. The strengths of the position within ACPA have been the access to orthopaedic centres across the UK and the contact with practitioners in Canada and Australia, thereby enlarging the perspective of current THA surveillance.

This research did not include any service users in the study design or evaluation of the results obtained. Although there has been some work on the evaluation of patient satisfaction with a service provided by non-medical health professionals (Daker-White et al. 1999), this aspect requires further evaluation. Other aspects relating directly to the patients would require input from service users to consider the implications of a screening service and asymptomatic failure.

6.4 Further research

There are a number of concepts emerging from this study which warrant further research. As a primary consideration, service user input to future work would be essential in view of the asymptomatic nature of some of the changes around a THA. A screening service that identifies 'hidden' changes has ethical implications, and these must be considered in the planning of further research and dissemination of the results.

In relation to the arthroplasty practitioner role, it would be useful to examine the implications of a service provided by such health professionals. It may be possible to conduct a randomised controlled trial of early follow up comparing review by orthopaedic doctors with arthroplasty practitioners. Outcomes such as patient satisfaction with the service, outcomes of the review, health economics and adverse event occurrence, in addition to any outcomes identified from service user input, would provide information useful for future service planning.

A prospective study of the outcomes of hip arthroplasty review in the longer term (for example, from 15 to 20 years) would be useful to estimate costs and cost savings. Comparison of a group of patients with clinical follow-up could be compared with a group who received virtual follow-up and a group with no follow-up to estimate costs incurred and costs saved through timely referral for revision surgery. This may need to involve multiple orthopaedic units to obtain a sufficiently large sample but would benefit from the use of registry data and organisations such as ACPA.

In a separate consideration of future service planning, further research is required to explore the optimal time for the first and subsequent reviews. Although there is evidence that the adverse radiographic changes first appear in the mid-term period, many units still follow traditional recommendations to review in the early postoperative period. Patients requiring further treatment in the first few postoperative years are more likely to initiate the process themselves as their problem will be symptomatic. Consequently, it may be that a surveillance service is not needed for some patients in this early period.

Further research is also needed to determine how the information from hip arthroplasty review is best used for clinical decision making. The present study explored the relationship between a simple summation of radiographic changes and PROMs. It would be valuable to

develop and validate a simple radiographic scoring tool and then explore the magnitude of change which, in combination with validated patient-reported outcome scores, might indicate a need to consider revision surgery. Following a pilot study, this type of research would need to include multiple orthopaedic centres and prostheses for the results to have external validity. The inclusion of information on patient demographics would enhance the application of the results in future service planning. For instance, a ninety year old gentleman with an all cemented THA might have different requirements to a sixty year old female with a hybrid hip replacement that had been in place for ten years. Input from service users would provide a valuable patient perspective on this situation in addition to the orthopaedic view.

The morphometric grid has potential to contribute valuable information to these studies but further research is needed to establish its sensitivity to change through measuring a number of osteolytic lesions over time. In addition, information gained from its use by a number of different observers in different centres would facilitate calculation of a meaningful MDC to apply to clinical use. A further study of the relationship between the size (measured with the morphometric grid) and the location of lesions in patients who had subsequently received revision surgery may provide useful information for clinical decision making. These studies may be possible retrospectively with increasing use of digitised images although standardisation of patient position is required to obtain comparable images.

When the morphometric grid is transferred to a digital platform, it would be useful to conduct a feasibility study of its use in a clinical situation. A comparison between commonly used techniques such as 'eyeballing' or linear measurement could be made with measurements taken using the grid, the outcomes being the sensitivity to change and the time taken. In addition, a study of criterion validity would be useful - measuring areas of known osteolytic lesions, comparing them with the area as measured on a CT scan and exploring the relationship with volume.

Many of these suggestions for future work could be incorporated into a programme of research to investigate long term follow up in hip arthroplasty review. There would be two major streams to this research – the 'how' of the assessment process and the 'when' of the timing of review. The first would include the evaluation of the arthroplasty practitioners, the development of the tools used and the comparison of clinical and virtual surveillance. The second stream

would investigate the optimal timing for review in relation to variables such as patient demographics and type of THA, and the health economics. As stated earlier, the input of service users would be essential to this work, as would the professional perspective through existing organisations, specialist groups and expert opinion.

The aim of such a research programme would be to produce recommendations for follow up in hip arthroplasty review. The results of such research could be used to inform clinically and economically valid guidelines for long-term surveillance. The guidance produced could be presented in easy reference form for clinical use and may provide a framework for economic planning by orthopaedic units. The benefit to patients would be the information on which to base a decision about timely and appropriate clinical review. This potentially empowers a patient to take responsibility for requesting an arthroplasty review at a pre-determined time in order to monitor the long term status of their hip replacement.

Chapter 7

Conclusions

In this thesis, the assessment of hip arthroplasty has been considered in order to add to the evidence base for THA surveillance. In the absence of specific guidelines for hip arthroplasty review, the initial objective of the study was to search the current orthopaedic literature to establish the methods used in the assessment of THA.

In a healthcare environment where there appears to be a growing need for joint replacements alongside increasing economic constraints, the employment of non-medical health professionals to conduct arthroplasty review provides an alternative model of care to the traditional medical model. The development of this role requires additional skills and the methods by which these are currently acquired are varied. The second objective of the research was to describe a method used to develop a new skill in image interpretation of arthroplasty x-rays, and it may provide a model for training future arthroplasty practitioners.

The third objective of the research was the development of a clinical tool for use with x-rays. This tool was designed for the measurement of the area of osteolytic lesions which may be observed around a failing total hip replacement. These lesions are potentially threatening to the stability of the arthroplasty components and often appear without accompanying symptoms. The monitoring of osteolytic lesions and any change in size is an important component of hip arthroplasty review. This tool provides a reliable method of quantifying the size in order to monitor changes over time.

The final objective of the research was to explore the relationship between some of the tools currently used in hip arthroplasty review. A cohort of patients was selected for clinical review at mid-term, all of whom had either a fully cemented hip replacement or an uncemented acetabular cup with a cemented femoral stem. This cohort was representative of hip arthroplasty practice in the UK at the time of initial surgery between the years of 2000 and 2003. The results of the study showed that, although the OHS is sensitive to change at mid-term review, it cannot be used to predict the state of the underlying hip replacement as seen on x-ray images. The conclusion was that hip arthroplasty review at mid-term cannot be conducted using only a joint specific PROM but must include an x-ray to assess the state of the artificial joint.

The information gained from each part of this thesis is important for future service delivery. It contributes to the scientific evidence base of arthroplasty review and stimulates ideas for further research.

References

Arthroplasty Care Practitioners' Association (2006) Our aims. [online], *The Arthroplasty Care Practitioners Association*. Available from: http://www.acpa-uk.net/ [Accessed 28 February 2011].

Aiken, A.B., Atkinson, M., Harrison, M.M. and Hope, J. (2007) Reducing hip and knee replacement wait times: an expanded role for physiotherapists in orthopedic surgical clinics. *Healthcare Quarterly (Toronto, Ont.),* 10(2), 88-91.

Aiken, A.B., Harrison, M.M. and Hope, J. (2009) Role of the advanced practice physiotherapist in decreasing surgical wait times. *Healthcare Quarterly,* 12(3), 80-83.

Altenburg, A.J., Callaghan, J.J., Yehyawi, T.M., Pedersen, D.R., Liu, S.S., Leinen, J.A., Dahl, K.A., Goetz, D.D., Brown, T.D. and Johnston, R.C. (2009) Cemented total hip replacement cable debris and acetabular construct durability. *Journal of Bone & Joint Surgery - American Volume,* 91(7), 1664-1670.

Altman, D.G. and Bland, J.M. (1983) Measurement in medicine: the analysis of method comparison studies. *The Statistician,* 32, 307-317.

American Association of Orthopedic Surgeons (2010) American Joint Replacement Registry [online], *AAOS*. Available from: http://orthodoc.aaos.org/ajrr/ [Accessed 28 February 2011].

Amstutz, H.C., Campbell, P., Kossovsky, N. and Clarke, I.C. (1992) Mechanism and clinical significance of wear debris-induced osteolysis. *Clinical Orthopaedics & Related Research,* 276, 7-18.

Amstutz, H.C., Thomas, B.J., Jinnah, R., Kim, W., Grogan, T. and Yale, C. (1984) Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 66(2), 228-241.

Archibeck, M.J., Berger, R.A., Jacobs, J.J., Quigley, L.R., Gitelis, S., Rosenberg, A.G. and Galante, J.O. (2001) Second-generation cementless total hip arthroplasty: Eight to eleven-year results. *Journal of Bone & Joint Surgery - American Volume,* 83(11), 1666-1673.

REFERENCES

ASTM International. (2004) Standard test method for determining femoral head penetration into acetabular components of total hip replacement using clinical radiographs (F2385-04). *American Society for Testing and Materials International,* 1430-1434.

Atkinson, T. and Claxton, G., eds. (2000) *The Intuitive Practitioner.* Buckingham, UK: Open University Press.

Australian Orthopaedic Association (2010) National Joint Replacement Registry [online], *AOA*. Available from: http://www.dmac.adelaide.edu.au/aoanjrr/ [Accessed 24 August 2010].

Barrack, R.L., Folgueras, A., Munn, B., Tvetden, D. and Sharkey, P. (1997) Pelvic lysis and polyethylene wear at 5-8 years in an uncemented total hip. *Clinical Orthopaedics & Related Research,* 335, 211-217.

Barrack, R.L., Mulroy, R.D.,Jr and Harris, W.H. (1992) Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. *Journal of Bone and Joint Surgery - British Volume,* 74(3), 385-389.

Bellamy, N., Buchanan, W. and Goldsmith, C. (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *Journal of Orthopaedic Rheumatology,* 1, 95-98.

Benger, J.R. (2002) Can nurses working in remote units accurately request and interpret radiographs? *Emergency Medicine Journal,* 19(1), 68-70.

Berry, D.J., Harmsen, W.S. and Ilstrup, D.M. (1998) The natural history of debonding of the femoral component from the cement and its effect on long-term survival of Charnley total hip replacements. *Journal of Bone & Joint Surgery - American Volume,* 80(5), 715-721.

Bhabra, G., Sood, A., Fisher, B., Cartwright, L., Saunders, M., Evans, W.H., Surprenant, A., Lopez-Castejon, G., Mann, S., Davis, S.A., Hails, L.A., Ingham, E., Verkade, P., Lane, J., Heesom, K., Newson, R. and Case, C.P. (2009) Nanoparticles can cause DNA damage across a cellular barrier. *Nature Nanotechnology,* 4(12), 876-883.

Biring, G.S., Masri, B.A., Greidanus, N.V., Duncan, C.P. and Garbuz, D.S. (2007) Predictors of quality of life outcomes after revision total hip replacement. *Journal of Bone and Joint Surgery - British Volume,* 89(11), 1446-1451.

Bjorgul, K., Novicoff, W.M. and Saleh, K.J. (2010) Evaluating comorbidities in total hip and knee arthroplasty: available instruments. *Journal of Orthopaedics and Traumatology,* 11(4), 203-209.

REFERENCES

Bland, M.J. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet,* (February 8), 307-310.

Bolz, K.M.D., Crawford, R.W., Donnelly, B., Whitehouse, S.L. and Graves, N. (2010) The cost-effectiveness of routine follow-up after primary total hip arthroplasty. *Journal of Arthroplasty,* 25(2), 191-196.

Boud, D., Keough, R. and Walker, D., eds. (1985) *Reflection: Turning experience into learning.* London: Kogan Page.

Bowling, A. (2002) Research methods in health: Investigating health and health services. 2nd ed. Maidenhead: Open University Press.

Bozic, K.J., Kurtz, S.M., Lau, E., Ong, K., Vail, T.P. and Berry, D.J. (2009) The epidemiology of revision total hip arthroplasty in the United States. *Journal of Bone & Joint Surgery - American Volume,* 91(1), 128-133.

Brand, R.A. (2010) Biographical sketch: Sir John Charnley MD, 1911-1982. *Clinical Orthopaedics & Related Research,* 468:3147-8.

British Orthopaedic Association and British Association for Surgery of the Knee. (1999) *Knee replacement - A guide to good practice.* London: British Orthopaedic Association.

British Orthopaedic Association. (2006) *Primary total hip replacement: A guide to good practice.* London: British Orthopaedic Association.

British Orthopaedic Association. (2010) Arthroplasty care practitioners [online], *BOA*. Available from: http://www.boa.ac.uk/ [Accessed 11 February 2011].

Brooker, A.F., Bowerman, J.W., Robinson, R.A. and Riley, L.H.,Jr. (1973) Ectopic ossification following total hip replacement. Incidence and a method of classification. *Journal of Bone & Joint Surgery - American Volume,* 55(8), 1629-1632.

Brookfield, S.D. (1987) *Developing critical thinkers*. Maidenhead: Open University Press.

Bryant, M.J., Kernohan, W.G., Nixon, J.R. and Mollan, R.A. (1993) A statistical analysis of hip scores. *Journal of Bone and Joint Surgery - British Volume,* 75-B(5), 705-709.

Burns, A.W.R. and Bourne, R.B. (2006) Economics of revision total hip arthroplasty. *Current Orthopaedics,* 20(3), 203-207.

Burwell, R.G. (1997) Anthropometry. In: P.B. Pynsent, J.C.T. Fairbank and A.J. Carr (eds) *Assessment methodology in orthopaedics.* Oxford: Butterworth-Heinemann, pp 131-166.

REFERENCES

Cahir, J.G. and Toms, A.P. (2009) CT and MRI of hip replacements. *Orthopaedics and Trauma,* 23(2), 101-108.

Callaghan, J.J., Bracha, P., Liu, S.S., Piyaworakhun, S., Goetz, D.D. and Johnston, R.C. (2009) Survivorship of a Charnley total hip arthroplasty: A concise follow-up, at a minimum of thirty-five years, of previous reports. *Journal of Bone & Joint Surgery - American Volume,* 91(11), 2617-2621.

Callaghan, J.J., Templeton, J.E., Liu, S.S., Pedersen, D.R., Goetz, D.D., Sullivan, P.M. and Johnston, R.C. (2004) Results of Charnley total hip arthroplasty at a minimum of thirty years: A concise follow-up of a previous report. *Journal of Bone & Joint Surgery - American Volume,* 86(4), 690-695.

Calvert, G.T., Devane, P.A., Fielden, J., Adams, K. and Horne, J.G. (2009) A double-blind, prospective, randomized controlled trial comparing highly cross-linked and conventional polyethylene in primary total hip arthroplasty. *Journal of Arthroplasty,* 24(4), 505-510.

Campbell, D. (2011) NHS bans on operations gamble with patients' health, senior surgeon warns. *Guardian* [online] 17 January. Available from: http://www.guardian.cu.uk/2011/jan/17/nhs-bans-operations-surgeon-warns/.

Capello, W.N., D'Antonio, J.A., Feinberg, J.R. and Manley, M.T. (2003) Ten-year results with hydroxyapatite-coated total hip femoral components in patients less than fifty years old: A concise follow-up of a previous report. *Journal of Bone & Joint Surgery - American Volume,* 85(5), 885-889.

Centre for Evidence Based Medicine (2010) Levels of evidence #2 [online], *Centre for evidence based medicine.* Available from: http://www.cebm.net/ [Accessed 17 November 2010].

Charlson, M.E., Pompei, P., Ales, K.A. and MacKenzie, C.R. (1987) A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases,* 40(5), 373-383.

Charnley, J. (1972) The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *Journal of Bone and Joint Surgery - British Volume,* 54(1), 61-76.

Charnley, J. and Halley, D.K. (1975) Rate of wear in total hip replacement. *Clinical Orthopaedics & Related Research,* 112, 170-179.

Chartered Society of Physiotherapy (2009) The interactive CSP [online], *CSP.* Available from: http://www.interactivecsp.org.uk/ [Accessed 12 May 2009].

Chougle, A., Hemmady, M.V. and Hodgkinson, J.P. (2005) Severity of hip dysplasia and loosening of the socket in cemented total hip replacement: A long-term follow-up. *Journal of Bone and Joint Surgery - British Volume,* 87(1), 16-20.

REFERENCES

Claus, A.M., Engh, J., C.A, Sychterz, C.J., Xenos, J.S., Orishimo, K.F. and Engh, S., C.A. (2003) Radiographic definition of pelvic osteolysis following total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 85(8), 1519-1526.

Claxton, G., Atkinson, T., Osborn, M. and Wallace, M., eds. (1996) *Liberating the Learner.* London: Routledge.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioural Sciences*. 2nd ed. Hillsdale, USA: Lawrence Erlbaum Associates.

Corner, J.A., Rawoot, A. and Parmar, H.V. (2008) The thrust plate prosthesis in the treatment of osteoarthritis of the hip. Clinical and radiological outcome with minimum 5 year follow up. *Hip International,* 18(2), 88-94.

Crawford Adams, J. and Hamblen, D.L. (2001) *Outline of orthopaedics.* London: Churchill Livingstone.

Critical Appraisal Skills Programme. (2010) CASP tool for cohort studies [online], *Solutions for Public Health, National Health Service.* Available from: http://www.sph.nhs.net/ [Accessed 17 September 2010].

Daker-White, G., Carr, A.J., Harvey, I., Woolhead, G., Bannister, G., Nelson, I. and Kammerling, M. (1999) A randomised controlled trial. Shifting boundaries of doctors and physiotherapists in orthopaedic outpatient departments. *Journal of Epidemiology and Community Health,* 53(10), 643-650.

Dall, D.M., Learmonth, I.D. and Solomon, M. (1992) A scoring system to determine radiological loosening in cemented total hip arthroplasty. *International Orthopaedics, 16, 339-343.*

Daniel, J., Pynsent, P.B. and McMinn, D.J.W. (2004) Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *Journal of Bone and Joint Surgery - British Volume,* 86(2), 177-184.

Darrah, C. (2006) Follow up after Joint Replacement Surgery: A survey of orthopaedic surgeons and professionals allied to medicine. In: *Proceedings of the British Hip Society Annual Meeting, Edinburgh, 2-3 March 2006.*

Davies, A.P., Willert, H.G., Campbell, P.A., Learmonth, I.D. and Case, C.P. (2005) An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *Journal of Bone & Joint Surgery - American Volume,* 87(1), 18-27.

Dawson, J., Fitzpatrick, R., Carr, A. and Murray, D. (1996a) Questionnaire on the perceptions of patients about total hip replacement. *Journal of Bone and Joint Surgery - British Volume,* 78(2), 185-90.

REFERENCES

Dawson, J., Fitzpatrick, R., Murray, D. and Carr, A. (1996b) The problem of 'noise' in monitoring patient-based outcomes: generic, disease-specific and site-specific instruments for total hip replacement. *Journal of Health Services Research and Policy,* 1(4), 224-231.

Dawson, J., Fitzpatrick, R., Frost, S., Gundle, R., McLardySmith, P. and Murray, D. (2001) Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *Journal of Bone and Joint Surgery - British Volume,* 83(8), 1125-1129.

Dawson, J., Fitzpatrick, R., Chruchmane, D., Verjee-Lorenz, A. and Clayson, D. (2010) *User manual for the Oxford Hip Score (OHS). Version 1.0.* Oxford: Isis Innovation Ltd. Available from: http://www.isis-innovation.com/outcomes/index.html [Accessed 30 March 2011].

DeLee, J.G. and Charnley, J. (1976) Radiological demarcation of cemented sockets in total hip replacement. *Clinical Orthopaedics & Related Research,* 121, 20-31.

Della Valle, C.J. and Paprosky, W.G. (2002) The middle-aged patient with hip arthritis: The case for extensively coated stems. *Clinical Orthopaedics & Related Research,* 405, 101-107.

Della Valle, C.J. and Paprosky, W.G. (2004) The femur in revision total hip arthroplasty evaluation and classification. *Clinical Orthopaedics & Related Research,* 420, 55-62.

Della Valle, C.J., Mesko, N.W., Quigley, L., Rosenberg, A.G., Jacobs, J.J. and Galante, J.O. (2009) Primary total hip arthroplasty with a porous-coated acetabular component. A concise follow-up, at a minimum of twenty years, of previous reports. *Journal of Bone & Joint Surgery - American Volume,* 91(5), 1130-1135.

Department of Health (2000) *Meeting the challenge: A strategy for the Allied Health Professions.* London: Department of Health.

Department of Health (2004a) *A compendium of solutions to implementing the Working Time Directive for doctors in training from August 2004.* London: Department of Health. Available from: http://www.dh.gov.uk/ [Accessed 10 February 2011].

Department of Health (2004b) *Patient and public involvement in health: The evidence for policy implementation.* London: Department of Health. Available from: http://www.dh.gov.uk/ [Accessed 2 April 2011].

Department of Health. (2008) *Guidance on the routine collection of patient reported outcome measures (PROMS)* [online]. London: Department of Health. Available from: http://www.dh.gov.uk/ [Accessed 17 March 2010].

Dobzyniak, M., Fehring, T.K. and Odum, S. (2006) Early failure in total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 447, 76-78.

REFERENCES

Donoghue, D. and Stokes, E.K. (2009) How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *Journal of Rehabilitation Medicine,* 41, 343-346.

Dorr, L.D. and Wan, Z. (1996) Comparative results of a distal modular sleeve, circumferential coating, and stiffness relief using the anatomic porous replacement II. *Journal of Arthroplasty,* 11(4), 419-428.

Duffy, P.J., Masri, B.A., Garbuz, D.S. and Duncan, C.P. (2005) Evaluation of patients with pain following total hip replacement. *Journal of Bone & Joint Surgery - American Volume,* 87(11), 2566-2575.

Dumbleton, J.H., Manley, M.T. and Edidin, A.A. (2002) A literature review of the association between wear rate and osteolysis in total hip arthroplasty. *The Journal of Arthroplasty,* 17(5), 649-661.

El Masri, F., Kerboull, L., Kerboull, M., Courpied, J.P. and Hamadouche, M. (2010) Is the so-called 'French paradox' a reality?: Long-term survival and migration of the Charnely-Kerboull stem cemented line-to-line. *Journal of Bone and Joint Surgery - British Volume,* 92(3), 342-348.

Elias, H., Hennig, A. and Schwartz, D.E. (1971) Stereology: Applications to biomedical research. *Physiological Reviews,* 51(1), 158-200.

Eliasziw, M., Young, S.L., Woodbury, M.G. and FrydayField, K. (1994) Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Physical Therapy,* 74(8), 777-788.

Elke, R., Berli, B., Wagner, A. and Morscher, E.W. (2003) Acetabular revision in total hip replacement with a press-fit cup. *Journal of Bone and Joint Surgery - British Volume,* 85(8), 1114-1119.

Emms, N.W., Stockley, I., Hamer, A.J. and Wilkinson, J.M. (2010) Long-term outcome of a cementless, hemispherical, press-fit acetabular component: Survivorship analysis and dose-response relationship to linear polyethylene wear. *Journal of Bone and Joint Surgery - British Volume,* 92(6), 856-861.

Engh, C.A., Bobyn, J.D. and Glassman, A.H. (1987) Porous-coated hip replacement. The factors governing bone ingrowth, stress shielding, and clinical results. *Journal of Bone and Joint Surgery - British Volume,* 69(1), 45-55.

Engh, C.A., Massin, P. and Suthers, K.E. (1990) Roentgenographic assessment of the biologic fixation of porous-surfaced femoral components. *Clinical Orthopaedics & Related Research,* 257, 107-128.

REFERENCES

Eskelinen, A., Helenius, I., Remes, V., Ylinen, P., Tallroth, K. and Paavilainen, T. (2006) Cementless total hip arthroplasty in patients with high congenital hip dislocation. *Journal of Bone & Joint Surgery - American Volume,* 88(1), 80-91.

Ethgen, O., Bruyere, O., Richy, F., Dardennes, C. and Reginster, J.-. (2004) Health-related quality of life in total hip and total knee arthroplasty: A qualitative and systematic review of the literature. *Journal of Bone & Joint Surgery - American Volume,* 86(5), 963-974.

EuroQol Group. (2010) EQ-5D: A standardised instrument for use as a measure of health outcome [online], *EuroQol Group*. Available from: http://www.euroqol.org/ [Accessed 9 October 2010].

Field, R.E., Cronin, M.D. and Singh, P.J. (2005) The Oxford hip scores for primary and revision hip replacement. *Journal of Bone and Joint Surgery - British Volume,* 87(5), 618-622.

Fitzpatrick, R., Morris, R., Hajat, S., Reeves, B., Murray, D.W., Hannen, D., Rigge, M., Williams, O. and Gregg, P. (2000) The value of short and simple measures to assess outcomes for patients of total hip replacement surgery. *Quality in Health Care,* 9, 146-50.

Furmanski, J., Anderson, M., Bal, S., Greenwald, A.S., Halley, D., Penenberg, B., Ries, M. and Pruitt, L. (2009) Clinical fracture of cross-linked UHMWPE acetabular liners. *Biomaterials,* 30(29), 5572-5582.

Furnes, O., Lie, S.A., Espehaug, B., Vollset, S.E., Engesaeter, L.B. and Havelin, L.I. (2001) Hip disease and the prognosis of total hip replacements: A review of 53 698 primary total hip replacements reported to the Norwegian Arthroplasty Register 1987-99. *Journal of Bone and Joint Surgery - British Volume,* 83(4), 579.

Garbuz, D.S., Xu, M. and Sayre, E.C. (2006) Patients' outcome after total hip arthroplasty. A comparison between the Western Ontario and McMaster Universities index and the Oxford 12-item hip score. *Journal of Arthroplasty,* 21(7), 998-1004.

Geerdink, C.H., Grimm, B., Vencken, W., Heyligers, I.C. and Tonino, A.J. (2009) Cross-linked compared with historical polyethylene in THA: an 8-year clinical study. *Clinical Orthopaedics & Related Research,* 467, 979-984.

Ghoz, A. and Macdonald, D. (2008) New trends in total hip replacement: follow-up is it required and who pays? *Current Orthopaedics,* 22(3), 173-176.

Gonzalez Della Valle, A., Zoppi, A., Peterson, M.G.E. and Salvati, E.A. (2004) Clinical and radiographic results associated with a modern, cementless modular cup design in total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 86(9), 1998-2004.

Gosling, S. (1999) Physiotherapy and postgraduate study: A follow-up discussion paper. *Physiotherapy,* 85(3), 117-121.

REFERENCES

Green, A., Perry, J. and Harrison, K. (2008) The influence of a postgraduate clinical master's qualification in manual therapy on the careers of physiotherapists in the United Kingdom. *Manual therapy,* 13(2), 139-147.

Griffith, M.J., Seidenstein, M.K., Williams, D. and Charnley, J. (1978) Socket wear in Charnley low friction arthroplasty of the hip. *Clinical Orthopaedics & Related Research,* 137, 37-47.

Grübl, A., Chiari, C., Giurea, A., Gruber, M., Kaider, A., Marker, M., Zehetgruber, H. and Gottsaunder-Wolf, F. (2006) Cementless total hip arthroplasty with the rectangular titanium Zweymüller stem: A concise follow-up, at a minimum of fifteen years, of a previous report. *Journal of Bone & Joint Surgery - American Volume,* 88(10), 2210-5.

Gruen, T.A., McNeice, G.M. and Amstutz, H.C. (1979) "Modes of failure" of cemented stem-type femoral components: a radiographic analysis of loosening. *Clinical Orthopaedics & Related Research,* 141, 17-27.

Haddad, F.S., Ashby, E. and Konangamparambath, S. (2007) Should follow-up of patients with arthroplasties be carried out by general practitioners? *Journal of Bone & Joint Surgery - British Volume,* 89(9), 1133-1134.

Hallan, G., Dybvik, E., Furnes, O. and Havelin, L.I. (2010) Metal-backed acetabular components with conventional polyethylene: A review of 9113 primary components with a follow-up of 20 years. *Journal of Bone and Joint Surgery - British Volume,* 92(2), 196-201.

Hamilton, W.G., Hopper, J.,Robert H., Ginn, S.D., Hammell, N.P., Engh, J., Anderson, C. and Engh, C.A. (2005) The effect of total hip arthroplasty cup design on polyethylene wear rate. *Journal of Arthroplasty,* 20(Supplement 3), 63-72.

Han, C., Choe, W. and Yoo, J. (1999) Effect of polyethylene wear on osteolysis in cementless primary total hip arthroplasty: Minimal 5-year follow-up study. *Journal of Arthroplasty,* 14(6), 714-723.

Hardy, M.L. and Barrett, C. (2003) Requesting and interpreting trauma radiographs: a role extension for accident & emergency nurses. *Accident & Emergency Nursing,* 11(4), 202-213.

Harris, W.H. (1969) Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *Journal of Bone & Joint Surgery - American Volume,* 51(4), 737-755.

Harris, W.H. (2004) Conquest of a worldwide human disease: particle-induced periprosthetic osteolysis. *Clinical Orthopaedics & Related Research,* 429, 39-42.

Harris, W.H., McCarthy, J., J.C and O'Neill, D.A. (1982) Femoral component loosening using contemporary techniques of femoral cement fixation. *Journal of Bone & Joint Surgery - American Volume,* 64(7), 1063-1067.

REFERENCES

Hart, A.J., Sabah, S., Henckel, J., Lewis, A., Cobb, J., Sampson, B., Mitchell, A. and Skinner, J.A. (2009) The painful metal-on-metal hip resurfacing. *Journal of Bone and Joint Surgery - British Volume,* 91(6), 738-744.

Hartofilakidis, G., Georgiades, G., Babis, G.C. and Yiannakopoulos, C.K. (2008) Evaluation of two surgical techniques for acetabular reconstruction in total hip replacement for congenital hip disease: Results after a minimum ten-year follow-up. *Journal of Bone and Joint Surgery - British Volume,* 90(6), 724-730.

Harvie, P., Pollard, T.C.B., Chennagiri, R.J. and Carr, A.J. (2005) The use of outcome scores in surgery of the shoulder. *Journal of Bone and Joint Surgery - British Volume,* 87(2), 151-154.

Hernigou, P., Zilber, S., Filippini, P. and Poignard, A. (2009) Ceramic-ceramic bearing decreases osteolysis: a 20-year study versus ceramic-polyethylene on the contralateral hip. *Clinical Orthopaedics & Related Research,* 467, 2274-2280.

Hodgkinson, J.P., Shelley, P. and Wroblewski, B.M. (1988) The correlation between the roentgenographic appearance and operative findings at the bone-cement junction of the socket in Charnley low friction arthroplasties. *Clinical Orthopaedics & Related Research,* 228, 105-109.

Holm, I., Bolstad, B., Lutken, T., Ervik, A., Rokkum, M. and Steen, H. (2000) Reliability of goniometric measurements and visual estimates of hip ROM in patients with osteoarthrosis. *Physiotherapy Research International,* 5(4), 241.

Holt, G., Murnaghan, C., Reilly, J. and Meek, R.M.D. (2007) The biology of aseptic osteolysis. *Clinical Orthopaedics & Related Research,* 460, 240-252.

Honey, P. (1997) *Improve your people skills*. 2nd ed. London: Chartered Institute of Personnel Development.

Horne, G., Culliford, N., Adams, K. and Devane, P. (2007) Hybrid total hip replacement: Outcome after a mean follow up of 10 years. *ANZ Journal of Surgery,* 77(8), 638-641.

Howard, P.W. (2009) Reconstruction/revision of acetabular failure in revision hip replacement. *Orthopaedics and Trauma,* 23(5), 322-326.

Hozack, W.J., Mesa, J.J., Carey, C. and Rothman, R.H. (1996) Relationship between polyethylene wear, pelvic osteolysis, and clinical symptomatology in patients with cementless acetabular components: A framework for decision making. *The Journal of Arthroplasty,* 11(7), 769-772.

REFERENCES

Huddleston, J.I., Harris, A.H.S., Atienza, C.A. and Woolson, S.T. (2010) Hylamer vs conventional polyethylene in primary total hip arthroplasty: A long-term case-control study of wear rates and osteolysis. *Journal of Arthroplasty,* 25(2), 203-207.

Hugill, L., Foguet, P. and Costa, M.L. (2010) Setting up a 'virtual' arthroplasty surveillance plan in the modern NHS: problems, pitfalls and persistence. In: *Proceedings of the British Association for Surgery to the Knee, Annual Meeting, Oxford, 25-26 March 2010.*

Huo, M.H., Stockton, K.G., Mont, M.A. and Parvizi, J. (2010) What's new in total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 92, 2959-2972.

Ihle, M., Mai, S., Pfluger, D. and Siebert, W. (2008) The results of the titanium-coated RM acetabular component at 20 years: A long-term follow-up of an uncemented primary total hip replacment. *Journal of Bone and Joint Surgery - British Volume,* 90(10), 1284-1290.

Iida, H., Matsusue, Y., Kawanabe, K., Okumura, H., Yamamuro, T. and Nakamura, T. (2000) Cemented total hip arthroplasty with acetabular bone graft for developmental dysplasia: Long-term results and survivorship analysis. *Journal of Bone and Joint Surgery - British Volume,* 82(2), 176-184.

Illich, I. (1995) *Limits to medicine. Medical nemesis: The expropriation of health*. London: Marion Boyars Publishers Ltd.

Incavo, S.J., Beynnon, B.D. and Coughlin, K.M. (2008) Total hip arthroplasty with the Secur-Fit and Secur-Fit Plus femoral stem design. A brief follow-up report at 5 to 10 Years. *Journal of Arthroplasty,* 23(5), 670-676.

Ito, H., Matsuno, T., Aok, Y. and Minami, A. (2004) Total hip arthroplasty using an Omniflex modular system: 5 to 12 years follow up. *Clinical Orthopaedics & Related Research,* 419, 98-106.

Jacobs, C.A., Christensen, C.P., Greenwald, A.S. and McKellop, H. (2007) Clinical performance of highly cross-linked polyethylenes in total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 89(12), 2779-2786.

Jarrett, B.T., Cofske, J., Rosenberg, A.E., Oral, E., Muratoglu, O. and Malchau, H. (2010) In vivo biological response to vitamin E and vitamin-E-doped polyethylene. *Journal of Bone & Joint Surgery - American Volume,* 92(16), 2672-2681.

Jensen, G.M., Gwyer, J. and Shepard, K.F. (2000) Expert practice in physical therapy. *Physical Therapy,* 80(1), 28.

Johnston, R.C. and Larson, C.B. (1969) Results of treatment of hip disorders with cup arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 51(8), 1461-1479.

REFERENCES

Johnston, R.C., Fitzgerald, R.H., Harris, W.H., Poss, R., Muller, M.E. and Sledge, C.B. (1990) Clinical and radiographic evaluation of total hip replacement. A standard system of terminology for reporting results. *Journal of Bone & Joint Surgery - American Volume,* 72(2), 161-168.

Joshi, R.P., Eftekhar, N.S., McMahon, D.J. and Nercessian, O.A. (1998) Osteolysis after Charnley primary low-friction arthroplasty: A comparison of two matched paired groups. *Journal of Bone and Joint Surgery - British Volume,* 80(4), 585-590.

Kaiser, L. (1989) Adjusting for baseline: change or percentage change? *Statistics in Medicine,* 8, 1183-1190.

Karanicolas P.J., Bhandari M., Kreder H., Moroni A., Richardson M., Walter S.D., Norman G.R., Guyatt G.H. and Collaboration for Outcome Assessment in Surgical Trials (COAST) Musculoskeletal Group. (2009) Evaluating agreement: conducting a reliability study. *Journal of Bone & Joint Surgery - American Volume,* 91(Suppl 3), 99-106.

Karrholm, J. (2010) The Swedish hip arthroplasty register. *Acta Orthopaedica,* 81(1), 3-4.

Karrholm, J. (2011) Personal communication.

Keener, J.D., Callaghan, J.J., Goetz, D.D., Pederson, D.R., Sullivan, P.M. and Johnston, R.C. (2003) Twenty-five-year results after Charnley total hip arthroplasty in patients less than fifty years old: a concise follow-up of a previous report. *Journal of Bone & Joint Surgery - American Volume,* 85(6), 1066-1072.

Kersten, P., McPherson, K., Lattimer, V., George, S., Breton, A. and Ellis, B. (2007) Physiotherapy extended scope of practice - who is doing what and why? *Physiotherapy,* 93(4), 235-242.

Kim, Y., Oh, S., Kim, J. and Koo, K. (2003) Contemporary total hip arthroplasty with and without cement in patients with osteonecrosis of the femoral head. *Journal of Bone & Joint Surgery - American Volume,* 85(4), 675-681.

Kim, Y.H. (2005) Comparison of polyethylene wear associated with cobalt-chromium and zirconia heads after total hip replacement. A prospective, randomized study. *Journal of Bone & Joint Surgery - American Volume,* 87(8), 1769-1776.

Kim, Y.H. (2008) The results of a proximally-coated cementless femoral component in total hip replacement: a five- to 12-year follow-up. *Journal of Bone & Joint Surgery - British Volume,* 90(3), 299-305.

Kim, Y.H., Kook, H.K. and Kim, J.S. (2002) Total hip replacement with a cementless acetabular component and a cemented femoral component in patients younger than fifty years of age. *Journal of Bone & Joint Surgery - American Volume,* 84(5), 770-774.

REFERENCES

Kind, P., Hardman, G. and Macran, S. (1999) *UK Population norms for EQ-5D* (DP 172). Centre of Health Economics, University of York.

Kitamura, N., Pappedemos, P.C., Duffy, P.R.,3rd, Stepniewski, A.S., Hopper, R.H., Jr, Engh, C.A.,Jr and Engh, C.A. (2006a) The value of anteroposterior pelvic radiographs for evaluating pelvic osteolysis. *Clinical Orthopaedics & Related Research,* 453, 239-245.

Kitamura, N., Sychterz-Terefenko, C.J. and Engh, S.,Charles A. (2006b) The temporal progression of pelvic osteolysis after uncemented total hip arthroplasty. *Journal of Arthroplasty,* 21(6), 791-795.

Kobayashi, A., Donnelly, W.J., Scott, G. and Freeman, M.A.R. (1997) Early radiological observations may predict the long-term survival of femoral hip prostheses. *Journal of Bone and Joint Surgery - British Volume,* 79(4), 583-589.

Kolb, D.A. (1984) *Experiential Learning*. London: Prentice-Hall International.

Langlais, F., Kerboull, M., Sedel, L. and Ling, R.S.M. (2003) The 'French Paradox'. *Journal of Bone and Joint Surgery - British Volume,* 85(1), 17-20.

Langton, D.J., Jameson, S.S., Joyce, T.J., Hallab, N.J., Natu, S. and Nargol, A.V.F. (2010) Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: A consequence of excess wear. *Journal of Bone and Joint Surgery - British Volume,* 92(1), 38-46.

Latimer, H.A. and Lachiewicz, P.F. (1996) Porous-coated acetabular components with screw fixation. Five to ten-year results. *Journal of Bone & Joint Surgery - American Volume,* 78(7), 975-81.

Lavernia, C.J. (1998) Cost-effectiveness of early surgical intervention in silent osteolysis. *The Journal of Arthroplasty,* 13(3), 277-279.

Learmonth, I.D. and Cavendish, V.J. (2005) Outcome assessment following total hip replacement. *Orthopedics,* 28(8), 827-30.

Levine, B.R., Sporer, S., Poggie, R.A., Della Valle, C.J. and Jacobs, J.J. (2006) Experimental and clinical performance of porous tantalum in orthopedic surgery. *Biomaterials,* 27(27), 4671-4681.

Li, L.C., Westby, M.D., Sutton, E., Thompson, M., Sayre, E.C. and Casimiro, L. (2009) Canadian physiotherapists' views on certification, specialisation, extended role practice, and entry-level training in rheumatology. *BMC Health Services Research,* 9.

REFERENCES

Littlejohn, F., Nahna, M., Newland, C., Robins, S. and Helford, C. (2006) What are the protocols and procedures for imaging referral by physiotherapists? *New Zealand Journal of Physiotherapy,* 34(2), 81-87.

Livermore, J., Ilstrup, D. and Morrey, B. (1990) Effect of femoral head size on wear of the polyethylene acetabular component. *Journal of Bone & Joint Surgery - American Volume,* 72(4), 518-528.

LoBiondo-Wood, G. and Haber, J. (1998) *Nursing research: methods, critical appraisal, and utilization.* 4th ed. St. Louis, Missouri: Mosby, Inc.

Lusty, P.J., Walter, W.L., Walter, W.K. and Zicat, B. (2007) Cementless hip arthroplasty in Paget's disease at medium-term follow-up (average of 6.7 years). *Journal of Arthroplasty,* 22(5), 692-696.

Mabry, T.M., Prpa, B., Haidukewych, G.J., Harmsen, W.S. and Berry, D.J. (2004) Long-term results of total hip arthroplasty for femoral neck fracture nonunion. *Journal of Bone & Joint Surgery - American Volume,* 86(10), 2263-2267.

Makela, K., Eskelinen, A., Pulkkinen, P., Paavolainen, P. and Remes, V. (2008) Cemented total hip replacement for primary osteoarthritis in patients aged 55 years or older: results of the 12 most common cemented implants followed for 25 years in the Finnish Arthroplasty Register. *Journal of Bone & Joint Surgery - British Volume,* 90(12), 1562-1569.

Malchau, H., Garellick, G., Eisler, T., Karrholm, J. and Herberts, P. (2005) The Swedish hip registry: Increasing the sensitivity by patient outcome data. *Clinical Orthopaedics & Related Research,* 441, 19-29.

Malchau, H., Herberts, P., Eisler, T., Garellick, G. and Soderman, P. (2002) The Swedish total hip replacement register. *Journal of Bone & Joint Surgery - American Volume,* 84(2), 2-20.

Maloney, W.J., Galante, J.O., Anderson, M., Goldberg, V., Harris, W.H., Jacobs, J., Kraay, M., Lachiewicz, P., Rubash, H.E., Schutzer, S. and Woolson, S.T. (1999) Fixation, polyethylene wear, and pelvic osteolysis in primary total hip replacement. *Clinical Orthopaedics & Related Research,* 369, 157-164.

Mancuso, C.A., Jout, J., Salvati, E.A. and Sculco, T.P. (2009) Fulfillment of patients' expectations for total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 91(9), 2073-2078.

Martell, J.M., Pierson III, R.H., Jacobs, J.J., Rosenberg, A.G., Maley, M. and Galante, J.O. (1993) Primary total hip reconstruction with a titanium fiber-coated prosthesis inserted without cement. *Journal of Bone & Joint Surgery - American Volume,* 75(4), 554-571.

REFERENCES

Massin, P., Schmidt, L. and Engh, C.A. (1989) Evaluation of cementless acetabular component migration: An experimental study. *Journal of Arthroplasty,* 4(3), 245-251.

Matthews, J.B., Besong, A.A., Green, T.R., Stone, M.H., Wroblewski, B.M., Fischer, J. and Ingham, E. (2000) Evaluation of the response of primary human peripheral blood mononuclear phagocytes to challenge with in vitro generated clinically relevant UHMWPE particles of known size and dose. *Journal of Biomedical Materials Research,* 52(2), 296-307.

McBride, T.J. and Prakash, D. (2010) How to read a postoperative total hip replacement radiograph. *Postgraduate Medical Journal,* 87, 101-109.

McCalden, R.W., MacDonald, S.J., Rorabeck, C.H., Bourne, R.B., Chess, D.G. and Charron, K.D. (2009) Wear rate of highly cross-linked polyethylene in total hip arthroplasty: A randomized controlled trial. *Journal of Bone & Joint Surgery - American Volume,* 91(4), 773-782.

McCalden, R.W., Naudie, D.D., Yuan, X. and Bourne, R.B. (2005) Radiographic Methods for the Assessment of Polyethylene Wear After Total Hip Arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 87(10), 2323-2334.

McCombe, P. and Williams, S.A. (2004) A comparison of polyethylene wear rates between cemented and cementless cups. A prospective, randomised trial. *Journal of Bone and Joint Surgery - British Volume,* 86(3), 344-349.

McMinn, D.J., Daniel, J., Ziaee, H. and Pradhan, C. (2008) Results of the Birmingham Hip Resurfacing dysplasia component in severe acetabular insufficiency: a six- to 9.6-year follow-up. *Journal of Bone & Joint Surgery - British Volume,* 90(6), 715-723.

McPherson, K., Kersten, P., George, S., Lattimer, V., Breton, A., Ellis, B., Kaur, D. and Frampton, G. (2006) A systematic review of evidence about extended roles for allied health professionals. *Journal of Health Services & Research Policy,* 11(4), 240-247.

McWilliams, T.G. and Parker, J.R. (2008) Alternative bearing surfaces for hip arthroplasty. *Orthopaedics and Trauma,* 23(1), 2-7.

Mehin, R., Yuan, X., Haydon, C., Rorabeck, C.H., Bourne, R.B., McCalden, R.W. and MacDonald, S.J. (2004) Retroacetabular osteolysis: When to operate? *Clinical Orthopaedics & Related Research,* 428, 247-255.

Mehin, R. (2011) Letter to the Editor: Use of morphometry to quantify osteolysis after total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 469, 617.

Merle D'Aubigne, R.M. and Postel, M. (1954) Functional results of hip arthroplasty with acrylic prosthesis. *Journal of Bone & Joint Surgery - American Volume,* 36(3), 451-475.

REFERENCES

Miura, H., Matsuda, S., Okazaki, K., Kawano, T., Kawamura, H. and Iwamoto, Y. (2005) Validity of an oblique posterior condylar radiographic view for revision total knee arthroplasty. *Journal of Bone & Joint Surgery - British Volume,* 87(12), 1643-1646.

Morris, R.W. (1997) Assessing the reliability of clinical measurement. In: P.B. Pynsent, J.C.T. Fairbank and A.J. Carr, eds, *Assessment methodology in orthopaedics.* Oxford: Butterworth-Heinemann, pp 1-18.

Mu, Z., Tian, J., Wu, T., Yang, J. and Pei, F. (2009) A systematic review of radiological outcomes of highly cross-linked polyethylene versus conventional polyethylene in total hip arthroplasty. *International Orthopaedics,* 33(3), 599-604.

Muller, U., Gautier, E., Roeder, C. and Busato, A. (2003) The relationship between cup design and the radiological signs of aseptic loosening in total hip arthroplasty. *Journal of Bone and Joint Surgery - British Volume,* 85(1), 31-36.

Mulroy, W.F., Estok, D.M. and Harris, W.H. (1995) Total hip arthroplasty with use of so-called second-generation cementing techniques. A fifteen-year-average follow-up study. *Journal of Bone & Joint Surgery - American Volume,* 77(12), 1845-1852.

Murray, D.W., Fitzpatrick, R., Rogers, K., Pandit, H., Beard, D.J., Carr, A.J. and Dawson, J. (2007) The use of the Oxford hip and knee scores. *Journal of Bone & Joint Surgery - British Volume,* 89(8), 1010-1014.

National Joint Registry (2010) *7th annual report* [online]. London: Department of Health. Available from: http://www.njrcentre.org.uk/ [Accessed 14 February 2011].

Needham, J., Burns, T. and Gerlinger, T. (2008) Catastrophic failure of ceramic-polyethylene bearing total hip arthroplasty. *Journal of Arthroplasty,* 23(4), 627-630.

Neill, J. (2008) *Why use effect sizes instead of significance testing in program evaluation?* [online]. Available from: http://wilderdom.com/effectsizes.html [Accessed 9 April 2011]

NHS Choices Information (2010) Hip replacement [online], *NHS Choices.* Available from: http://www.nhs.uk/conditions/hip-replacement/ [Accessed 23 March 2011].

NHS National Services Scotland. (2010) The Scottish Arthroplasty Project [online], *NHS Scotland*. Available from: http://www.arthro.scot.nhs.uk/ [Accessed 14 February 2011].

Nilsdotter, A.K., Roos, E.M., Westerlund, J.P., Roos, H.P. and Lohmander, L.S. (2001) Comparative responsiveness of measures of pain and function after total hip replacement. *Arthritis and Rheumatism,* 45(3), 258-262.

REFERENCES

Nixon, M., Taylor, G., Sheldon, P., Iqbal, S.J. and Harper, W. (2007) Does bone quality predict loosening of cemented total hip replacements? *Journal of Bone & Joint Surgery - British Volume,* 89(10), 1303-1308.

Norton, M.R., Yarlagadda, R. and Anderson, G.H. (2002) Catastrophic failure of the Elite Plus total hip replacement, with a Hylamer acetabulum and Zirconia ceramic femoral head. *Journal of Bone and Joint Surgery - British Volume,* 84(5), 631-635.

Ogino, D., Kawaji, H., Konttinen, L., Lehto, M., Rantanen, P., Malmivaara, A., Konttinen, Y.T. and Salo, J. (2008) Total hip replacement in patients eighty years of age and older. *Journal of Bone & Joint Surgery - American Volume,* 90(9), 1884-1890.

Ollivere, B., Darrah, C., Brankin, R.C., Donnell, S.T. and Walton, N.P. (2009) The continued value of clinical and radiological surveillance: The Charnley elite plus hip replacement system at 12 years. *Journal of Bone and Joint Surgery - British Volume,* 91(6), 720-724.

Olyslaegers, C., Defoort, K., Simon, J. and Vandenberghe, L. (2008) Wear in conventional and highly cross-linked polyethylene cups: A 5-Year Follow-Up Study. *Journal of Arthroplasty,* 23(4), 489-494.

O'Neill, D.A. and Harris, W.H. (1984) Failed total hip replacement: Assessment by plain radiographs, arthrograms, and aspiration of the hip joint. *Journal of Bone & Joint Surgery - American Volume,* 66(4), 540-546.

Ong, A., Wong, K.L., Lai, M., Garino, J.P. and Steinberg, M.E. (2002) Early failure of precoated femoral components in primary total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 84(5), 786-792+Adv116.

Oral, E., Malhi, A.S., Wannomae, K.K. and Muratoglu, O.K. (2008) Highly cross-linked ultrahigh molecular weight polyethylene with improved fatigue resistance for total joint arthroplasty. *Journal of Arthroplasty,* 23(7), 1037-1044.

Orishimo, K.F., Claus, A.M., Sychterz, C.J. and Engh, C.A. (2003) Relationship between polyethylene wear and osteolysis in hips with a second-generation porous-coated cementless cup after seven years of follow-up. *Journal of Bone & Joint Surgery - American Volume,* 85(6), 1095-1099.

Ostendorf, M., van Stel, H.F., Buskens, E., Schrijvers, A.J.P., Marting, L.N., Verbout, A.J. and Dhert, W.J.A. (2004) Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. *Journal of Bone and Joint Surgery - British Volume,* 86(6), 801-808.

Pallant, J. (2007) *SPSS Survival manual*. 3rd ed. Maidenhead: Open University Press.

REFERENCES

Pandit, H., Glyn-Jones, S., McLardy-Smith, P., Gundle, R., Whitwell, D., Gibbons, C.L.M., Ostlere, S., Athanasou, N., Gill, H.S. and Murray, D.W. (2008) Pseudotumours associated with metal-on-metal hip resurfacings. *Journal of Bone and Joint Surgery - British Volume,* 90(7), 847-851.

Paprosky, W.G., Burnett, R.S.J. and Martin, E.L. (2001) Current topics in revision total hip replacement: Acetabular deficiency and its management. *Techniques in Orthopaedics,* 16(3), 227-236.

Paprosky, W.G., Weeden, S.H. and Bowling, J.W.,Jr. (2001) Component removal in revision total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 393, 181-193.

Parry, M.C., Whitehouse, M.R., Mehendale, S.A., Smith, L.K., Webb, J.C., Spencer, R.F. and Blom, A.W. (2010) A comparison of the validity and reliability of established bone stock loss classification systems and the proposal of a novel classification system. *Hip International,* 20(1), 50-55.

Peat, J.K., Mellis, C., Williams, K. and Xuan, W. (2002) *Health science research: A handbook of quantitative methods*. London: Sage Publications.

Perka, C., Fischer, U., Taylor, W.R. and Matziolis, G. (2004) Developmental hip dysplasia treated with total hip arthroplasty with a straight stem and a threaded cup. *Journal of Bone & Joint Surgery - American Volume,* 86(2), 312-319.

Petrie, A. (2006) Statistics in orthopaedic papers. *Journal of Bone and Joint Surgery - British Volume,* 88(9), 1121-1136.

Petsatodis, G.E., Papadopoulos, P.P., Papavasiliou, K.A., Hatzokos, I.G., Agathangelidis, F.G. and Christodoulou, A.G. (2010) Primary cementless total hip arthroplasty with an alumina ceramic-on-ceramic bearing: results after a minimum of twenty years of follow-up. *Journal of Bone & Joint Surgery - American Volume,* 92(3), 639-644.

Piper, K.J. and Paterson, A. (2009) Initial image interpretation of appendicular skeletal radiographs: a comparison between nurses and radiographers. *Radiography,* 15(1), 40-48.

Pollard, T.C.B., Baker, R.P., Eastaugh-Waring, S.J. and Bannister, G.C. (2006) Treatment of the young active patient with osteoarthritis of the hip: A five- to seven-year comparison of hybrid total hip arthroplasty and metal-on-metal resurfacing. *Journal of Bone and Joint Surgery - British Volume,* 88(5), 592-600.

Pospischill, M. and Knahr, K. (2005) Cementless total hip arthroplasty using a threaded cup and a rectangular tapered stem: follow-up for ten to 17 years. *Journal of Bone and Joint Surgery - British Volume,* 87(9), 1210-1215.

REFERENCES

Prevrhal, S., Klifa, C.S., Shepherd, J.A., Ries, M.D. and Genant, H.K. (2008) Computer-assisted technique to delineate osteolytic lesions around the femoral component in total hip arthroplasty. *Journal of Arthroplasty,* 23(6), 833-838.

Price, A. (ann.price@wwl.nhs.uk), (19 July 2010) Management of arthroplasty patients. Personal communication.

Pynsent, P.B. (2001) Choosing an outcome measure. *Journal of Bone and Joint Surgery - British Volume,* 83(6), 792-794.

Pynsent, P.B., Adams, D.J. and Disney, S.P. (2005) The Oxford hip and knee outcome questionnaires for arthroplasty: Outcomes and standards for surgical audit. *Journal of Bone and Joint Surgery - British Volume,* 87(2), 241-248.

Pynsent, P.B., Fairbank, J.C.T. and Carr, A.J. (2004) *Outcome measures in orthopaedics and orthopaedic trauma*. 2nd ed. London: Arnold.

Rajaratnam, S.S., Jack, C., Tavakkolizadeh, A., George, M.D., Fletcher, R.J., Hankins, M. and Shepperd, J.A. (2008) Long-term results of a hydroxyapatite-coated femoral component in total hip replacement: a 15- to 21-year follow-up study. *Journal of Bone & Joint Surgery - British Volume,* 90(1), 27-30.

Rankin, G. and Stokes, M. (1998) Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation,* 12(3), 187-199.

Rapp, S.M. (2011) National total joint registers foster quality care, patient satisfaction. *Orthopaedics Today Europe, January/February 2011*.

Revell, M.P., McBryde, C.W., Bhatnagar, S., Pynsent, P.B. and Treacy, R.B.C. (2006) Metal-on-metal hip resurfacing in osteonecrosis of the femoral head. *Journal of Bone & Joint Surgery - American Volume,* 88(3), 98-103.

Roder, C., Eggli, S., Aebi, M. and Busato, A. (2003) The validity of clinical examination in the diagnosis of loosening of components in total hip arthroplasty. *Journal of Bone and Joint Surgery - British Volume,* 85(1), 37-44.

Royal College of Nursing, Society and College of Radiographers, Chartered Society of Physiotherapy, NHS Alliance and Royal College of Radiologists. (2006) *Clinical imaging requests from non-medically qualified professionals.* London: Royal College of Nursing.

Ruston, S.A. (2008) Extended scope practitioners and clinical specialists: A place in rural health? *Australian Journal of Rural Health,* 16(3), 120-123.

REFERENCES

Saleh, K.J., Holtzman, J., Gafni, A., Saleh, L., Davis, A., Resig, S. and . Gross, A.E. (2001) Reliability and intraoperative validity of preoperative assessment of standardized plain radiographs in predicting bone loss at revision hip surgery. *Journal of Bone & Joint Surgery - American Volume,* 83(7), 1040-1046.

Santori, F.S. and Santori, N. (2010) Mid-term results of a custom-made short proximal loading femoral component. *Journal of Bone and Joint Surgery - British Volume,* 92(9), 1231-1237.

Schmalzried, T. and Harris, W. (1993) Hybrid total hip replacement. A 6.5-year follow-up study. *Journal of Bone and Joint Surgery - British Volume,* 75(4), 608-615.

Shaver, S.M., Brown, T.D., Hillis, S.L. and Callaghan, J.J. (1997) Digital edge-detection measurement of polyethylene wear after total hip arthroplasty. *Journal of Bone & Joint Surgery - American Volume,* 79(5), 690-700.

Shrout, P.E. and Fleiss, J.L. (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin,* 86(2), 420-428.

Sim, J. and Wright, C. (2000) *Research in health care: concepts, designs and methods*. Cheltenham: Nelson Thornes.

Skills for health (2009) *Rethinking roles and services* [online]. Bristol: Skills for Health. Available from: http://www.skillsforhealth.org.uk/ [Accessed 28 February 2011].

Smith, L.K., Langkamer, V.G., Shannon, M.N., Dixon, J.H., Case, R., Mahajan, A.J. and Spencer, R.F. (2008) Performance of the Cenator and EPF-Plus acetabular components at three years. *Journal of Bone and Joint Surgery - British Volume*, 90(Supplement 2), 310.

Smith, L.K., Cramp, F., Palmer, S., Coghill, N. and Spencer, R.F. (2010) Use of morphometry to quantify osteolysis after total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 468, 3077-3083.

Smith, L.K. (2011) Reply to letter to the Editor: Use of morphometry to quantify osteolysis after total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 469, 618.

Smith, T.O., Hedges, C., MacNair, R. and Schankat, K. (2009) Early rehabilitation following less invasive surgical stabilisation plate fixation for distal femoral fractures. *Physiotherapy,* 95(2), 61-75.

Soderman, P. and Malchau, H. (2001) Is the Harris hip score system useful to study the outcome of total hip replacement? *Clinical Orthopaedics & Related Research,* 384, 189-197.

Soderman, P., Malchau, H. and Herberts, P. (2001) Outcome of total hip replacement: A comparison of different measurement methods. *Clinical Orthopaedics & Related Research,* 390, 163-172.

REFERENCES

Spangehl, M.J., Berry, D.J., Trousdale, R.T. and Cabanela, M.E. (2001) Uncemented acetabular components with bulk femoral head autograft for acetabular reconstruction in developmental dysplasia of the hip: results at five to twelve years. *Journal of Bone & Joint Surgery - American Volume,* 83(10), 1484-1489.

Spangehl, M.J., Masri, B.A., O'Connell, J.X. and Duncan, C.P. (1999) Prospective analysis of preoperative and intraoperative investigations for the diagnosis of infection at the sites of two hundred and two revision total hip arthroplasties. *Journal of Bone & Joint Surgery - American Volume,* 81(5), 672-683.

Stamenkov, R., Howie, D., Taylor, J., Findlay, D., McGee, M., Kourlis, G., Carbone, A. and Burwell, M. (2003) Measurement of bone defects adjacent to acetabular components of hip replacement. *Clinical Orthopaedics & Related Research,* 412, 117-124.

Stratford, P.W. (2004) Getting more from the literature: Estimating the standard error of measurement from reliability studies. *Physiotherapy Canada,* 56, 27-30.

Streiner, D.L. and Norman, G.R. (2003) *Health Measurement Scales: a practical guide to their development and use*. 3rd ed. Oxford: Oxford University Press.

Stulberg, S.D., Wixson, R.L., Adams, A.D., Hendrix, R.W. and Bernfield, J.B. (2002) Monitoring pelvic osteolysis following total hip replacement surgery: an algorithm for surveillance. *Journal of Bone & Joint Surgery - American Volume,* 84(Supplement 2), 116-122.

Sugano, N., Nishii, T., Miki, H., Yoshikawa, H., Sato, Y. and Tamura, S. (2007) Mid-term results of cementless total hip replacement using a ceramic-on-ceramic bearing with and without computer navigation. *Journal of Bone and Joint Surgery - British Volume,* 89(4), 455-460.

Szende, A., Oppe, M. and Devlin, N. (eds) (2007) *EQ-5D Value sets: Inventory, comparative review and user guide.* Dordrecht, The Netherlands: Springer.

Tabachnick, B.G. and Fidell, L.S. (2001) Using multivariate statistics. 4th ed. Boston, USA: Allyn and Bacon.

Tegner, Y. and Lysholm, J. (1985) Rating systems in the evaluation of knee ligament injuries. *Clinical Orthopaedic & Related Research,* 198, 43-49.

The Chartered Society of Physiotherapy (2009) *The voice of physiotherapy* [Online]. London: The Chartered Society of Physiotherapy. Available from: http://www.csp.org.uk [Accessed 15 July 2010].

The Journal of Bone and Joint Surgery (American volume) (2010) *Instructions to authors* [online]. Available from: http://www.jbjs.org/ [Accessed 20 November 2010].

REFERENCES

The NHS Information Centre (2011) *HESonline: patient reported outcome measures - monthly summary* [online]. London: National Health Service. Available from: http://www.hesonline.nhs.uk/ [Accessed 10 February 2011].

The Royal College of Radiologists. (2006) *Standards for the reporting and interpretation of imaging investigations.* London: Royal College of Radiologists.

Thomas, A., Epstein, N.J., Stevens, K. and Goodman, S.B. (2007) Utility of judet oblique x-rays in preoperative assessment of acetabular periprosthetic osteolysis: a preliminary study. *The American Journal of Orthopedics,* 36(7), 107-110.

Thomas, J.R. and Nelson, J.K. (1990) *Research methods in physical activity*. 2nd ed. Champaign, Illinois, USA: Human Kinetics.

Toms, A.P., Botchu, R. and Nolan, J.F. (2009) Diagnostic plain film radiology of the failing hip replacement. *Orthopaedics and Trauma,* 23(2), 88-100.

Utting, M.R., Raghuvanshi, M., Amirfeyz, R., Blom, A.W., Learmonth, I.D. and Bannister, G.C. (2008) The Harris-Galante porous-coated, hemispherical, polyethylene-lined acetabular component in patients under 50 years of age: A 12- to 16-year review. *Journal of Bone and Joint Surgery - British Volume,* 90(11), 1422-1427.

Walton, M.J., Walton, J.C., Bell, M. and Scammell, B.E. (2008) The effectiveness of physiotherapist-led arthroplasty follow-up clinics. *Annals of the Royal College of Surgeons of England,* 90(2), 117-119.

Walton, N.P., Darrah, C., Shepstone, L., Donell, S.T. and Phillips, H. (2005) The Elite Plus total hip arthroplasty: The need for radiological surveillance. *Journal of Bone and Joint Surgery - British Volume,* 87(4), 458-462.

Wangen, H., Lereim, P., Holm, I., Gunderson, R. and Reikerås, O. (2008) Hip arthroplasty in patients younger than 30 years: excellent ten to 16-year follow-up results with a HA-coated stem. *International Orthopaedics,* 32(2), 203-208.

Ware, J.E.,Jr and Sherbourne, C.D. (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care,* 30(6), 473-483.

Ware, J.E.,Jr, Kosinski, M. and Keller, S.D. (1996) A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care,* 34(3), 220-233.

Warwick, R. and Williams, P.L. (eds) (1973) *Gray's Anatomy*. 35th ed. London: Longman

Weissman, J. (2003) *Presenting to win*. Upper Saddle River, NJ, USA: Financial Times Prentice Hall.

REFERENCES

Willert, H., Buchhorn, G.H., Fayyazi, A., Flury, R., Windler, M., Koster, G. and Lohmann, C.H. (2005) Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints: A clinical and histomorphological study. *Journal of Bone & Joint Surgery - American Volume,* 87(1), 28-36.

Williams, H.D., Browne, G., Gie, G.A., Ling, R.S., Timperley, A.J. and Wendover, N.A. (2002) The Exeter universal cemented femoral component at 8 to 12 years. A study of the first 325 hips. *Journal of Bone & Joint Surgery - British Volume,* 84(3), 324-334.

Wilson, P.D.,Jr, Amstutz, H.C., Czerniecki, A., Salvati, E.A. and Mendes, D.G. (1972) Total hip replacement with fixation by acrylic cement. A preliminary study of 100 consecutive McKee-Farrar prosthetic replacements. *Journal of Bone & Joint Surgery - American Volume,* 54(2), 207-236.

Wright, J.G. and Feinstein, A.R. (1992) Improving the reliability of orthopaedic measurements. *Journal of Bone & Joint Surgery - British Volume,* 74(2), 287-291.

Wroblewski, B.M., Siney, P.D., Fleming, P.A. and Bobak, P. (2000) The calcar femorale in cemented stem fixation in total hip arthroplasty. *Journal of Bone and Joint Surgery - British Volume,* 82(6), 842-845.

Wroblewski, B.M., Siney, P.D. and Fleming, P.A. (2002) Charnley low-frictional torque arthroplasty in patients under the age of 51 years. *Journal of Bone and Joint Surgery - British Volume,* 84(4), 540-543.

Wroblewski, B.M., Siney, P.D. and Fleming, P.A. (2007) Charnley low-friction arthroplasty: survival patterns to 38 years. *Journal of Bone and Joint Surgery - British Volume,* 89(8), 1015-1018.

Wroblewski, B.M., Siney, P.D. and Fleming, P.A. (2009) Charnley low-frictional torque arthroplasty: Follow-up for 30 to 40 years. *Journal of Bone and Joint Surgery - British Volume,* 91(4), 447-450.

Wylde, V.V., Learmonth, I.D. and Cavendish, V.J. (2005) The Oxford hip score: The patient's perspective. *Health Qual.Life.Outcomes, .*

Yates, P.J., Burston, B.J., Whitley, E. and Bannister, G.C. (2008) Collarless polished tapered stem: Clinical and radiological results at a minimum of ten years' follow-up. *Journal of Bone and Joint Surgery - British Volume,* 90(1), 16-22.

Zicat, B., Engh, C.A. and Gokcen, E. (1995) Patterns of osteolysis around total hip components inserted with and without cement. *Journal of Bone & Joint Surgery - American Volume,* 77(3), 432-9.

Appendices

List of appendices

Appendix I

Correspondence

Correspondence from abroad: Australia

From: ███████████████████████████████
Sent: Wednesday, 12 January 2011 21:28
To: Harding, Paula
Subject: Arthroplasty practitioners

Dear Paula,
I hope you do not mind me getting in touch but I am the current President of the Arthroplasty Care Practitioners Association in the UK.

I was aware of your visit to the UK and have read your subsequent report with interest. I would like to hear of any developments in your geographical area as a result of your work, and wonder if you would mind letting me know. I am currently in the final stages of a PhD on the subject of hip arthroplasty follow-up and feel strongly that we need to develop the international links as well as the national profile of this work.

I would also like to refer to your work in my thesis, and wondered if you would give permission for this? I will reference it appropriately so that the credit goes to you!

Finally, from you experiences in Canada, is there one person that you would particularly recommend for me to get in touch to develop the links?

Well done for an excellent piece of work and I look forwards to hearing from you in due course.

Kind regards,

Lindsay Smith

APPENDIX 1. CORRESPONDENCE

From: "Harding, Paula" ███████████████████
Date: 14 January 2011 03:58:48 GMT
To: ████████████████
Subject: RE: Arthroplasty practitioners

Hi Lindsay,

The arthroplasty role is very new to our public hospital system. My hospital - The Alfred was awarded a grant to implement the arthroplasty review clinic from the Victoria Government. it was a one off block of funding. Two other hospitals in Melbourne - St Vincents and The Royal Melbourne were also provided with a grant to trial clinics. We have been fortunate to have received funding from the hospital to continue. I have enclosed the final report I had to submit to the Victorian Government. it was one of 4 reports for the year so if you were interested in the other documents let me know and I can send it on.

Just before xmas I presented the results of the clinic to another metropolitan hospital in Melb who were interested in introducing such a role. I am not aware of any of the other states in Australia doing any of this work. I have just been appointed to a national committee for advancing scope of practice for physiotherapy and are due to meet for the first time in a few weeks - I should have the opportunity to find out more about what the other states are doing via this committee.

In regards to my report I am happy for you to include this in your thesis. The only thing to consider is the physios in Toronto Canada were reluctant to have all the information I included in the spreadsheet about them in the report published on the internet and asked if I could just reference their website.

Good luck with your PhD (I am still a few years away from completing mine)
Let me know if you would like any further info.

Kind Regards
Paula

Paula Harding
B. Physio, M.Manip
Grade 4 Musculoskeletal Stream Leader
Physiotherapy

██████ ████████████████████

The Alfred
55 Commercial Road
Melbourne 3000 Australia

Correspondence from abroad: Canada

From: Lindsay Smith
Sent: Wed 09/02/2011 5:27 AM
To: Robarts, Susan; Kennedy, Deborah
Subject: Advanced practice in arthroplasty

Dear Susan and Deborah,

I am a UK trained physiotherapist working as an orthopaedic researcher and as an arthroplasty practitioner. I have been given your names by Paula Harding, Australian physiotherapist.

I have looked at your work in developing the APP role in arthroplasty and find it very interesting. As you probably know, a number of hospitals have been using physiotherapists and orthopaedic nurses in similar roles for a number of years, and so we have formed an organisation to try and bring together those involved. It is called ACPA - Arthroplasty Care Practitioners Association (http://www.acpa-uk.net) and I am currently the President.

I would be very interested to know if you are still expanding the role in Canada and if you are training others to do this type of work? Is long term follow up of arthroplasty a continuing part of the role?

If you have a moment to answer these questions, I would very much appreciate any information you can supply. I am currently completing my doctoral thesis on the surveillance of hip arthroplasty and am about to present some of the work to the UK hip specialist surgeons at their annual conference.

Yours sincerely,

Lindsay Smith

APPENDIX 1. CORRESPONDENCE

From: "Robarts, Susan"
Date: 9 February 2011 14:15:50 GMT
To: "Lindsay Smith"
Subject: RE: Advanced practice in arthroplasty

Hi Lindsay!

It's great to hear from you! We met the past president, Morag, and her team, at the Edinburgh hospital when we came over for a site visit in 2006. We also visited the Glasgow Royal Infirmary: Helen Findlay and the arthroplasty team were so welcoming and a tremendous support to us as we developed the role.

We are very much still involved in long term follow up. The Advanced Practice Physiotherapists (and one Occupational Therapist) saw over 4000 patients for follow up last year. We also provide telephone support following discharge from hospital. We recently published an article about patient satisfaction in Physiotherapy Canada which was exciting to see in print. The Commentary was written by a leading Oncologist and wait time expert who is very supportive of these roles. We are trying to get papers into the published domain to help maintain the role despite the poor economy!

In addition, our Advanced Practice team triages all the referrals to the Centre (about 200 a month) and performs the initial comprehensive assessment to determine surgical candidates – about 56 assessments a week. A nurse and physiotherapy assistant contribute to this assessment.

We have expanded into other areas: the Shoulder Program, Spine, Fracture Clinic and Hip Fracture....all with lots of new challenges but having a template to work with, makes all the difference.

We have assisted many sites across Canada in developing the role and it is taking off!!

I'm very happy to be in touch with you! I am just about to see a patient but wanted to send a quick response! - Let's stay in touch!! I would love to hear about your PhD.

Susan

Susan Robarts MSc,BHScPT,BSc
Team Leader, Advanced Practice Physiotherapist
Holland Orthopaedic & Arthritic Centre
Sunnybrook Hospital, Toronto

Appendix II

Image interpretation

Direct Observation of Procedural Skills (DOPS):

Interpretation of plain radiographs in arthroplasty review

ASSESSOR
NAME: TITLE:

PRACTITIONER
NAME:

TYPE OF ARTHROPLASTY:

DATE:

	Unsatisfactory	**Satisfactory**	**Above expected**
1. **Comments on adequacy of film**			
2. **Verbal report**			
3. **Drawing up appropriate management plan**			

Strengths?

Areas for development?

Assessors signature …………………………………………………………...

Studies and reports from the literature to identify image interpretation by non-medical health professionals

Source	Title	Journal	Design	Level of evidence*
Benger et al, 2002	Can nurses working in remote units accurately request and interpret radiographs?	Emergency Medicine Journal	Diagnostic study (exploratory, cohort)	4
Hardy and Barrett, 2003	Requesting and interpreting trauma radiographs: a role extension for accident and emergency nurses	Accident and Emergency Nursing	Survey	N/A
Piper and Paterson, 2009	Initial image interpretation of appendicular skeletal radiographs: A comparison between nurses and radiographers	Radiography	Diagnostic study (exploratory, cohort)	4
Littlejohn et al, 2006	What are the protocols and procedures for imaging referral by physiotherapists?	New Zealand Journal of Physiotherapy	Survey	N/A
Daker-White et al, 1999	A randomised controlled trial. Shifting boundaries of doctors and physiotherapists in orthopaedic outpatient departments	Journal of Epidemiology and Community Health	Randomised controlled trial	2b
McPherson et al, 2006	A systematic review of evidence about extended roles for allied health professionals	Journal of Health Services Research and Policy	Systematic review	2a-
The Royal College of Radiologists, 2006	Standards for the reporting and interpretation of imaging investigations	Royal College of Radiologists, London	Report	N/A
Royal College of Nursing et al, 2006**	Clinical imaging requests from non-medically qualified professionals	Royal College of Nursing	Report	N/A

* Levels of evidence (March 2009), Centre for evidence based medicine (available at www.cebm.net)

** The report from the Royal College of Nursing was produced in association with representatives from the Society and College of Radiographers, the General Chiropractic Council, the General Osteopathic Council, the Chartered Society of Physiotherapy, the NHS Alliance and the Royal college of Radiologists

Protocol for requesting radiological investigations for arthroplasty clinics

Background

Arthroplasty clinics provide a routine review of all patients who have received a joint replacement at Weston General Hospital, as recommended by the British Orthopaedic Association[1]. This occurs at five year intervals or more frequently if symptoms change or if a new prosthesis is being monitored (recommended by NICE[2]). The assessment includes x-rays, completion of validated outcome questionnaires and a clinical examination. The patient benefits from the exclusion of pathology or the option of suitable intervention if adverse changes have occurred.

The clinics are operated by an orthopaedic researcher who is an advanced practice physiotherapist. There are no orthopaedic consultants present but they can be readily accessed in simultaneous clinics.

The x-rays are viewed with the patient but any information imparted indicates that a consultant orthopaedic opinion will be obtained. The images are subsequently reviewed in detail in the radiology department by the researcher and results discussed with an orthopaedic consultant. A radiologist report is not routinely requested.

In cases where deterioration of the arthroplasty is suspected, further investigations are ordered in consultation with, and in preparation for, an orthopaedic consultant. This may require an ultrasound scan, a CT scan, an isotope bone scan or an MRI scan.

Protocol

The practitioner will:

- Complete IR(ME)R training and updates as required by the Radiology Department and be familiar with current local clinical imaging protocols[3].
- Consider the needs of each patient before ordering x-rays appropriate to the arthroplasty (e.g. exclude patients with recent images of the same joint; exclude any woman with possibility of pregnancy).
- Provide clear clinical information and patient identification relevant to the investigation requested[4].
- Confirm the necessity for investigations other than x-ray with a senior orthopaedic colleague before requesting.
- Review all x-ray images of joint replacements and record findings with reference to the clinical signs and symptoms.
- Discuss the findings and management of the patient with an orthopaedic consultant.
- Recognise their own limitations and obtain an expert opinion when in doubt.
- In cases where unsuspected pathology is noted, inform the GP in accordance with RCR standards[4].
- Review any reports of additional investigations (e.g. bone scan) in conjunction with a specialist orthopaedic consultant.
- Provide evidence of continuing professional development in the interpretation of the images obtained.

---

[1] British Orthopaedic Association 2006. *Primary total hip replacement: A guide to good practice.*
[2] National Institute for Clinical Excellence 2000. *Guidance on the selection of prostheses for primary total hip replacement.* Technology Appraisal Guidance No.2
[3] Royal College of Nursing 2006. *Clinical imaging requests from non-medically qualified professionals.*
[4] The Royal College of Radiologists 2006. *Standards for the reporting and interpretation of imaging investigations.*

Case study: Unilateral hip replacement

What?

RH is a 69 year old male with a right total hip replacement (THR) completed in 2003.

At a routine five year review appointment in September 2008, this patient presented with a history of intermittent pain over the anterior of the right hip over the last three years. Pain appeared on commencing walking, then disappeared and then reappeared after 15 minutes. He had no back problems. He required analgesia for the pain.

On examination, he was tender over the greater trochanter and the anterior groin. No sacro-iliac joint or low back pain was produced by tests or movements.

The x-rays showed moderate polyethylene wear of 0.16mm/yr. There was a possible screw hole granuloma seen behind the acetabular cup. Radiolucencies were noted in Gruen zones 1, 7,8,14 but all were less than 2mm width and were at the cement-bone interface.

So what?

On reflection, LKS decided the possibilities were:

Differential diagnosis possibilities	Test	Action
Low grade infection	Inflammatory markers	Nil if normal
Wear debris granuloma and/or synovitis	Aspiration and injection of local anaesthetic	
Loosening of femoral component	X-ray analysis	Monitor
Greater trochanter bursitis	Steroid injection	Assess pain response
Unclear		GP to review pain medication; LKS to monitor

Low grade infection was unlikely as he had experienced a two year pain-free period following his joint replacement. However, in order to exclude it, a blood test for inflammatory markers was ordered. The results were all within normal parameters and so this diagnosis was dropped.
Femoral loosening was unlikely despite noting radiolucencies in Gruen zones as above. The radiolucencies were all less than 2mm in width and were not continuous around the prosthesis. They will be monitored for possible progression but no further action was necessary at this time. The wear debris synovitis and the bursitis remained as possibilities.

The patient was referred to a hip specialist and seen in November 2008. On examination, the patient had tenderness located over the greater trochanter. So, in view of normal blood markers and no major radiographic changes, it was decided to try a local anaesthetic and steroid injection to the trochanteric region. If this was unsuccessful in relieving the patients pain, further investigations for possible wear debris synovitis would be instigated.

On 16th December 2008, a phone call was made to the patient. He reported that his symptoms had decreased but were not abolished and that he now had low back pain. This suggested some element of bursitis but was not a complete explanation for his pain.

On 2nd July 2009, he attended the arthroplasty clinic for a further monitoring appointment. He reported that his hip symptoms were now minimal and did not cause him any problem. However, his low back pain had increased and was his main complaint. From his described history and clinical examination, his symptoms appeared to be from arthritis of sacro-iliac joints. He was given appropriate advice for managing this condition in conjunction with intermittent use of non-steroidal anti-inflammatory medicine prescribed by his GP. His hip replacement will be monitored for further changes.

What now (learning points)?

This is a complex scenario in which the presenting symptoms and the x-rays did not indicate the same diagnosis. It required diagnosis by exclusion and that involved some time. It was an excellent case to challenge my personal knowledge base in order to resolve the problems. As a result, my practice has changed as follows:

- My objective assessment in similar cases will more specifically examine the area of the greater trochanter and its overlying bursa.
- I will take into account that some time may be needed to make a diagnosis by exclusion
- I will continue to contrast and compare the signs on x-ray with the symptoms as described by the patient

Appendix III

Radiographic assessment tool

Paper copy of morphometric grid
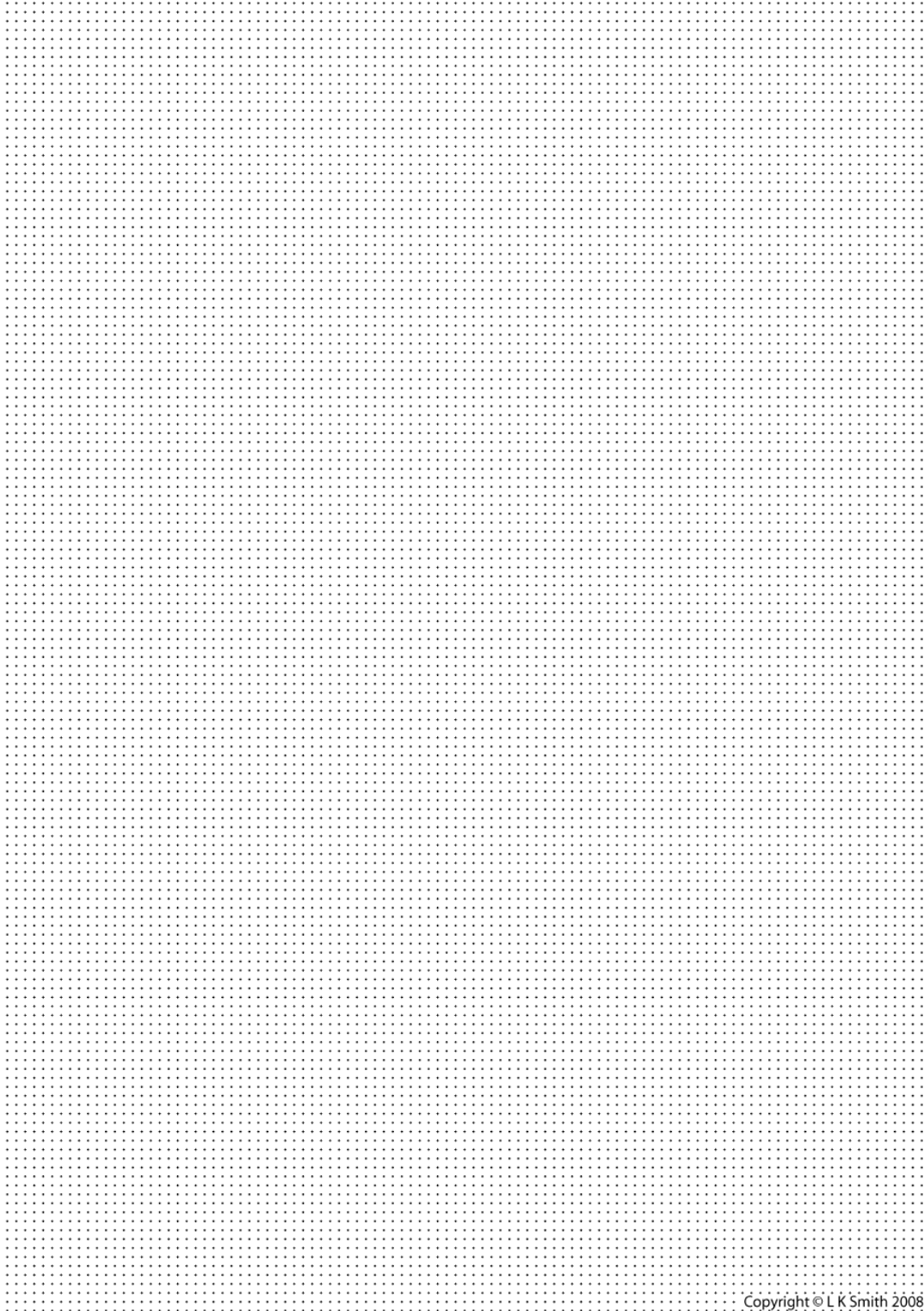
Formula to adjust morphometric grid for differences in magnification

If the morphometric grid is to be used on sequential plain radiographs of differing magnification, a simple formula can be used to adjust the number of points observed. For Film 1, the femoral head diameter = F1 and the size of the lesion on the grid = N1. For Film 2, the femoral head diameter = F2 and the size of the lesion on the grid = N2. The number of points for N1 can then be adjusted using the following formula: true number for N1 = (F2/F1)2 x N1. This true number for N1 can be compared with the number of points recorded for the other lesion (N2) to assess any change in size.

Appendix IV
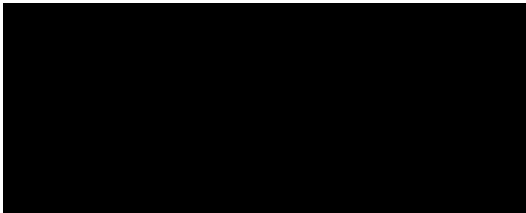
Research governance

National Research Ethics Service - Ethical application

Weston Area Health **NHS**
NHS Trust

24th April 2007

Dear Mr. Ashby,

Re. E342 – Steps to NICE compliance: the issue of the acetabulum

I am writing to ask for your advice on a proposed extension to an earlier study. This study was approved by the Weston Local Research Ethics Committee on 25th June 2003 and was completed in December 2005. Patients attended Weston General Hospital orthopaedic department for one appointment three years after their total hip replacement. At this appointment, an x-ray of the joint replacement was taken and the patient completed two short scores. The study has led to the establishment of a routine service which now follows up all patients who have had a joint replacement performed by one of the Weston orthopaedic consultants.

As a result of this new service, we will be recalling all the patients from the E.342 study for routine follow-up at seven or eight years after their hip replacement, approximately four years since their last visit. However, following their routine appointment, we would like to use the information obtained (scores and x-rays) to compare with results from the earlier study. This will form part of a PhD study registered with the University of the West of England. The project is entitled: Development of essential criteria for hip arthroplasty review.

The data set produced will be anonymised before it leaves Weston General Hospital and all statistical analysis will be conducted on anonymised data at the University of the West of England. The key to the data will be stored at Weston. No other organisations or commercial companies will be involved.

This study will not involve any extra visits to the hospital for the patients. All the relevant data will be collected as part of the routine follow-up of patients with joint replacements in weekly clinics. It is proposed that patients are informed in advance of clinic attendance about the further study of their results, and that written consent is requested at the clinic appointment. The results obtained will contribute to both local and national discussions on long term care of patients with hip replacements through presentation at orthopaedic conferences and through the Arthroplasty Care Practitioners Association (a network for health professionals involved in long term care of patients with joint replacements).

In view of the fact that the data needed for this study will be collected routinely in a follow-up clinic, I am writing to ask for your advice about ethical approval. There is no intervention or change in treatment and no new data being collected beyond routine attendance at the hospital. I would appreciate your guidance and advice in this matter.

Yours sincerely, Lindsay K. Smith MSc., M.C.S.P. Orthopaedic Research Assistant

NRES opinion on clinical study

National Research Ethics Service

North Somerset & South Bristol Research Ethics Committee

Assembly Rooms
UBHT Headquarters
Marlborough Street
Bristol BS1 3NU
Tel: 0117 928 2513
Email: naaz.nathoo@ubht.swest.nhs.uk
Facsimile: 0117 928 3724

14 May 2007

Lindsay K Smith
Orthopaedic Research Assistant
Orthopaedic Department
Weston General Hospital
Grange Road, Uphill
Weston-super-Mare
BS23 4TQ

Dear Lindsay,

Re: E342 –Steps to NICE compliance: the issue of the acetabulum

Thank you for your letter dated 24 April 2007, contents of which have been reviewed by a Sub-Committee of the NS&SB Research Ethics Committee.

It was felt that the proposed project entitled "Development of essential criteria for hip arthroplasty review" did not require formal ethics review, provided patients are informed about the proposed use of their (anonymised) results and appropriate consent is in place.

With best wishes,

Yours sincerely,

REC Coordinator

University of the West of England - Ethical approval

Our ref: SE/lt

13 December 2007

Faculty of Health and Social Care

BRISTOL

Lindsay K Smith
Highfield
85 West Hill
Portishead
Bristol   BS20 6LQ

Health Training and Research Centre

Frenchay Campus
Coldharbour Lane
Bristol BS16 1QY

Telephone 0117 958 5655
Facsimile 0117 32
Direct Line 0117 32

Dear Lindsay

Application number: HSC/07/12/119
Application title: Development of essential criteria for hip arthroplasty review

Your ethics application was considered by the Faculty Ethics Sub-Committee and based on the information provided was given ethical approval to proceed with the following conditions:

1. Please add UWE logo to the information sheet and consent form, given that you are carrying out this study as a UWE student.
2. Please state on the information sheet that this study is part of a PhD at UWE.
3. Item 11 of the information sheet should also state that the study has been reviewed by an ethics committee at the University of the West of England.
4. The information sheet needs to include details of a complaints mechanism for participants.
5. The committee suggests checking that participants are not deceased prior to sending out invitations to take part.
6. Please send a copy of the revised information sheet to Leigh Taylor.

If these conditions include providing further information please do not proceed with your research until you have full approval from the committee. You must notify the committee in advance if you wish to make any significant amendments to the original application.

Please note that all information sheets and consent forms should be on UWE headed paper.

If you have to terminate your research, please inform the Faculty Ethics Sub-Committee within 14 days, indicating the reasons for early termination.

-2-

Please be advised that as principal investigator you are responsible for the secure storage and destruction of data at the end of the specified period. A copy of the faculty data handling guidelines are enclosed for your information.

We wish you well with your research.

Yours sincerely

Chair
Faculty Ethics Sub-Committee

c.c. Fiona Cramp

Research & Development approval – Weston General Hospital

K Drive/R&D/R&D Approval
Approval let.

JA/VP

27ᵗʰ October 2008

Weston Area Health NHS

NHS Trust

Research & Development
Weston General Hospital
Grange Road, Uphill
Weston-super-Mare
Somerset
BS23 4TQ

Mrs. L. Smith
Associate Orthopaedic Researcher
Orthopaedic Dept
Weston General Hospital
Grange Road
Weston super Mare
BS23 4TQ.

Tel: 01934 881135 / Ext 5035
Fax: 01934 881139 / Ext 5039

Dear Lindsay,

RE: Project No. 2109 – Arthroplasty Review

I am please to tell you that the above project has been approved by Weston Area Health Trust (WAHT) and can now proceed.

It is essential that this project be carried out according to Good Clinical Practice and within the guidelines of the NHS Research Governance Framework for Health and Social Care and that an Investigator Site File is maintained. (Full information is available on http://www.dh.gov.uk/PolicyAndGuidance.) You have a responsibility for ensuring that all participants sign informed consent and that the protocol agreed by the Research Ethics Committee (REC) is adhered to by yourself and any co-workers. Please ensure that the version of the protocol and supporting documents held by the Research and Development Department is the same as the one approved by the REC and provide up-to-date versions if necessary. This letter should be kept in the site file for this project with other appropriate documentation.

If your study involves an intervention in the treatment of patients then you must ensure that any serious adverse events, regardless of whether you believe the event is related to the research or the intervention, are reported to the R&D Department as soon as possible using the standard WAHT incident form or alternatively, a photocopy of the Serious Adverse Events Form if your study is a commercially funded study. Note that you must also follow any SAE reporting requirements stipulated by the Sponsor.

Congratulations on initiating this research study. We wish you every success. We are very keen to support good research at WAHT and are pleased that you have decided to conduct your study here.

If you require any support please do not hesitate to contact myself or our R&D Co-ordinator on ▮▮▮▮▮ or by email▮▮▮▮▮

Yours sincerely,

▮▮▮▮▮

Consultant-Palliative Medicine
Associate Medical Director for R&D

Chairman: Chris Creswick   Chief Executive: Lorene Read

An Associated Teaching Hospital of the University of Bristol

214

National Research Ethics Service – Reliability study of morphometric grid

From: Lindsay Smith
Sent: 29 April 2009 14:31
To: NRES Queries Line
Subject: Advice please

Dear Sir/Madam,
I am a PhD student as well as an orthopaedic researcher based in the NHS and would like some advice about part of my PhD project.

My PhD research study involves observing the changes seen in patients who have had hip replacements approximately seven years previously. The patients are all attending routine follow-up clinics held in our department, regardless of any participation in a study, and there is no intervention or difference from other attendees in the way that they are screened. This involves a check x-ray, completion of validated questionnaires and a clinical assessment. The group of 200 whom I am in the process of observing was part of an earlier study with local NRES approval and for which they all gave informed consent. They are giving informed consent again for use of their results in this extension to the study.

I have received ethical approval from my university ethics committee (University of the West of England) and a letter from the local NRES stating that ethical approval is not required for this extension to the original study (see attached). However, some of my supervisors are raising the issue of additional ethical approval for a sub-section of my PhD.

As part of my PhD, I have developed a tool to use on x-rays to measure some of the changes that are seen. It is a simple grid which is superimposed on existing x-rays to count off the size of an area of interest. The development of the tool involved a statistical reliability study using some of the x-rays. The following points summarize relevant considerations:

- The x-rays used were already in existence so no time or money was spent on obtaining them and they were not required for any other purpose
- The retrieval of the x-rays did not involve any NHS staff as I did it myself
- The facilities used to view the x-rays were not patient facilities so no patient was compromised
- The viewing of the x-rays by myself and three colleagues was all done in our own time – no NHS time was used.
- All x-rays were used anonymously so patient privacy was protected
- No patient treatment or management was changed in any way as a result of taking the measurements – it was purely a quantitative exercise
- All results were completely anonymous, just figures on a piece of paper

The results from this statistical study support the use of the grid in further measurements of the x-rays of the same cohort. It is part of the pathway in the PhD study.

I feel caught between opinions and would really appreciate your advice on the necessity for further ethical approval or not.

Yours faithfully, Lindsay Smith

APPENDIX IV. RESEARCH GOVERNANCE

From: NRES Queries Line [mailto:queries@nres.npsa.nhs.uk]
Sent: 29 April 2009 15:06
To: █████████████████████
Subject: RE: Advice please

Your query was reviewed by our Queries Line Advisers.

Thank you for your email seeking advice on [whether your project requires REC review] [determining whether your project should be classified as research requiring REC review or as some other type of activity such as audit or service evaluation].

Review by a Research Ethics Committee is required only for research that falls within paragraph 3.1 of the Department of Health's Governance Arrangements for Research Ethics Committees (available   at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4005727). Legislation such as the Clinical Trials Regulations, Human Tissue Act and Mental Capacity Act also require ethical approval from an appropriately recognised REC even where the research is taking place outside of the NHS.

NRES has produced a leaflet on "Defining Research", which will help you to determine whether or not your project is research or audit or service evaluation.

If your project will be taking place within the NHS, your local R&D office will be able to advise on whether the project should be classified as research and requires management within the Research Governance Framework for Health and Social Care, including ethical review by a REC. The R&D office can also advise on local governance procedures for other types of project such as audit or service evaluation. More detailed guidance on categorising projects is also available on the website of the NHS R&D Forum at http://www.rdforum.nhs.uk/docs/categorising_projects_guidance.doc.

So that we might further consider your query, please email an A4 summary (**one side only as a Word document**) outlining your proposal to Queries Line. For ease of reference please include your request in the covering email.

Queries Line
National Research Ethics Service
National Patient Safety Agency
4-8 Maple Street
London
W1T 5HD

Website: www.nres.npsa.nhs.uk
Email: queries@nres.npsa.nhs.uk

Ref: 04/02

APPENDIX IV. RESEARCH GOVERNANCE

From: Lindsay Smith ███████████████████████
Sent: 29 April 2009 16:32
To: NRES Queries Line
Subject: RE: Advice please

Thank you for this reply.

My summary is already included below but to clarify my question further, it is:

I have all the required approvals (NRES, university and R&D) for my PhD study. Do I need further ethical approval for a subsection of my PhD study which involves a reliability study of a tool which is superimposed on x-rays?

Lindsay Smith

Your query was reviewed by our Queries Line Advisers.

If you have a favourable opinion from your local REC for your study, I don't think further action is needed for this reliability assessment.

I hope this helps.

Regards

Queries Line
National Research Ethics Service
National Patient Safety Agency
4-8 Maple Steet
London
W1T 5HD

Website: www.nres.npsa.nhs.uk
Email: queries@nres.npsa.nhs.uk

Ref: 04/02

Appendix V

Clinical study documentation

Participant invitation letter

WESTON AREA HEALTH NHS TRUST
ORTHOPAEDIC DEPARTMENT
GRANGE ROAD, UPHILL
WESTON SUPER MARE
SOMERSET
BS23 4TQ

TEL: 01934 636363
DL: 01934 647206
Fax: 01934 647219

Dear   (Date)

We are writing to inform you about the results of the research clinic on hip replacements which you attended a few years ago.

First of all, we would like to offer our thanks for your participation when you attended the Orthopaedic Department at Weston General Hospital for an x-ray and some questionnaires. When we looked at the results from over 200 people who had received hip replacements at Weston, we found that the types of implant that we had used were performing well. We are sure that you will be glad to know this!

The other positive outcome of our research is that we are now able to offer routine follow up to all patients who have had a hip or knee or shoulder replacement with one of our local orthopaedic surgeons. Everyone is invited to attend after an appropriate number of years for x-rays and questionnaires. This allows you, the patient, to check your progress and allows us to monitor any changes in your joint replacement.

As a result, we have included an appointment for you to have a further check up on your hip. It is now over five years since you attended the clinics and we would like to assess if there have been any changes. We would do this in the same way as before with an x-ray and some questionnaires which tell us how you are getting on. We will then write to your GP with details of your progress.

In addition, we would like to use your results to compare with your previous results. This will help us to develop some standards for measuring hip replacements. This is very important as so many people now have hip replacements and some of them will need a second replacement later in their life. The information we collect will be used at both local and national level. It will, however, all be anonymous once it leaves this hospital.

Letter of invitation to participants continued…

If you consent to being included in this second part to the study, you will benefit by having your hip closely checked by specialists in the subject. Also, it will contribute to the work in this area across the British Isles. There are no risks to you from being involved. The study has been approved by the local ethics committee and is being organised by the same people as before – Mr. R.F. Spencer, Consultant Orthopaedic Surgeon, and Lindsay Smith, Research Associate.

As previously stated, we have included an appointment for you at the Orthopaedic Department in Weston General Hospital. If the time or date is inconvenient, please feel free to phone (01934 647206) and change it to suit you. You will be asked when you attend the clinic if you agree to us using your results as described.

Thank you, again, for your help with the initial study, and we look forward to seeing you in clinic.

Yours sincerely,

Lindsay K. Smith MSc, MCSP
Research Associate in Orthopaedics.

Patient information

University of the West of England BRISTOL

Weston Area Health NHS NHS Trust

DEVELOPMENT OF ESSENTIAL CRITERIA FOR ARTHROPLASTY REVIEW

INFORMATION FOR PATIENTS ……………………….October 2007. Version 3.

1. **Study Title**
   This study could also be called 'Developing standards for measuring hip replacements'.

2. **What is the purpose of the study?**
   To find some simple methods that can be used in orthopaedic clinics to measure the progress of a hip replacement.

3. **Why have I been chosen?**
   Because you participated in the earlier study of hip replacements and so we already have some useful information about your progress.

4. **Who is organising the study?**
   The same people as before – Mr. R. F. Spencer, Consultant Orthopaedic Surgeon, and Lindsay Smith, Orthopaedic Research Associate.

5. **What will happen to me if I take part?**
   After you have attended the orthopaedic clinic for a check-up on your hip replacement, we will use the information you give us to compare with your previous results. However, you are free to withdraw from the study at any time (by phoning or writing to us) and we will then remove your results from our analysis.

6. **Are there any disadvantages in taking part in this study?**
   No, there are none.

7. **What are the possible risks of taking part?**
   There are no risks to you from this study.

8. **What are the possible benefits of taking part?**
   Your hip joint and its performance will be closely assessed by specialists.

Patient information continued…

9. Is my doctor being paid for including me in this study?
No.

10. Confidentiality – who will know I am taking part in this study?
All information which is collected about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will be anonymised so that you cannot be recognised from it.

11. Ethical approval
This study has been approved by the North Somerset and South Bristol Research Ethics Committee and by an ethics committee at the University of the West of England.

12. What will happen to the results of this study?
This study is part of a PhD at the University of the West of England. The information will be published as a research study, will be presented at conferences and will be used in discussions at national level about future follow up of hip replacements.

13. Contact for further information

Lindsay Smith, Orthopaedic Research Associate, Weston General Hospital, Grange Road, Uphill, Weston-s-Mare, BS23 4TQ.

Phone: 01934 647206

If you should require an independent opinion, please contact:

Lesley Pattenden, Senior Physiotherapist, Physiotherapy Department, Weston General Hospital

Phone: 01934 647131

If you should have any concerns or complaint, please contact the **Patient Advocacy and Liaison Service** in the main Outpatient Department of the hospital.

Phone: 01934 647216

Thank you for being willing to take part in this extension of the earlier study with which you were involved.

Consent form for clinical study

University of the
West of England
BRISTOL

Weston Area Health NHS
NHS Trust

Patient Information Number for this trial:

CONSENT FORM

Title of Project: 'Development of essential criteria for hip arthroplasty review'

Name of Researchers: Robert F. Spencer. Consultant Orthopaedic Surgeon.

Lindsay K. Smith.   Orthopaedic Research Associate.

1. I confirm that I have read and understand the information sheet dated 'October 2007. Version 2' for the above study ……………………………………………………… ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without my medical care or legal rights being affected ………………………………   ☐

3. I am willing to allow access to my medical records but understand that strict confidentiality will be maintained. The purpose of this is to check that the study is being carried out correctly.   …………………………………………………………………………….. ☐

4. I agree to take part in the above study. …………………………………………………..   .☐

……………………………………………………………………………………………………..
Name of patient   Date   Signature

……………………………………………………………………………………………………..
Name of person taking consent   Date   Signature

Clinical record form for radiographic review

UWE **BRISTOL** University of the West of England

Weston Area Health **NHS**
NHS Trust

RADIOGRAPHIC ASSESSMENT – HIP

MRN……………..Initials ………….Side ……… Date …………………Yrs. post op. ……

KEY: **R = Radiolucency, PR = Progressive radiolucency, O = Osteolysis, E = Erosion,**
G = granuloma, CMD = Cement mantle deficiency, SS = Stress shielding,
CH = Cortical hypertrophy, RL = Reactive Line, PD = Pedestal formation, SW =
Spot welding, A=Atrophic, SHG = Screw hole granuloma

<u>**CUP**</u>   Comments:

Inclination (degrees, horizontal)	< 40 / 40 – 50 / > 50	**Wear** (AP, sup and inf, mm)	
		Migration (from teardrop)	

ZONES	I	IIA	IIB	III	IV	V	VI	Screws

<u>**STEM**</u>   Comments:

Inclination (AP view; degrees, Varus/valgus)	Neutral / VL <4 / VL >4 / VR <4 / VR >4	**Calcar resorption**	
Subsidence (AP; mm)	None / <1.5 / 1.5-3.0 / >3.0	**Calcar rounding**	

AP ZONES	1	2	3	4	5	6	7

Lateral ZONES	8	9	10	11	12	13	14

<u>**Heterotopic ossification**</u> (Brooker I - IV):

<u>**X-RAY OPINION:**</u>   **POOR**   **MOD. CHANGES**   **SLIGHT CHANGES**   **GOOD**

Appendix VI

Results

Data analysis

Oxford Hip Score: construction of the variable

Examination of the change in Oxford Hip Score (OHS) showed that the data had a mean value of 1.07 (95% CI; -0.01 to 2.15; range -13.0 to 23.0) but a median of zero. The trimmed mean was 0.72, indicating that outliers were affecting the mean. All outliers were checked for data accuracy and no discrepancies were found.

The Kolmogorov-Smirnov statistic was significant (0.18, df 154, $p < 0.001$) suggesting non-normality of the data although this is not unusual in larger samples (Pallant 2007). The skewness was 0.96 (standard error 0.195) but on visual examination of the histogram, it was reasonably normal which allayed fears about the outliers (see Figure A.7). In large samples, data can be accepted as normally distributed based on the visual appearance as deviation due to skewness has little effect on analysis (Tabachnick et al. 2001).

A decision was made to retain the outliers as they were not unattached from the other data and the shape of distribution approximated normal. The OHS change score data was considered acceptable to use as the dependent variable for the statistical analysis.

A.1. Bland-Altman plots for inter-observer reliability of morphometric grid on x-ray images
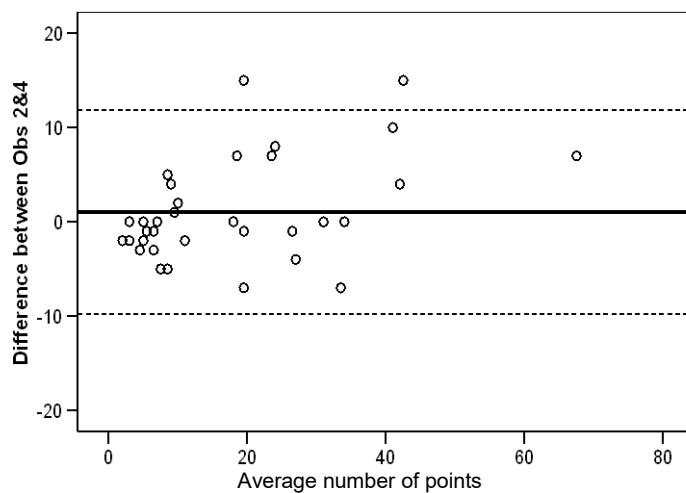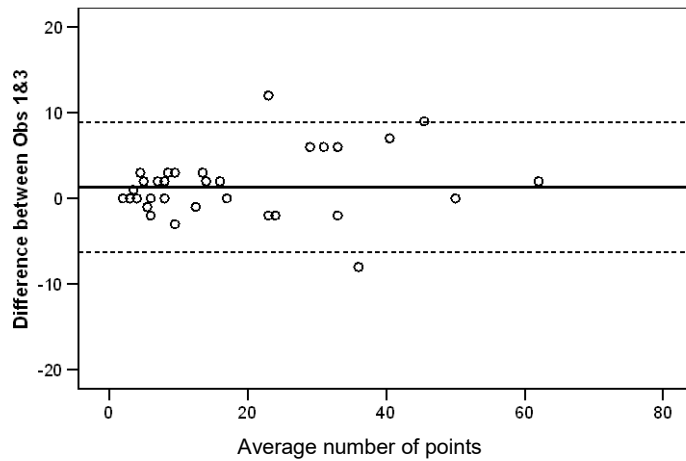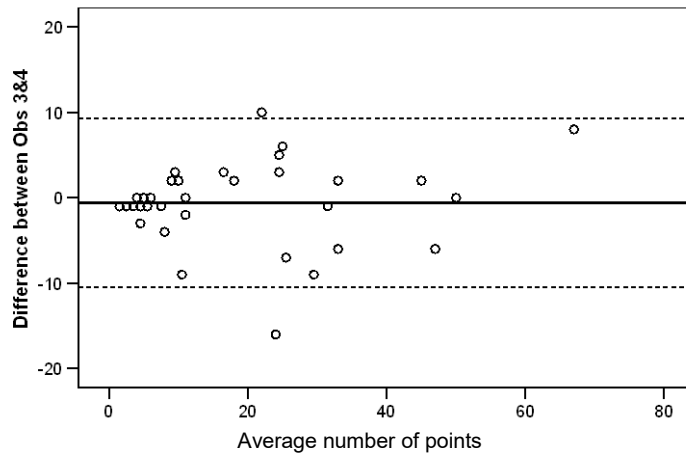
Bland-Altman distribution plots for inter-observer reliability of the morphometric grid when used to obtain measurements of osteolytic lesions on 35 radiographs. The mean of the differences between each pair of observers is shown as a solid black line; 95% limits of agreement are shown as dotted lines. The observer pairs are indicated on the y-axis.

A.2. Bland-Altman plots for test-retest reliability of morphometric grid on x-ray images
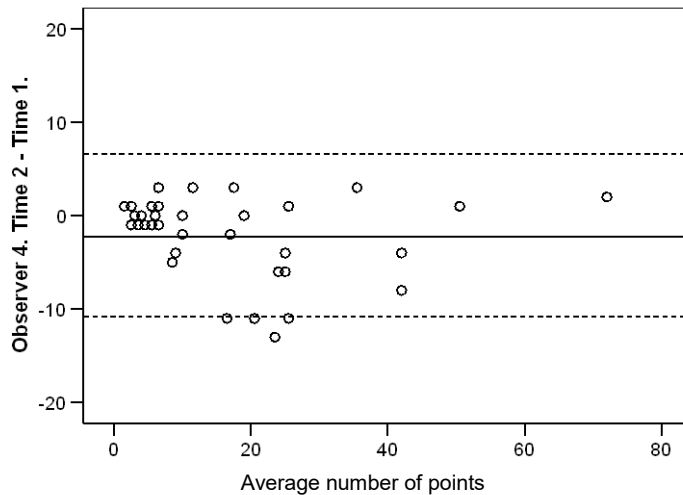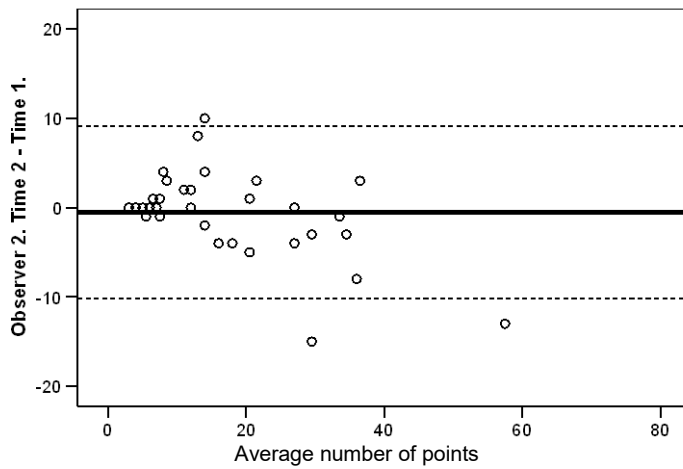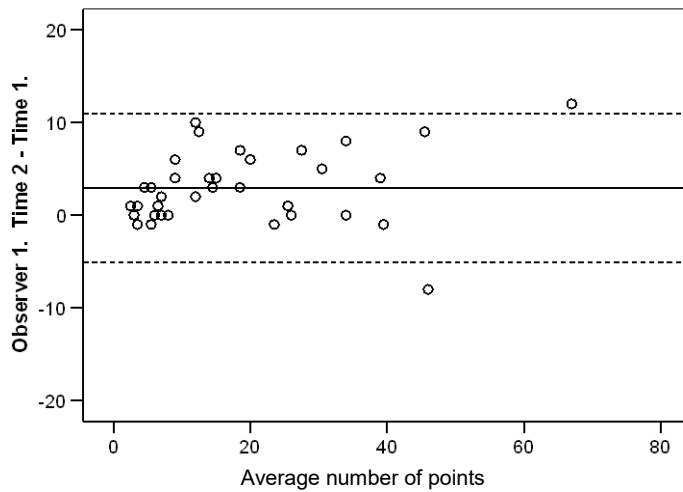
Bland-Altman distribution plots for test-retest reliability of the morphometric grid when used to obtain measurements of osteolytic lesions on 35 radiographs (observers one, two and four). The mean of the differences between each set of measurements is shown as a solid black line; the 95% limits of agreement are shown as dotted lines. The observer is indicated on the y-axis.

A.3. Bland-Altman plot for interobserver reliability of wear rate measurements
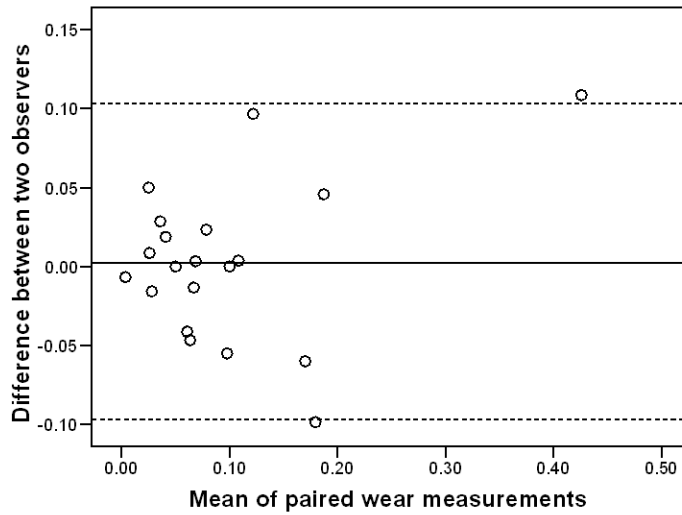
The mean of the differences between each set of measurements is shown as a solid black line; the 95% limits of agreement are shown as dotted lines.

A.4. Bland-Altman plot for test retest reliability of osteolysis measurements
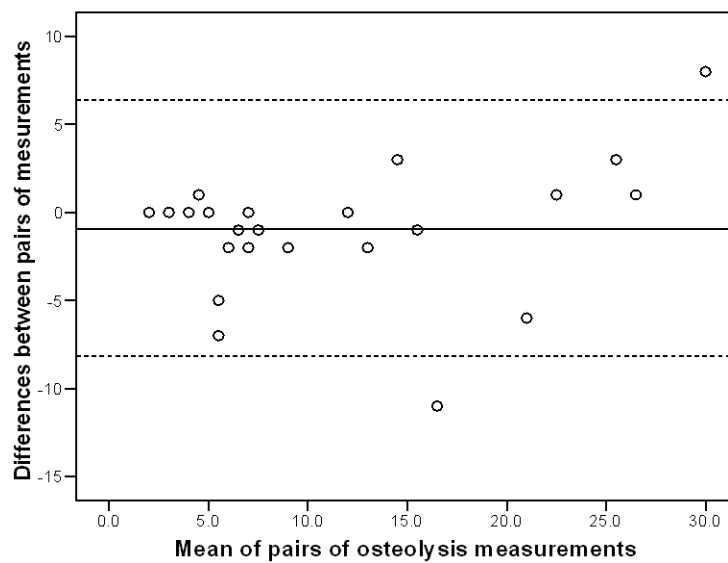
The mean of the differences between each set of measurements is shown as a solid black line; the 95% limits of agreement are shown as dotted lines.

A.5. Scatter plot of OHS change score against OHS score at three years
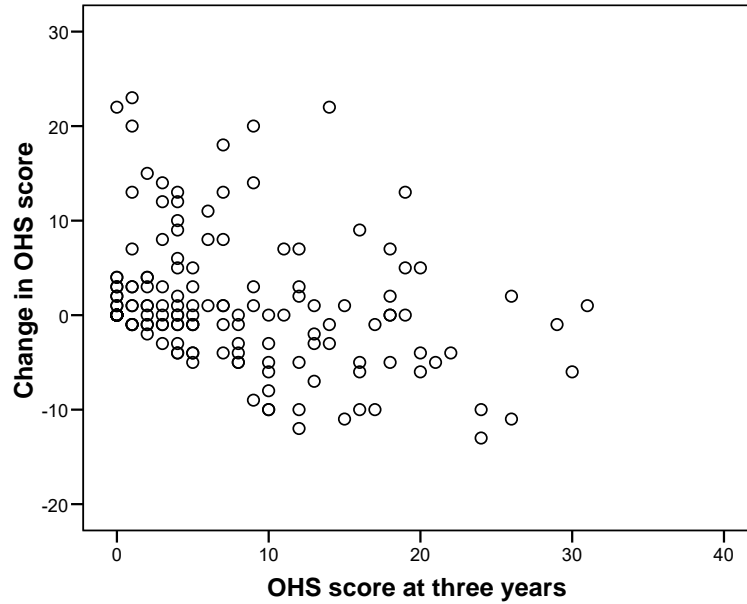
A.6. Scatter plot of percentage change in OHS against OHS score at three years
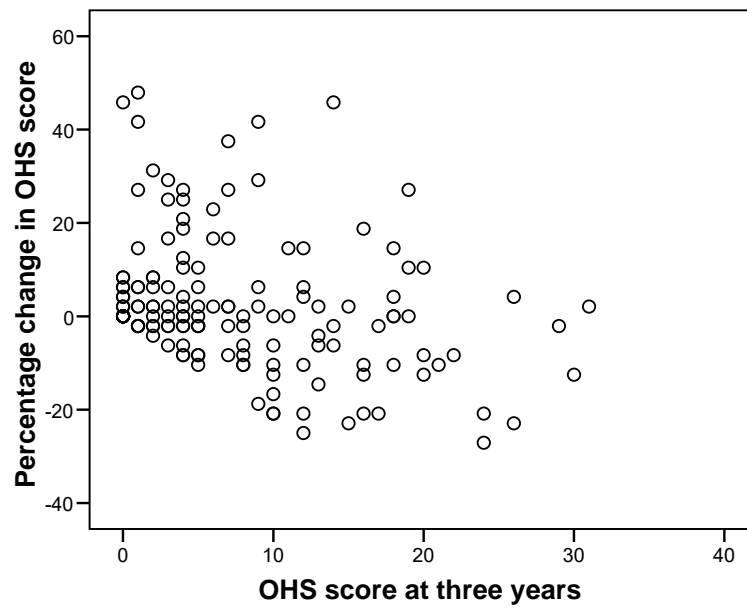
A.7. Histogram of OHS change scores
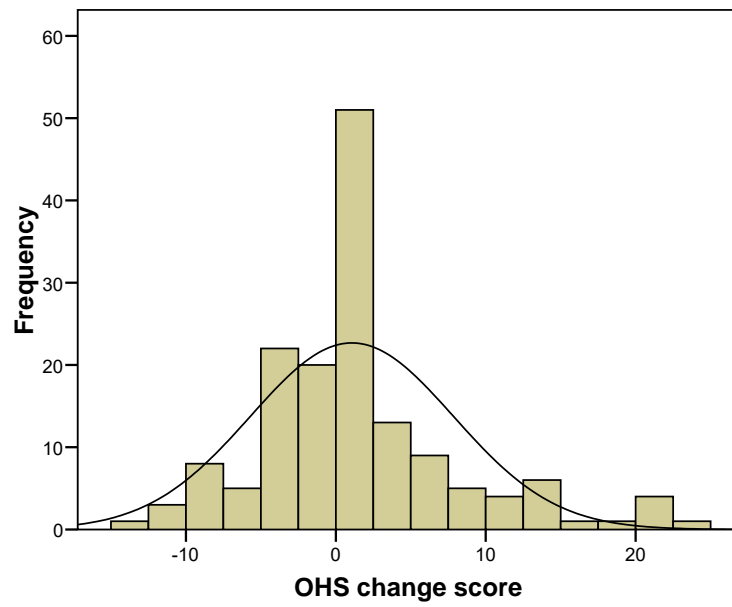
A.8. Histogram of ages of participants
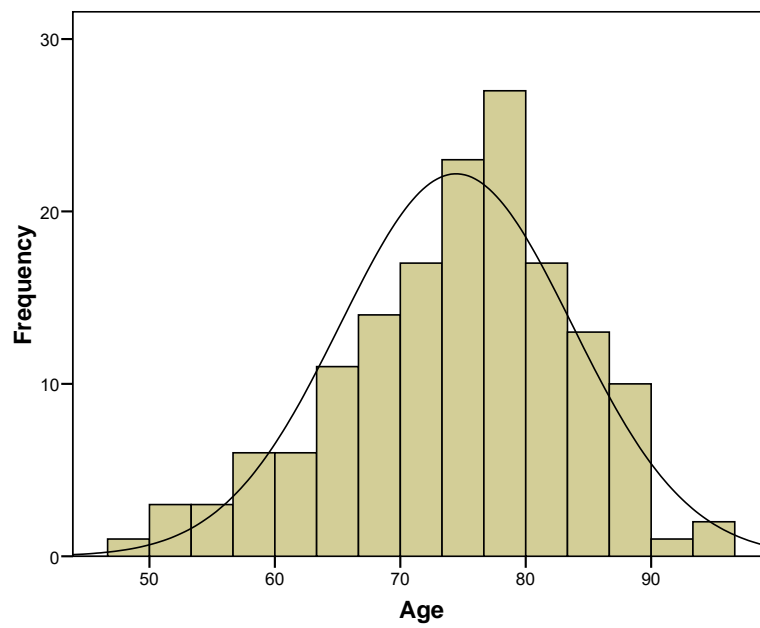
A.9. Histogram of EQ-5D scores
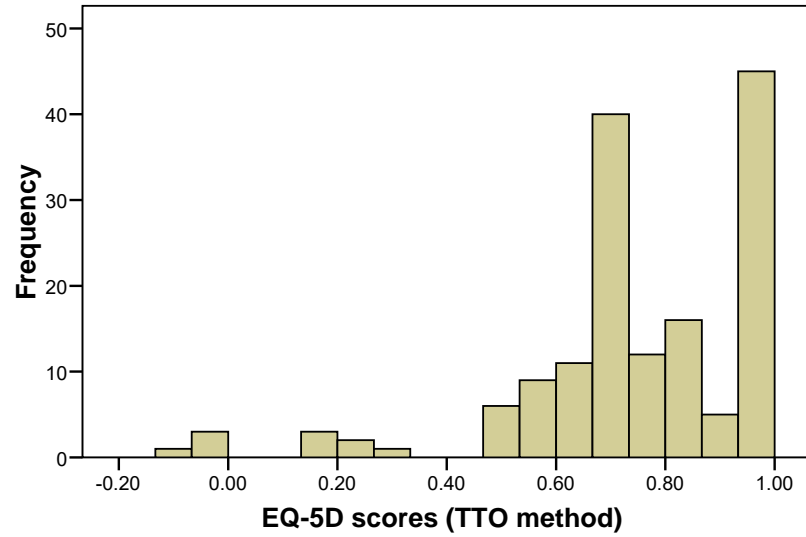
A.10. Histogram of Charlson comorbidity scores
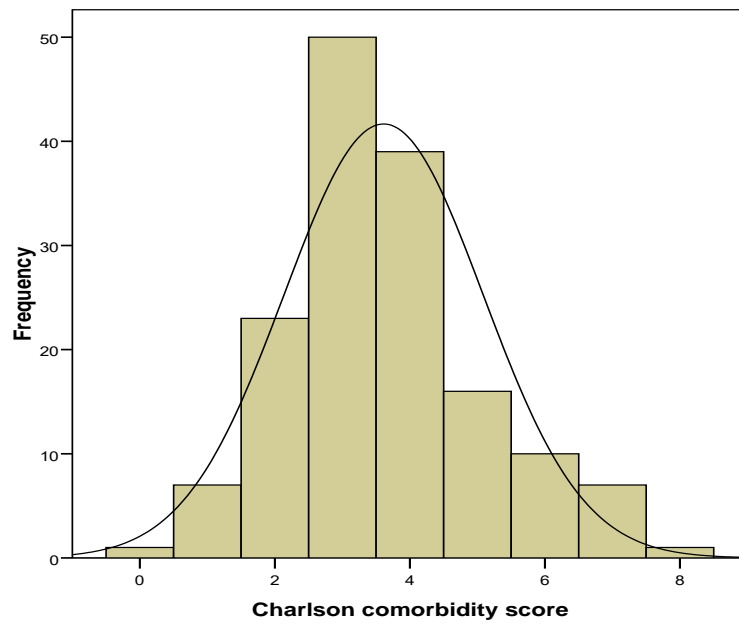
A.11. Histogram of the number of radiographic changes
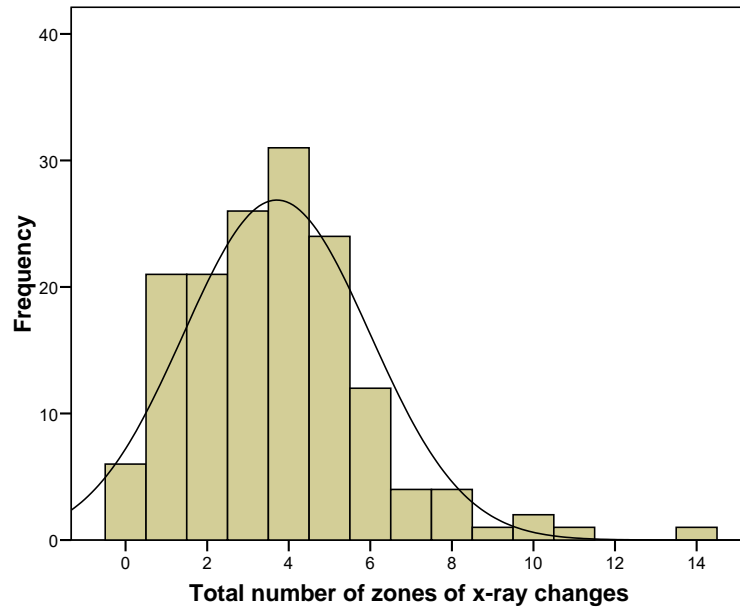
A.12. Histogram of the steady state wear rates of the polyethylene
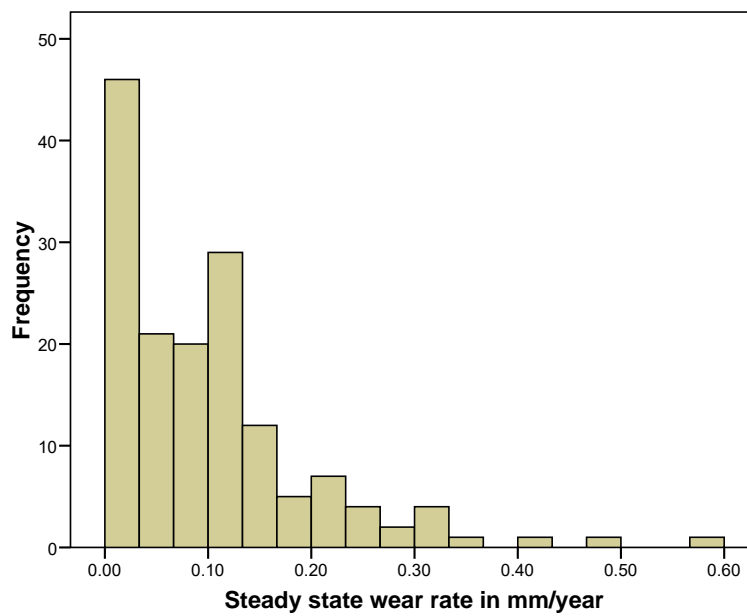
A.13. Box plot of radiographic changes for each decade of age
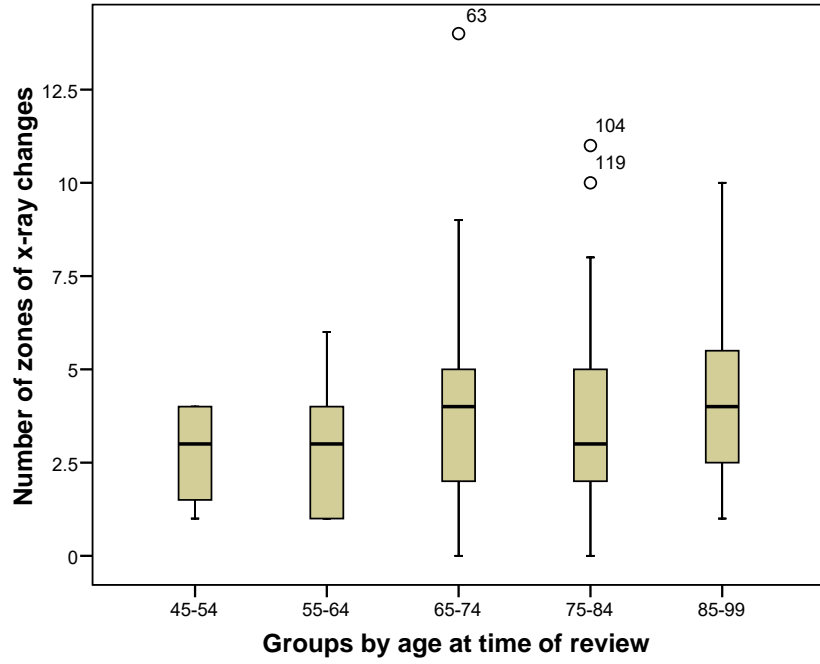
Appendix VII

Results of literature review: failing THA

APPENDIX VII. RESULTS OF LITERATURE REVIEW

Key for literature review

#	Fracture
ALVAL	Aseptic lymphocytic vasculitis associated lesion
AP	Antero-posterior view
CDH	Congenital dysplasia of the hip
CoCr	Cobalt chrome
CRP	C-reactive protein test
DDH	Developmental dysplasia of the hip
DEXA	Dual energy x-ray absorptiometry
EBRA	Ein-bild-roentgen-analyse method of measuring change in THA component position
EQ-5D	EuroQol 5 dimension questionnaire
ESR	Erythrocyte sedimentation rate
FU	Follow up
HHS	Harris hip score
LLD	Leg length discrepancy
MoM	Metal on metal articulation
n/a	not available
OHS	Oxford Hip Score
PE	Polyethylene
pt	Patient
RHR	Revision hip replacement
RL	Radiolucencies
SF12	Short form 12 questionnaire
SF36	Short form 36 questionnaire
THA	Total hip arthroplasty
UCLA	University of California in Los Angeles activity score
UHMWPE	Ultra high molecular weight polyethylene
VAS	Visual analogue scale
VL	Valgus
VR	Varus
WOMAC	Western Ontario and McMaster University Osteoarthritis index

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Altenburg et al 2009, USA	Non-randomized controlled cohort	3	15-18yrs av	130	56 yrs cable group, 59yrs wire group	151 hybrid	15% cable, 9% wire	54% cable group, 30% wire group	Johnston questionnaire	Osteolysis = any non linear RL at bone-cement interface ≥ 5mm; wear - edge detection Shaver 1997; migration of acetabular component; RL around acetabulum (definite = # cement or migration, probable = 100% RL, possible = 50-99%RL loosening);	Revision of femoral or acetabular component; radiographic loosening of either from RL or osteolysis.	Third body wear from cable accelerated PE wear of cup	44% and 26% acetabular osteolysis; 70% and 48% proximal femoral osteolysis; 54% and 30% loosening at 10 yrs. Recommendation to avoid use of CoCr cable near bearing surface.
Amin et al 2006, Scotland	Systematic review	4	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	High BMI may influence symptoms of hip OA rather than radiological signs; may get lower function after THA; no long term impact known.
Archibeck et al 2001, USA	Follow-up study	3	8-11yrs	85	52 yrs	92 uncemented	2.0%	5pts required liner and head exchange for osteolysis or wear;	HHS, physical examination; thigh pain;	A-P and lateral; RL categorized by width (,1mm, 1-2mm, >2mm); subsidence;VR/VL;calcar rounding, CH, pedestal; heterotopic bone; osteolysis by location and small/large (small = <1cm, large= >1cm, combined = 2 or more; peripheral, retroacetabular or ischial);	Distal osteolysis in stem threatening stability; wear debris seen at imtermediate FU causing osteolysis.	osteolysis; (recurrent dislocation, trauma);	0.16mm/yr wear; osteolysis 16% acetabular; wear and osteolysis associated (0.015 cp with 0.27mm/yr); RL and pedestal not bad if not progressive; good HHS but acetabular wear and osteolysis remain problem for potential loosening; less distal osteolysis in femur than other cementless stems; one pathologic fracture due to osteolysis in greater trochanter at 12 years.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Bennett et al 2008, Northern Ireland	Cross-sectional study	4	10 yrs	100	60.3 yrs	100 cemented	0.0%	0.0%	Pedometer for 2 weeks	Wear using Griffiths 1978	n/a	n/a	Age not related to wear despite decreasing activity with age; 0.14mm/yr wear rate.
Berry et al 2002, USA	Follow-up study	3	25 yrs	1689	63.5 yrs	2000 Charnley cemented	19.0%	n/a	questionnaire on further surgery	n/a	Re-operation	aseptic loosening 13% of cohort; 6% other reasons;	Acetabular aseptic loosening increased in later period; femoral aseptic loosening was linear. Osteoarthritic patients at more risk of aseptic loosening than rheumatoid; younger age at surgery and male sex increased risk of aseptic loosening. General health better than average for first 20 yrs post THA.
Biring et al 2007, Canada	Follow-up study	3	1.8 yrs	235	65 yrs	235 revision THA	n/a	n/a	WOMAC, SF12, UCLA	n/a	n/a	aseptic loose 84%;	Predictors of better outcomes: better WOMAC function pre-op, 60-70yrs old, male, lower Charnley class. SF12 poorly correlated with post op WOMAC.
Bizot et al 2004, France	Follow-up study	3	9 yrs	62	46 yrs	71 hybrid ceramic on ceramic	6.0%	1 pt with symptom-free osteolysis, monitored	Merle d'Aubigne and Postel	RL lines and osteolysis by zones; migration of components.	Re-operation	infection, pain, # alumina head, aseptic on one cup;(4 pts)	Good mid-term results in younger pts as osteolysis only in one pt with acetabular trauma; socket remained well-fixed although some RL seen; pts with osteolysis had good score.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Buckwalter et al 2006, USA	Follow-up study	3	25 yrs	49	56.2 yrs	52 Charnley cemented (339 originally)	10.6% femoral components	17% loosening of femoral component	WOMAC at final follow up and other questionnaire	Loosening, osteolysis and wear (modified Shaver et al); femoral cementing by Schmalzried and Harris.	Revision and radiographic loosening	Aseptic loosening, recurrent dislocation, infection	Wear mean 0.134 mm/yr (0.213 mm/yr in revised hips and 0.098 in stable hips). Osteolysis significantly associated with wear.
Burns & Bourne 2006, Canada	Opinion	5	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Cost of RHR increasing annually as no. of operations increases; reduction of infection and dislocation are important to reduce need for RHR; improvement of materials and designs also important but must be evidence-based; reimbursement for RHR is often insufficient if bone graft and costly implants are required.
Callaghan et al 2004, USA	Follow-up study	3	30 yrs	262 (27 alive)	54 yrs	330 Charnley cemented	12% of original cohort; 32% of final cohort revised after 30 yr for aseptic loose Dx on x-ray:	16.4% total; 45% cups and 19% stems in final cohort:	WOMAC	Loosening -RL lines and osteolysis by zones; wear by Shaver 1997;	Revision and radiographic loosening	aseptic loosening 8.6% (cups at 13.4 yrs> stems at 15 yrs); infection, dislocation;	68% Charnleys still intact at 30yrs; aseptic loosening main reason for failure and revision. Linear wear 0.1mm/yr.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Callaghan et al 2009, USA	Follow-up study	3	35 yrs	12	51 yrs	15 Charnley cemented (330 originally)	(7 hips) 27% single revision, 13% two revisions, Girdlestone in 7%	9 hips	WOMAC at final follow up and other questionnaire	Loosening, osteolysis and wear (modified Shaver et al); femoral cementing by Schmalzried and Harris;	Revision and radiographic loosening	Aseptic loosening and sepsis.	78% survival at 35 yrs - benchmark for THA. Increasing age and decreasing activity and implant durability need to be considered.
Capello et al 2003, USA	Follow-up study	3	10 yrs	97	<50 yrs	111 uncemented	4.5% stems; 30% cups or liners;	47% with osteolysis in 1,7,8,14;	HHS, pain, walking aids	Stem stability by Engh 1990; RL <1mm, 1-2mm >2mm in Gruen zones; calcar resorption, stem subsidence (>2mm), cortical hypertrophy, cancellous condensation and osteolysis ; VR or VL stem change >2 degrees.	Revision	Osteolysis, excessive polyethylene wear, aseptic loosening;	Stem successful but not cup.
Chougle et al 2006, UK	Follow-up study	3	average 15.7 yrs (2 to 31 yrs)	206	43yrs	292 cemented for DDH	9% at 10yrs, 37% at 20 yrs	n/a	n/a	A-P x-ray; Crowe and Hartofilakidis dysplasia; Hodgkinson for loosening and migration; wear by Charnley and Halley 1975.	Acetabular revision	Aseptic loosening	DDH categorisation affects survival of acetabular component; younger age, increased wear, offset bore cup and previous acetabular surgery increase risk.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Collis & Mohler 2002, USA	Follow-up study	3	5-6yrs		68 yrs	244 hybrid (122 rough stems, 122 polished stems)	3% rough stems at 7 yrs	5% of rough stems at 7 yrs	Iowa score	Barrack cement grading, wear, RL, osteolysis (x by x mm description)	Revision or impending revision	Aseptic loosening and pain with substantial bone lysis with scalloping and erosive destruction and thinning of cortex	Rough stem stimulated lysis from bone cement disease.
Corner et al 2008, UK	Follow-up study	4	5 yrs +	38	56 yrs	41 uncemented (Thrust plate prosthesis)	7.0%	n/a	HHS, OHS,	A-P and lateral x-rays; Loosening and osteolysis	Revision and radiographic loosening	infection (1), # (1), aseptic loosening (1)	Thrust plate prosthesis is useful for selected young patients
de Jong et al 2004, The Netherlands	Follow-up study	3	10.5 yrs, 15 yrs	253	68.4 yrs	293 cemented (Weber prosthesis)	10% (30)	95/236 hips loose radiographically	HHS	Dall et al for acetabular and femoral loosening	Revision and radiographic loosening	Aseptic loosening (24 hips), infection, fracture.	Poor sensitivity of HHS to radiographic loosening. 88.7% survival at 10 yrs, 67.8% at 15 yrs (includes radiographic loosening). Stem was worse than socket in this group. Second generation cementing techniques improved survival.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Della Valle et al 2002, USA	Follow-up study	3	14.2 yrs	304	53 yrs	384 uncement-ed stems	0.6%	9.5% acetabular components	Merle d'Aubigne and Postel	RL, osseointegration by Engh 1990. Osteolysis by Zicat definition. Heterotopic ossification (Brooker). Resorptive remodelling of cortical bone (Engh 1987).	Revision or proposed revision of stem for aseptic loosening. Revision cup.	Instability of stem. Cup - aseptic loosening, osteolysis, wear or shell damage.	Need immediate stability for osseointegration to occur. Extensively coated stems with distal diaphyseal fit allow this or a stable fibrous fixation by means of dense fibrous tissue ingrowth if not. Proximally fitting implant relies on good cortical contact immediately. Stress shielding greater in extensively coated femoral components but does not affect clinical results.
Della Valle et al 2009, USA	Follow-up study	3	20 yrs	111	52 yrs	124 uncemen-ed	8.0%	16.0%	HHS (excellent 90-100, good 80-89, fair 70-79, poor <70).	Osteolysis by Archibeck et al (small= <1cm, large= >1cm, combined= 2 or more; peripheral, retroacetabular or ischial), Martell 1993 for x-rays.	Revision for / radiographic evidence of aseptic loosening	20% with osteolysis.	Reoperation rate for polyethylene wear and osteolysis has increased with time. Wear related complications significantly associated with younger age (<65 yrs) at time of surgery. Recommend cementing highly cross-linked liner into shell for liner replacement.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Dobzyniak et al 2006, USA	Follow-up study	3	n/a	692	n/a	745 revision THA	39% within 5 yrs		n/a	n/a	Revision	33% instability, 30% aseptic loosening, 14% infection, 55 osteolysis, 15% failed hemiarthroplasties, 3% peri-prosthetic fractures	High level of aseptic failure decreased over time, possibly as surgeons became more familiar with use of uncemented stems in which most of these failures had occurred.
Elke et al 2003, Switzerland	Follow-up study	3	7.4 yrs average	119	n/a	123 revision THA	7.3% (stems)	13% but no cups for aseptic loosening	Merle d'Aubigne and Postel	RL and osteolysis, cup angle (>5 degrees), cup anteversion (Dorr et al 2000), cup migration (Sutherland et al 1982) (>3mm), Heterotopic (Brooker et al 1973)	Cup with complete RL around it.	dislocation, aseptic loosening of stem, infection, peri-prosthetic #	Some migration but acceptable as part of press-fit concept and does not threaten long term stability. Screws more of a threat with potential for osteolysis.
Emms et al 2010, UK	Follow-up study	3	7.6 to 18 yrs, 11.6 yrs average	250	54 yrs	280 hybrid	22.0%	17% of surviving shells with osteolysis	n/a	A-P and lateral; osteolysis and RL lines in zones; wear by Wilkinson's own method using EBRA;	Hodgkinson acetabular loosening; Harris femoral loosening;	osteolysis>aseptic loosening>wear	Mid-term survival is a poor guide to long-term survival (same as Swedish and Finnish registers). Suggest dose-response relationship between wear and osteolysis rather than a wear threshold; not connected to age, gender, diagnosis, femoral head size or material, liner thickness.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Eskelinen et al 2006, Finland	Follow-up study	3	12.3 yrs	56	54 yrs	68 uncemented (CDH)	12% cup revised.	41.0%	HHS at each visit; abductor muscle strength; Trendelenburg test; satisfaction score.	A-P and frog-leg lateral: RL > 1mm by zone; osteolytic lesions by zone and size; alignment. Femoral - bone-ingrown, fibrous stable, unstable (Engh et al). Heterotopic ossification (Brooker). Liner thickness with mm paper and magnifying glass. Nonunion of greater trochanter > 15mm gap.	Femoral subsidence> 3mm. Acetabular cup - change in alignment > 2° or position ≥ 2mm. Excessive wear. RL in all three zones of cup.	aseptic loosening, excessive liner wear	Design of cup is of more importance in long term survival for this condition than wear and osteolysis.
Eswaramoorthy et al 2008, UK	Follow-up study	3	10.8 yrs	82	61.6 yrs (44-84)	85 cemented and hybrid (MoM)	7.1%	n/a	OHS	RL by zones if >50% of zone; migration of cup; heterotopic ossification (Brooker);	Migration >3mm, RL lines >2mm significant;	one cup for aseptic loosening; 1infection, 4 for pain (2 had ALVAL).	Failure rate comparable with metal on polyethylene THA; calcar loss like osteolysis but propose it is part of ALVAL -type response; aetiology of RL unknown; recommend follow-up.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Furnes et al 2001, Norway	Registry study	2	10 yrs	n/a	n/a	53698 cemented and uncement-ed prostheses	n/a	n/a	n/a	n/a	Revision	Aseptic loosening of stem or cup, dislocation, infection, fracture of femur, pain. Inferior uncemented prostheses defined as smooth press-fit stems, threaded cups with smooth surface, porous-coated stems without circumferential coating.	CDH and previous hip # are risk factors for worse outcome after THA. Hybrids at 7-10 years have increasing revision rate due to wear and osteolysis. Charnley cemented were showing more stem than cup revisions at 10 yrs (possibly improved techniques have reduced the incidence of loosening but brought about a shift from cup to stem loosening). Young patients in all disease groups had good results with a cemented implant.
Gonzalez Della Valle et al 2004, USA	Follow-up study	3	4-7 yr	297	64 yrs	335 hybrid	2.4%	n/a	Hospital for Special Surgery hip score (pain, walking ability, motion and function, 0-40)	Cup inclination, anteversion, migration, wear (Livermore technique), osteolysis, RL lines, calcar resorption, calcar osteolysis, heterotopic ossification (Brooker).	Well-fixed = no migration or RL. Fibrous fixation= circumferential RL≤1mm, no migration. Loose= progressive RL and migration.	Aseptic loosening >infection.	Second generation uncemented acetabular component. Age and gender significantly associated with wear rate (0.09 mm/yr average, 0.11 for men, 0.07 for women). 5% prevalence of osteolysis at this stage - may have been affected by lack of cups with holes. Good or excellent clinical outcome.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Grubl et al 2002, Austria	Follow-up study	3	10 yrs	133	61yrs (22-84)	139 uncement-ed	9.0%	12.0%	HHS, pain, satisfaction,	RL by zones, bone density (stress shielding), cortical hypertrophy, heterotopic ossification, Engh femoral implant stability, osteolysis as defined by Zicat et al, wear by Charnley & Halley 1975	Revision for any reason	75% of failures were the cup at mean of 8 yrs; aseptic loosening and wear of cup were most frequent reasons.	No osteolysis and a wear rate of 0.1mm/yr. Stem satisfactory but cup not so. RL in zones 1,7,8 and 13 not associated with instability.
Hartofilakidi s et al 2008, Greece	Historically controlled study	4	16.6 (cemented group) and 13.4 (hybrid group) yrs	67	50 yrs (25-69) cemented gorup, 47 yrs (27-70) hybrid group	46 cemented and 47 hybrid	32.6% cemented cup; 27.7% of uncemented cups (but only 8.5% for aseptic loosening)	n/a	Merle d'Aubigne and Postel	Wear (Livermore); osteolysis (linear = RL space > 1mm; expansile (Zicat definition) in zones, measured by length and width).	Revision of acetabular component or liner exchange. Acetabular loosening - migration > 2mm and/or continuous RL line > 1mm at cement-bone or implant-bone surface.	Aseptic loosening, wear of liner, peri-acetabular osteolysis. Cemented at 13.7 yrs, uncemented cups between 10 and 14 years on average.	Lytic lesions were mostly linear in cemented cups (35%), expansile for uncemented cups (15%). Higher rate of aseptic loosening in cemented cups related to diagnosis and offset cup. Uncemented cups failure was mostly due to wear and osteolysis.
Howard PW 2009, UK	Opinion	5	n/a	n/a	n/a	Revision THA	n/a	n/a	n/a	n/a	Aseptic loosening with/without bone loss and/or osteolysis - most common.		Reconstruction is aimed at implant stability and is compromised by bone loss. Pre-operative investigations do not indicate bone quality, only presence of bone.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG- ERY	NO PTS	AV. AGE AT SURG- ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Huddleston et al 2010, USA	Non-randomized controlled cohort	3	10 yrs	80	57yrs and 60 yrs for two groups	80 uncemented, 8 hybrid (45 Hylamer liners, 43 conventional PE liners)	16% of hylamer group, 2% of conventional PE group	16% included impending revision	n/a	AP and lateral x-ray; cup abduction angle, osteolysis by length and width of lesion; wear by Livermore technique (modified as only one x-ray used).	Expansile pelvic osteolysis	Affect of particle size on biological reactivity from Hylamer;	Linear wear rates similar (0.21mm/yr in Hylamer, 0.20mm/yr in conventional) but more pelvic osteolysis in Hylamer group. Suggest that asymptomatic patients with expanding lesions should be advised to undergo surgery. Suggest that asymptomatic patients with low linear wear rate and no osteolysis can have longer follow-up intervals.
Ihle et al 2008 Germany	Follow-up study	3	19.3 yrs	80	52 yrs	93 uncemented	15.0%	1.0%	Merle d'Aubigne and Postel pre-op. HHS at follow up. (Used scores to review changes and between groups)	Acetabular migration (Massin et al); RL and osteolysis by zone (Engh criteria). Heterotopic ossification (Brooker).	Cup migration >3mm or tilted >8 degrees or RL >2mm circumferential.	Revision of cup for isolated acetabular osteolysis >aseptic loosening > infection.	THA needs to be followed up. Uncemented system successful in younger patients if loosening is considered but failure due to osteolysis from wear debris increased between 13-15 years post surgery. Loosening was linear in comparison to osteolysis which was expansile.
Iida et al 2000, Japan	Follow-up study	3	12.3 yrs	112	53 yrs (33-72)	133 cemented for DDH	7.5%	23.0%	Merle d'Aubigne and Postel; leg length	RL lines, migration, inclination. Acetabular loosening (Hodgkinson et al), Femoral loosening (Harris et al)	Revision or radiologically loose	Aseptic loosening	Bone grafting and socket placement improve outcome in DDH.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Incavo et al 2008, USA	Follow-up study	3	6.7 yrs	91	57 yrs (19-84)	105 uncement-ed	6.0%	4% acetabular cups (in addition to revisions)	HHS, OHS, LLD >7mm.	Cup loosening RL in all three zones >2mm. RL by zones, osteolysis (>5mm in length), spot welds.	Revision	Aseptic loosening of cup - possibly due to impingement. Osteolysis in pelvis and greater trochanter	Review has prompted attention to impingement and materials used.
Ito et al 2004, Japan	Follow-up study	3	8.6 yrs	49	44 yrs	57 cementless	9.0%	19.0%	HHS, thigh pain on initial walking	Biologic fixation (Engh et al), bone ingrowth, fibrous ingrowth and unstable (subsidence >3mm or change in alignment). Cup loosening if RL >2mm and circumferential or inclination change > 5 degrees or migration >2mm. Osteolysis (defined) by zone and graded small < 2cm or large >2 cm in any direction.	Loosening of stem or revision	Lack of ingrowth of bone indicated radiologically or osteolysis and aseptic loosening	Wear of PE and osteolysis were a problem in femur. Lack of intact proximal implant-bone interface and out-of-roundness of head may have contributed to this.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Ito et al 2005, Japan	Follow-up study	3	11 yrs	111	61 yrs	128 hybrid	5.0%	zero	HHS (pre-op tp present change)	Cementation (Mulroy et al), osteolysis (focal lesions or RL.>2mm width), wear (Livermore)	Revision surgery. Aseptic loosening of stem -(Harris et al 1982). Acetabular migration > 4mm.	1 infection, 1 cup aseptic loose, 2 cups dislocation, 2 cups liner displacement	Wear rate of 0.09mm/yr. Asymptomatic osteolysis present. Cement mantle needs to be good but other issues affect aseptic loosening. Multifactorial influences on survival - materials, design, operative technique, patient population.
Jana et al 2001, USA	Follow-up study	3	11.4 yrs	55 RA	55 yrs (24 to 80)	71 cementless stems	21% (14 cups, 1 stem)	37% (4 stems, 7 cups in addition to revisions)	Merle d'Aubigne	Engh et al for uncemented stem; osteolysis by multiplication of two measurements, using zones.	Cup loosening > 5 degrees tilting, migration > 2mm, circumferential RL line, bead shedding from porous cups. Femur loose by Engh classification.	Stem - infection; cups, aseptic loosening (1 dislocation).	Femora had osteolytic lesions but not greater than 1.5 sq cm. Acetabular expansile lesions in 14% of unrevised hips. Pain and walking improved on average.
Keener et al 2003, USA	Follow-up study	3	25 yrs	45	< 50 yrs	60 Charnley cemented	37.0%	13% femoral, 34% acetabular components	HHS, WOMAC	Debonding by Berry et al (any RL in zone 1 of acetabulum between prosthesis and cement); wear by Shaver et al; heterotopic ossification, femoral cement grade, osteolysis, loosening of both components.	Revision and radiographic loosening	infection, aseptic loosening, dislocation, femoral fracture	No significant difference in functional evaluation between retained, revised and loose hips. Presence of comorbid conditions had effect on HHS and WOMAC. Wear of acetabular component significantly associated with acetabular loosening and revision.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Keisu et al 2001, USA	Follow-up study	3	5 (2-11)yrs	86	80+ yrs	92 uncement-ed	zero	zero	Merle d'Aubigne and Postel (Charnley modification); HHS in 69/92 but only pre-op to post-op	Bone ingrowth around both components; wear of acetabular liner.	Revision and radiographic loosening	None revised	Wear of liner in 41% with some osteolysis (4%). Femoral stems had osseintegrated. Uncemented THA is safe in elderly.
Kim 2005, Korea	Non-randomized controlled cohort	3	7.1 yrs	52	44.2 yrs	104 bilateral uncement-ed (one zirconia head and one cobalt-chromium femoral head in each patient)	2.0%	2.0%	HHS, Tegner activity scale	Adequacy of intramedullary fill; axial subsidence > 3mm or varus or valgus shift of stem; anteversion and abduction of acetabular component; radiolucencies; areas of osteolysis by length, width and zone; proximal femoral bone resorption by Engh 1987; differentiation of cause of osteolysis in calcar by bone biopsy; linear wear of polyethylene liner using software to produce image.	Stem loosening if progressive axial subsidence > 3mm or varus or valgus shift. Possible loosening if complete radiolucent line around prosthesis. Acetabular loosening if >2mm shift or complete radiolucent line >2mm.	Aseptic loosening of undersized femoral stems	Wear rate 0.08mm/yr in zirconia group, 0.17mm/yr in cobalt-chromium group. Increased wear rate associated with younger age, male, increased activity level and increased abduction angle. Not able to detect difference between types of hip with scores - compared means of each group . No osteolysis present.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Kim 2008, Korea	Follow-up study	3	8.8 yrs	471	52.7 yrs	601 uncemented	0.3% (deep infection), 0.2% (recurrent dislocation)	zero	HHS, Tegner activity scale	Supine standardised A-P, Lateral. Femoral subsidence, valgus/varus shift, RL lines, radio dense lines. Acetabular anteversion, inclination, migration, RL >2mm.Osteolysis (any discreet localised RL not on previous radiographs) by zones. Ingrowth by direct contact of trabecular striation between bone and component. Proximal femoral bone resorption by Engh et al 1987. Brooker heterotopic bone.	Femoral loosening when subsidence >3mm, axial shift> 3 degrees, complete RL line on AP and lateral views. Acetabular loosening when >2mm change vertical or medial/lateral, RL >2mm on AP and lateral views.		Metaphyseal fit without distal stem in contact with endosteal surface in combination with ceramic-on-ceramic bearing (no wear particle disease) produces good results.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Kim et al 2002, Korea	Follow-up study	3	9.4 yrs (8-10)	55	43.4 yrs	64 hybrid (22mm ceramic femoral head)	2.0%	0.2%	HHS at each FU; VAS pain rating in thigh.	AP and lateral x-ray (frog-leg/cross-table/Judet views). Femur - stem alignment, cement grading (Barrack et al), calcar rounding and resorption, cortical hypertrophy, subsidence, osteolysis (length and width of areas, both scalloped and linear >2mm width) by zone. Centre of rotation, limb length, femoral neck length, femoral offset, abductor moment arm. Heterotopic ossification (Brooker). Acetabulum - inclination, migration >3mm vertically or horizontally, osteolysis by zone, radiolucencies. Linear wear by distance between femoral head centre and cup centre with baseline at 6 weeks post op.	Acetabular loosening - migration >3mm or continuous RL line >2mm on AP and lateral views. Stem - definite loosening if subsidence of stem or mantle, bending or breakage of stem or crack in cement mantle. Debonding if RL at stem-cement interface.	Infection	RL in zone 1and 7 in 11%. 9% osteolysis in zone 7. No acetabular osteolysis. Bedding-in by 3 years. 0.1mm/yr linear wear.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Kim et al 2003, Korea	Non-randomized controlled cohort	3	9.3 yrs	98	47-48 yrs	50 bilateral (one hybrid, one uncement-ed), 48 unilateral uncement-ed	3.0%	zero	HHS, VAS for thigh pain,	Stem - alignment, adequacy of intramedullary fill or cement mantle grade (Barrack et al 1992), radiolucent lines. Classified as bone ingrown, stable fibrous ingrown or unstable. Acetabular cup - radiolucent lines, inclination and change in position. Linear wear with software to produce image.	Definite, probable and possible evidence of loosening of stem by Barrack 1992. Acetabular loosening if >2mm shift or complete radiolucent line >2mm.	Sepsis (2), # femur (2), dislocation (4) - possibly due to 22mm heads and habitual squatting.	HHS not sensitive enough to detect differences between groups based on mean score of each (cemented v. cementless, unilateral v. bilateral). Wear rate of 0.22mm/yr in zirconia heads (in cemented stems) and 0.14mm/yr in cobalt-chromium (on uncemented stems) was not as expected. Wear greater in younger, male patients with higher inclination angle.
Korovessis et al, 2006, Greece	Follow-up study	3	77 mths (60-112)	194	55 yrs (25 - 70)	217 uncement-ed (MoM)	6.5%	n/a	Physical examination (no details). HHS; satisfaction; functional classification of walking handicap (unlimited walkers, active, limited, household only, wheelchair). Blood tests - ESR, CRP and haematocrit. Histology of periprosthetic tissue at revision.	Osteolysis (focal area bone resorption ≥2mm wide not present on post-op x-ray) by zones. Brooker grade of heterotopic ossification. RL lines by zone and width; stem and cup positions.	Acetabular loosening - migration >3mm or continuous RL line >2mm in 2 or more zones. Stem - subsidence >2mm, varus >3°, or continuous RL line >2mm in 2 or more zones.	Aseptic loosening 4% (9 hips), technical failure (2 hips), infection (3 hips).	Cobalt-chromium metallosis in capsule and periprosthetic tissues of all revised for aseptic loosening. Longer follow-up studies required for metal-on-metal articulations.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Mabry et al 2004, USA	Follow-up study	3	12.2 yrs (2-28.9)	84	68 yrs (36-92)	84 Charnley cemented (post nonunion of # neck of femur)	14.0%	34.0%	n/a	Acetabular loosening by Hodgkinson et al criteria; femoral loosening by Harris et al (but <2mm RL line at shoulder of prosthesis not considered to represent loosening).	RL at cement-bone interface, probable 100% and possible 50-100%.	Loosening; dislocation (9%).	Younger patients (nonunion # neck of femur) need more of a joint-sparing procedure than a Charnley with first generation cementing technique. Older patients may be OK with better cementing and Charnley but need to avoid dislocation risk.
Makela et al 2008, Finland (JBJS Am)	Registry study	2	6.9 yrs (0-25)	n/a	70 yrs (55-96)	50,968 uncement-ed, hybrid and cemented	n/a	n/a	n/a	n/a	Revision of any part of THA	Aseptic loosening	Cementless stems had better survival than cemented in patients 55-74 years old. Above 75 yrs, survival was no different in any of the THA combinations. Uncemented cups were associated with multiple revisions for liner wear in 55-74 yr age group. Liner wear and periprosthetic osteolysis are still major problem on acetabular side.
Makela et al 2008, Finland (JBJS Br)	Registry study	2	7	n/a	71.9 yrs	34549 cemented	n/a	n/a	n/a	n/a	Revision of any part of THA	Aseptic loosening (27%), infection (4%), dislocation (9%),plus others	Exeter universal stem best survival overall but in the 55-74 yr age group, cemented stems did not produce excellent long term survival. Advocate continued surveillance beyond 10 years.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LE VE L	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Malchau et al 2005, Sweden	Registry study (report)	2	1, 6, 10 yrs	n/a	n/a	229,031 uncement-ed, hybrid and cemented	9.9% cemented, 10.8% hybrid, 28.1% uncemented	13% additional THA	Charnley category; VAS pain; EQ-5D; VAS satisfaction;	Osteolysis around cup or stem; definite loosening of cup or stem; wear indicated by eccentric head position.	Revision and radiographic loosening	Aseptic loosening 73.9%; deep infection 7.9%; dislocation 7.5%; periprosthetic fracture 5.7%.	High risk or undocumented prostheses should be on more regular FU with symptoms or ominous radiologic findings leading to orthopaedic referral. 218 pts with positive radiological findings had the same pain, satisfaction and EQ-5D results as those without radiological changes - i.e. not related to symptoms.
Maloney et al 1999, USA	Follow-up study	3	6.75ys & 10 yrs	941 & 184	56 yrs (16-87)	1081 & 204 hybrid	2.0%	2.4%	n/a	Comparison of initial and final FU x-rays: cup migration (vertical and horizontal with reference to interteardrop line); radiolucencies; inclination of cup. Linear wear by Livermore method and wear per year.	Aseptic loosening of cup: change of position, migration> 4mm, or complete radiolucency around cup.	PE wear and osteolysis	PE liner wear and osteolysis are main cause of revision surgery - cups remained will fixed. Wear pattern behind cementless cup is more often expansile and asymptomatic - FU advocated. Average time to osteolysis 8.67 yrs. Wear rate higher in younger patients (0.22 mm/yr in 30-40's, 0.12 in 50-60's, 0.08 in 60-70's). Pelvic osteolysis and revision surgery higher in younger pts.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
McMinn et al 2008, UK	Follow-up study	3	7.8 yrs (6-9.6)	103	47.2 yrs (21-62)	110 resurfacing arthroplast-ies	2.7%	zero	OHS (12/60 average at final review).	Radiolucencies around either component (used Amstutz et al 2004 score). Neck thinning; heterotopic ossification (Brooker).	Revision for aseptic loosening or other reasons.	# Neck of femur, collapse head, infection (one of each).	Use of dysplasia cup with attached lugs for neutralisation screws allows good fixation and construction of acetabular cover for severe DDH.
Muller et al 2003, Switzerland	Registry study	2	up to 33 yrs	15430	67.1 yrs	17951 (47.7% hybrid, 24.6% cemented, 14.7% cementless, reverse hybrid 13%)	3-14% at 10 yrs	n/a	n/a	RL lines around cup, migration, change in cortical density, ossification, others.	Acetabular loosening: continuous RL in DeLee and Charnley zones 1 to 3, superior migration or medial ≥ 5mm, severe protrusion, progressive tilt, # of cup or cement mantle.	First signs of acetabular loosening.	1. Cemented PE cups have good primary mechanical stability but no transformation of fixation over time so tend to fail late (7.6 yrs in this study). 2. Press-fit and expansion shells rely on early mechanical stability being transformed into secondary stability by osseointegration - will fail early if this does not happen (RL lines seen); otherwise, they perform well. 3. Hemispherical shells with screws may achieve early stability but will fail through bone remodelling around screws if osseointegration fails to take place. Screws hold for while then late failure occurs as cup migrates.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Nixon et al 2007, UK	Non-randomized controlled cohort	3	7.7 yrs (6-10)	127	n/a	127 Charnley cemented (Elite plus)	zero	34% mild loosening, 28% severe loosening.	Pain on VAS scale; OHS. DEXA scan. BMI. Biochemical tests.	Loose or stable by Dall, Learmonth and Solomon 1992.	Score >10 points or revision surgery.		VAS score >OHS detected severe loosening but not the earlier signs. Risk factors for loosening include atrophic OA, history of smoking, fragility #, lower peri-prosthetic bone mineral density (possibly). These patient groups need radiological surveillance.
Norton et al 2002, Uk	Follow-up study	3	3-7 yrs	26	49.2 yrs (31 - 57)	29 Charnley cemented (Elite Plus, ceramic head, Hylamer cup)	31.0%	52.0%	HHS, OHS, Merle D'Aubigne and Postel	RL and osteolysis by zones.	RL in 2 or more acetabular zones, 3 or more femoral zones or progressive osteolysis.		No statistically significant difference in scores between failed and surviving THA (but small sample). Decreased wear resistance of Hylamer over time led to progressive radiolucencies, osteolysis and clinical failure of THA. Early revision advocated to avoid bone loss. Ongoing surveillance of this group is essential.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LE VE L	TIME POST SURG- ERY	NO PTS	AV. AGE AT SURG- ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Ogino et al 2008, Finland	Registry study	2	n/a	6540	82.7 yrs	6989 cemented (82%), hybrid or uncement- ed	2.8%	n/a	n/a	n/a	Revision for any removal, exchange or re-implantation of prosthetic components.	Aseptic loosening 46%, recurrent dislocation 20%, periprosthetic # 13%, infection 13%.	Mis-matched components did not appear to show worse survival. Women had better THA survival. Proportion of aseptic loosening in older pts was lower than younger population but other indications for revision were higher. Hybrid better than cementless if used in this population but cemented were better still.
Ollivere et al 2009, UK	Follow-up study	3	12 yrs	217	n/a	234 Charnley cemented (Elite plus)	8.0%	57.7%	OHS	Both components rated as well-fixed, possibly loose or definitely loose. Femoral component - Harris et al 1982, Kobayashi et al 1997. Acetabular component - Hodgkinson, et al 1988.	Revision	Aseptic loosening (18/19 pts), 17 stem and cup. 1 patient for infection.	Poorer OHS score in pts with loose femoral component; cup loosening not identified by low OHS. No loosening on x-ray at 6 yrs was predictive of no revision at 12 yrs (high negative predictive value). Advocate clinical scores and x-ray at mid-term to give indication of longer term survival.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Ong et al 2002, USA	Follow-up study	3	8.4 yrs and 13.5 yrs	n/a	55 yrs for precoat-ed stems, 59 yrs for uncoated stems	768 cemented (254 uncoated stems, 514 precoated stems)	9.5% of precoated stems @5.7yrs, 3.9% of uncoated stems @9.9 yrs.	2.2% acetabular cups	HHS	RL and osteolysis. Cement grading using Mulroy et al.	Revision or impending revision. Failure = progressive RL ≥2mm around cement in all zones or migration of prosthesis or # of prosthesis	Loosening at bone-cement interface plus evidence of cement-stem debonding. High proportion of osteonecrosis patients influenced failure as well as other factors.	Design and surface treatment of stem including surface roughness and precoating of methylmethacrylate led to early failure.
Perka et al 2004, Germany	Follow-up study	3	9.3 yrs (82-143 mths)	93	52.6 yrs (26-73)	121 uncement-ed (DDH)	3.0%	1% impending	HHS; Merle d'Aubigne; subjective satisfaction; functional limb length.	RL lines or osteolysis by zones; loosening if continuous RL line, migration if ≥3mm change. Cup - vertical and horizontal migration and change in cup inclination angle. Wear from two dimensional penetration measured at cup entrance; volume by Charnley and Halley method 1975). Heterotopic ossification by Brooker grade. Anatomically correct centre of rotation by Ranawat et al.	Continuous RL line or migration of component.	Aseptic loosening of cup in 3 THA, 1 infection.	Uncemented components with medialised cup provided good outcomes for DDH patients.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Petsatodis et al 2010, Greece	Follow-up study	3	20.8 yrs (20-24)	78	47.5 yrs	85 uncement-ed (ceramic)	7.1% (6 hips)	n/a	Merle d'Aubigne-Postel.	Subsidence of femoral component, migration of acetabular cup, change in hip centre, calcar resorption, radiolucencies, osteolysis (any focal endocortical erosion progressing over time)	Revision of cup or stem	Aseptic loosening of acetabular cup with focal osteolysis.	RL lines <2mm noted in 79% of cases but had not progressed from one year post-op. Overall good survival for this age group and length of time.
Pollard et al 2006, UK	Follow-up study	3	5-7 yrs	101	50 yrs (18-67)	54 hybrid, 54 resurfacing arthroplas-ty	6% resurfacing, 8% hybrids	12% hybrids and 8% resurfacing impending.	OHS, UCLA, EuroQol	Hybrids - RL by zones, migration, osteolysis, linear wear (Dorr and Wan method 1996). Resurfacing - zones 1 to 3 on AP, 4 (anterior) to 6 (posterior) on lateral. RL lines > 1mm, migration, osteolysis, change in trabeculae density, cortical hypertrophy.	Resurfacing - failure of femoral component. Hybrids - osteolysis (3), recurrent dislocation (1).	Possible late avascular necrosis affecting resurfacings but osteolysis affecting hybrids.	60% of resurfacings had pedestal sign. Regular long term FU advocated.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Pospischill and Knahr 2005, Austria	Follow-up study	3	14.4 yrs (10.2-17.1)	99	59.2 yrs (24-84)	103 uncemented (32mm ceramic femoral head and UHMWPE insert)	14.6%	n/a	HHS, WOMAC, SF-36	RL lines, cortical hypertrophy, subsidence (medial point of lesser trochanter and tip of stem); osteolysis (cystic lesion at implant-bone interface); heterotopic ossification (Brooker). Linear wear - half the difference between superior and inferior margins of acetabular component of femoral head.	Revision of any part of THA	One hip infection; all others were aseptic revisions related to acetabular wear or loosening with wear debris affecting one femoral component as well.	Low level of osteolysis in proximal femur (possibly due to ceramic on PE articulation). Poorer WOMAC and SF-36 scores than normative values possibly due to comorbid conditions. Diaphyseal press-fit design leads to proximal stress shielding (RL in zones 1 and 7 early but unchanged at 10 yrs) and cortical hypertrophy in diaphysis.
Rajaratnam et al 2008, UK	Follow-up study	3	17.5 yrs (15-21)	291	71.2 yrs (31-90)	331 uncemented (1/3) and hybrid (2/3) (with 32mm ceramic femoral head)	1.8% femoral component (9.1% acetabular component)	n/a	Merle d'Aubigne-Postel.	RL by zones, subsidence of femoral component, increased bone density (spot weld formation), pedestal formation (blastic reaction at tip)	Aseptic failure of femoral component	Femoral component - Infection, trauma or trunion fretting. Acetabular component - aseptic loosening, polyethylene wear, dislocation, infection.	Aseptic failure of acetabular component was main threat to survival
Revell et al 2006, UK	Follow-up study	3	6.1 yrs (2-12)	60	43 yrs (17-69)	73 resurfacing arthroplasty	6.0%	7.0%	n/a	Loosening by Amstutz 2004 and Pollard 2006 (femoral). Heterotopic ossification (Brooker).	Revision	Fractures, femoral head collapse, femoral loosening, infection.	Suggests good outcome for Ficat grade III and IV osteonecrosis.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LE VE L	TIME POST SURG-ERY	NO PTS	AV. AGE AT SURG-ERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Santori & Santori 2010, Italy	Follow-up study	3	8 yrs (5 to 14)	109	51 yrs	129 uncement-ed (short femoral prosthesis)	4% acetabular	4.0%	HHS pre-op and at final FU; WOMAC at latest FU.	Femoral alignment, subsidence (centre of femoral head to lesser trochanter, >2mm), radiolucent lines, calcar rounding (smoothing of sharp definition of osteotomy of femoral neck), proximal bone resorption (loss of trabecular bone density/mottling or loss of definition between cortical and cancellous bone), endosteal hypertrophy (spot welds= local deposition of new bony trabeculae bridging the cortex and surface of the implant), distal cortical hypertrophy (new bone of cortical density resulting in an increase of cortical thickness), osteolysis (endosteal erosion or progressive radiolucent lines), femoral fit (over, under or normal; diaphyseal and metaphyseal fit).	Stem alignment of >2° between longitudinal axis of stem and of femur. Shedding of porous coating, distal osteolysis.	Polyethylene wear (3x liner exchanges, 2x total cup replacement).	Early implants with diaphyseal stem associated with pedestal and some cortical hypertrophy. No distal osteolysis seen. Better results with removal of short diaphyseal stem. No implants failed or moved into varus (possibly due to lateral flare).

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Spangehl et al 2001, USA	Follow-up study	3	7.5 yrs (5-12)	35	39 yrs (12-67)	44 uncemented cups (DDH)	9% acetabular (7% femoral)	n/a	HHS	RL by zones. Migration of acetabular component, inclination. Resorption (or lack of) of bone graft (disappearance of femoral head-host bone interface and/or bridging trabeculae).	Revision. Loosening by change in position of cup or complete RL line ≥2mm at bone-implant interface.	Wear and osteolysis without loosening; loosening; fracture of cup.	Moderate acetabular deficiency in DDH can be treated successfully using femoral head autograft and cementless cup.
Sugano et al 2007, Japan	Follow-up study	3	6 yrs (5-8)	143	53 yrs (27-74)	180 uncemented (ceramic-on-ceramic)	2% (2 pts) in non-navigated group	3.0%	Merle d'Aubigne-Postel.	AP and lateral - only used AP with pubic symphysis and spinous processes formed vertical line perpendicular to teardrop. Acetabular - inclination and anteversion (axial view). Femoral - Engh's classification of bone ingrown, stable fibrous or unstable. Subsidence or varus (using line through midpoint 1cm below lesser trochanter and 1cm above stem tip). Osteolysis. Leg length discrepancy - comparison of two sides for distance between inter-teardrop line and medial prominence of lesser trochanter.	Subsidence >4mm; VR >2mm change; migration of cup >2mm change; Inclination angle >5° change.	Impingement	No statistically significant difference between dislocations of those inside Lewinnek's safe zone and those outside. No score difference between navigated and non-navigated groups. Impingement and LLD only in non-navigated group.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Utting et al 2008, UK	Follow-up study	3	13.6 yrs (12-16)	53	40 yrs (19-49)	70 hybrid	16.0%	21% cups, 39% stems impending failure.	OHS, HHS	AP and lateral views. Inclination of cup, stem subsidence, migration, RL lines by zone, osteolysis (RL zone>2mm wide). Fracture of component. Linear wear (Livermore method).	Revision or intention to revise for failed acetabular component defined as migration >2mm, 2° change in inclination, RL >2mm in all zones. Femoral stem - subsidence >5mm.	Aseptic loosening, recurrent dislocation, linear wear, deep infection, osteolysis, lack of osseointegration,	Non-progressive RL lines around cup <2mm in 44%. Osteolysis around cup in 19%. Higher rates of revision or planned revision after 10 years in hybrid THA. OHS and HHS remain good suggesting silent osteolysis. Regular clinical and radiological review recommended in hybrid THA under 50 yrs of age.
Wangen et al 2008, Norway	Follow-up study	3	13 yrs (10-16)	44	25 yrs (15-30)	49 uncement-ed	49.0%	n/a	WOMAC, EQ-5D, HHS	Bone remodelling - incorporation as bone-implant contact, periprosthetic bone formation and absence of migration. Femoral - cortical or endosteal bone formation, subsidence >5mm. RL lines; radio dense lines parallel to surface of implant but separated by RL lines = fibrous ingrowths. Osteolysis - focal areas of bone loss.	Radiological loosening or revision	Acetabular failure for loosening (wear particles→macr-ophages and osteoclasts→ acidity→ dissolution of HA coating→shear stresses→ delamination plus foreign body reaction), PE wear and subsequent osteolysis, dislocations.	Some RL and osteolysis found in Gruen zones 1 and 7 but osseointegration centrally and distally. Suggest WOMAC and EQ-5D are not both needed as correlated with each other but not with HHS.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Williams et al 2002, UK	Follow-up study	3	10 yrs (8-12)	201	67.55 yrs (24-87)	201 cemented (Exeter)	6% acetabular component	n/a	Merle d'Aubigne-Postel, HHS, OHS	Socket or stem migration (subsidence from RL in zone 1 or tip stem to inferior pole of centraliser), RL lines by zones (RL line adjacent to sclerotic line), changes in femoral neck, diaphyseal hypertrophy, localised lysis. Cement grading by Barrack et al;	Revision of component	No stem failure. Cups failed for dislocation or loosening.	80% stems in VR or VL up to 2°; 98% <3mm subsidence (most within first two years). Exeter stem functions by 'taper-slip' principle, sealing off effective joint space to fluid. Polished surface minimises debris production at cement interface.
Wroblewski, Siney, Fleming 2007, UK	Follow-up study	3 minus	31 yrs	####	67 yrs (12-93)	22066 Charnley cemented	45.0%	n/a	No details	No details	Revision	Loose acetabular component> loose stem> fractured stem> infection> dislocation	Wear and loosening of UHMWPE cup were main problem. Will revise for radiological changes alone to avoid complexity of complications.
Wroblewski, Siney, Fleming 2009, UK	Follow-up study	3 minus	17 yrs (1-40)	1016	41 yrs (12 - 50)	1332 Charnley cemented	23.0%	n/a	No details	Complete demarcation >1mm of cup or change in position. Wear measurements (Griffiths 1978).	Acetabular loosening	Aseptic loosening	Increasing loss of 2:1 ratio of cup diameter to head size reduces the low friction arthroplasty principle which is based on frictional torque. Also, wear of PE leading to stem impingement on cup and mechanical loosening.

AUTHORS, DATE, COUNTRY	STUDY TYPE	LEVEL	TIME POST SURGERY	NO PTS	AV. AGE AT SURGERY	NO. AND TYPE OF THA	% REVISED	NO. FAILED	REVIEW METHODS CLINICAL	REVIEW METHODS X-RAY	DEFINITION OF FAILURE	REASONS FOR FAILURE	KEY FINDINGS
Wroblewski, Siney, Fleming, Bobak 2000, UK	Follow-up study	3 minus	12.4 yrs (1-21) and 8.3 yrs (1-14)	358	40 yrs (17-51)	441 Charnley cemented	0.3% routine group, 0.9% in calcar femorale group	4.8% in routine group	No details	Stem fixation by Harris et al 1982.	Revision or definitely loose	Lack of adequate proximal cementation due to presence of calcar femorale	Advocates surgical clearing of calcar femorale. Advocates follow up beyond 10 yrs which highlighted the problem.
Yates et al 2008, UK	Follow-up study	3	11.8 yrs (10-12)	175	62.3 yrs (19-83)	175 cemented and hybrid (collarless polished taper stems)	No stems	1.0%	HHS	Cement mantle (Barrack et al); Alignment and subsidence of stem. RL lines >1mm wide and >5mm long by zones. Osteolysis (Joshi et al 1998). Heterotopic ossification (Brooker). Proximal medial bone resorption (Engh et al 1987). Wear (Latimer and Lachiewicz).	Stem loosening	Acetabular wear and loosening in hybrid (H/G) THA. Infection, dislocation. Others.	CPT stem subsides in cement, particularly in first year. Should have 2mm mantle. Heterotopic ossification limited movement. HHS scores not linked to radiological scores, age of pt or magnitude of subsidence.

Appendix VIII

Results of literature review: radiographic assessment

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Amstutz, Beaulé, Dorey, Duff, Campbell, Gruen 2004	JBJS Am 2004: 86A; 28-39.	Hybrid surface arthroplasty assessment	Femoral component - fixation on a scale of 0-9 points with 0=no lucencies; 9= migration, 1-3=lucency in one zone, 4-6=lucencies in 2 zones, 7=lucencies in 3 zones incomplete, 8=lucencies in 3 zones complete. Zone 1=superior stem, 2=tip stem, 3=inferior stem. Acetabular fixation in 3 zones with 0-9 points, 0=none, 9= migration and others as for stem (zones I to III). Heteroptopic ossification (Brooker)	RL score ≥7. Stem-shaft angle increase > 5° = valgus; stem-shaft decrease > 5° = varys. Sagittal stem axis deviation >10° (anterior or posterior).	Need to select patients carefully; follow up is essential to define results and indications for this procedure.
Barrack, Mulroy and Harris 1992	JBJS Br 1992	Cement grading	A= complete 'white-out' of medullary cavity by cement, B=slight radiolucency of cement-bone interface, C=50-99% radiolucency of cement-bone interface or defective/incomplete cement mantle, D=100% radiolucency at cement-bone interface or uncovered stem tip due to unfilled canal.	Definite loosening = migration or change in position of stem or cement (includes fracture of cement or RL at cement-stem interface not present on immediate post-op x-ray). Probably loose = 100% RL line at cement-bone interface without migration. Possibly loose = RL 50-99% of a zone.	
Berry, Harmsen, Ilstrup 1998	JBJS Am 1998: 80:715-21.	Charnley stem loosening	RL line at supero-lateral border of Charnley femoral component: Class I =< 0.5mm, class II = 0.5 to 1.9mm, class III = >2.0mm	>2mm associated with higher risk of aseptic loosening	<2mm RL line at supero-lateral border in the first five years is not significant in Charnley prostheses.
Brooker, Bowerman, Robinson, Riley 1973	JBJS Am 1973; 55 ;1629-32.	Heterotopic ossification	Class I = islands of bone in soft tissues of hip;, II = bone spurs from pelvis or proximal end of femur with at least one cm between opposing bone surfaces; III = bone spurs from pelvis or proximal femur with less than one cm between; IV = apparent bony ankylosis of hip		
Charnley & Halley 1975	Clin. Orthop.1975; 112:170-9.	Wear measurement - cemented cups (serial x-rays)	Volume calculation of wear from linear wear measured by Griffiths method		
Cornell and Ranawat 1986	J Arthroplasty 1986;1:197-202.	Cemented acetabular fixation	RL line around acetabular component scored by adding thickness of line in all three zones e.g. 1mm in all three zones = 3 points, 3mm thickness in two zones = 6 points.		

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Crowe, Mani, Ranawat 1979	JBJS Am 1979;61:15-23	Level of dislocation of dysplasia	Based on the extent of proximal migration of the femoral head: I = <50% subluxation, II = 50-75% , III = 75-100%, IV = >than 100% subluxation		
Dall, Learmonth, Solomon 1992	Int Orthopaedics 1992; 16:339-43.	Radiological loosening score	Socket score (total 10pts): Cement-bone RL (zones and width), migration. Stem score (15 pts): cement-bone RL (width and zones), subsidence within and without cement mantle, resorption of shaft	Grade 1 = slight changes, Grade 2 = Moderate changes, requires monitoring, Grade 3 = severe changes with impending failure	RL of 1mm = suspicion of loosening; 2mm = definite demarcation; 3+mm = more severe change; calcar resorption can be counted as zone 7 change. Authors suggest simple score to record changes to allow comparison over long term.
Dorr & Wan 1996	J of Arthroplasty 1996;11,4:419-28.	Wear measurement - cemented and uncemented cups (one x-ray possible)	Obtain the distance from the edge of cup to femoral head at the opening of cup both superiorly and inferiorly. Calculate the difference between these and divide by 2 to obtain measurement of linear wear.	Presence of radiolucent lines- accepted standard for evaluation of fixation of components, cemented or uncemented. Impending failure judged by occurrence of RL lines.	1 patient revised for osteolysis; the occurrence of RL lines was predictive of impending failure.
Dorr, Bechtol, Watkins, Wan 2000	J. Arthroplasty 2000:15;890-900.	Acetabular measurements in THA	Acetabular anteversion: AP of pelvis and AP centred over hip used to measure difference of the angle between a line across the cup opening and a line from the edge of the cup to the femoral head centre. Value of difference indicates anteversion (positive) or retroversion (negative) of cup.		
Engh, Bobyn and Glassman 1987	JBJS Br 1987;69:45-55.	Calcar bone resorption	Calcar bone resorption: Grade 1 - atrophy or rounding of calcar; Grade 2 - loss of density of calcar with preservation of medial wall to lesser trochanter; Grade 3 - loss of density with loss of medial wall to lesser trochanter; Grade 4 - loss of density of medial wall to below lessser trochanter.	*Bony fixation* = no subsidence and none/minimal radio-opaque line around stem; most bone-implant interface appears stable. *Stable fibrous ingrowth* = no progressive migration (even if some in early stages); extensive radio-opaque line formation around stem in parallel with stem and radiolucent space ≤ 1mm; no local cortical hypertrophy. *Unstable implant* = subsidence or migration in canal and diverging radio-opaque lines; cortical hypertrophy at tip and collar.	Resorptive remodelling (due to stress shielding) was more extensive in larger, more rigid stems

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Engh, Massin and Suthers 1990	CORR 1990;257:107-28.	Fixation of cementless femoral component	Fixation score: appearance of porous interface (reactive lines, RL), spot welds. Stability score:appearance of smooth interface (lines, RL), pedestal presence, calcar modelling, interface deterioration (increase in width of RL or reactive line), migration, particle shedding.	Lack of osseointegration shown by reactive lines around porous surface; instability of distal portion shown by canal widening, reactive lines and /or subsidence.	Need serial x-rays! X-rays are better predictor than intra-operative test of stability as it may only be a fibrous fixation. Osseointegration shown by absence of reactive lines adjacent to porous surface and presence of spot welds of endosteal new bone. RL around smooth surface not necessarily bad if porous surface ingrown. Pedestal acceptable if in contact with stem tip and no new RL lines.
Field, Singh, Latif, Cronin, Matthews 2006	JBJS Br 2006;88: 315-20.	Follow-up of cemented femoral stem with radiostereometric analysis for migration	Radio-opaque beads in region of greater trochanter at surgery; standardised positioning of pt for subsequent x-rays; bead to stem tip distance for migration.		Migration of 1.89mm average at 5 yrs (indication of an acceptable migration)
Griffiths, Seidenstein, Williams, Charnley 1978	CORR 1978;137:37-47.	Wear measurement - cemented cups (serial x-rays)	Standardised positioning of patient and x-ray tube; wear distance measured using caliper from centre of femoral head to wire marker at site of greatest wear; adjust for magnification.		Slight differences in set up do not have a significant impact on measurements obtained.
Harris, McCarthy and O'Neill et al 1982	JBJS Am 1982;64:1063-67.	Loosening of cemented femoral components	Definitely loose: migration with RL lines at stem-cement interface and shift of stem; crack of cement or # stem. Probably loose: RL line around 100% of cement-bone interface. Possibly loose: RL line around > 50% interface but <100%.		
Hodgkinson, Shelley, Wroblewski 1988	CORR 1988;228:105-9.	Loosening of cemented acetabular components	Assessment of bone-cement junction of socket: 0 = no demarcation, 1 = demarcation of outer 1/3 only, 2 = demarcation of outer and middle thirds, 3 = complete demarcation, 4 = socket migration; sclerotic line is not an indication of fixation. RL lines >1mm width and type 2 or 3 classification means that socket is loose.	Migration of cup or RL lines >1mm in all zones	Any RL line that is new, progressive or not apparent on initial x-ray is significant; the extent of demarcation is more important than gap size. One year x-ray useful for predicting long term results. Surgeon must be prepared to intervene for x-ray changes even if asymptomatic because of loss of bone stock leading to socket migration.

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Johnston, Fitzgerald, Harris, Poss, Müller, Sledge 1990	JBJS Am 1990;72:161-8.	System for THA evaluation	Acetabulum: migration, cement grade, RL width and zone, breakage of screws, wear, inclination, version. Femur: alignment in varus/valgus, subsidence, cement grade, stem #, RL width and zone, resorption, bone density, endosteal cavitation, ectopic ossification, greater trochanter bony quality. Add an evaluation of porous coating for uncemented components. Defines zones 8 - 14 in addition to zones 1 - 7.	Migration of prosthesis or cracks in cement. Radiolucencies at prosthesis-cement interface of femur indicates cracking and movement.	Width and extent of RL lines is important but meaning at bone-cement interface not known. Zones can be combined innovatively for reporting purposes.
Joshi, Eftekhar, McMahon, Nercessian 1998	JBJS Br 1998;80:585-90.	Osteolysis	Osteolysis = new expansile, cystic lesion with endosteal scalloping and/or migration. Recorded by zone; size with digitizer in sq. mm		Wear and instability of implant increase incidence of osteolysis which appears at approx 6 yrs. Recommend 6mm cement mantle around cup, 3mm around stem.
Kobayashi, Donnelly, Scott, Freeman 1997	JBJS Br 1997;79:583-89.	Predictors of femoral survival	Femoral component: RL lines, osteolytic lesions, migration, bony changes.	Predictors of femoral stem failure which should be examined: RL lines >2mm, osteolytic lesions >2mm, migration, osteopoenia, pedestal formation, femoral neck resorption, cortical hypertrophy.	Rule of two proposed: RL >2mm width and >2mm of migration at 2 yrs = increased risk of failure at 10 yrs. Add osteolysis >2mm at 5yrs and risk increases more. Migration >2.0mm had 96% negative predictive value NPV (26% PPV); RL>2mm at 2 yrs were not highly predictive of failure (NPV 96%, PPV 29%); osteolytic lesion >2mm at 5 yrs had 92% negative predictive value (53% PPV); cortical hypertrophy at 2 yrs had 93% NPV but only 29% PPV.
Latimer, Lachiewicz 1996	JBJS Am 1996;78:975-81.	Wear measurement - uncemented cups (serial x-rays)	Measurement of femoral head diameter to correct for magnification; then one measurement of polyethylene thickness in supero-medial direction from edge of femoral head to outer to outer acetabular wall. Most recent and post-op x-ray compared to obtain linear wear as the difference between the two.		

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Lewinnek, Lewis, Tarr, Compere, Zimmerman 1978	JBJS Am 1978;60:217-20.	Dislocations after THA (Safe zone of cup)	Meaurement of acetabular inclincation and anteversion	Dislocation	Anterior dislocations were associated with increased acetabular-component anteversion. There was no significant correlation between cup-orientation angle and posterior dislocation. The dislocation rate for cup orientation with anteversion of 15 +/- 10 degrees and lateral opening of 40 +/- 10 degrees was 1.5%, while outside this "safe" range the dislocation rate was 6.1%.
Livermore, Ilstrup and Morrey 1990	JBJS Am 1990;72:518-28.	Wear measurement - cemented cups (serial x-rays)	Templating to find centre of femoral head. Calipers used to measure distance from centre of femoral head to edge of cup. Identify the point of greatest wear as the shortest radial distance. Measure on post-op x-ray at the same point and calculate the difference in the measurement to obtain the linear wear.		Higher wear associated with resorption and lysis in proximal femur.
Martell, Pierson, Jacobs, Rosenberg, Maley, Galante 1993	JBJS Am 1993;75:554-570.	Evaluation of uncemented THA	Standardisation of AP and lateral views, with definition and subdivision of zones. Acetabular component: inclination, migration (vertical only), RL lines >2mm. Femoral component:stem angle, subsidence, remodeling of bone, cortical reaction, RL and sclerotic lines (>2mm), degree of canal fill, pedestal/canal plug, heterotopic bone (Brooker).	Possibly loose cup if RL in at least 2/3 of circumference and >2mm width in at least one zone. Cup migration>2mm definite instability. Femoral subsidence (>2mm), cortical hypertrophy and erosion, RL <2mm or >2mm, pedestal, canal filling.	Femoral revision in asymptomatic patient with well-fixed femoral prosthesis due to cortical erosion; three other femoral revisions due to radiographically assessed instability (all had poor Harris Hip Score).
Massin, Schmidt, Engh 1989	J of Arthroplasty 1989;4,3:245-51.	Cementless cup migration	Vertical migration of cup best measured between centre of cup and inter-teardrop line; horizontal migration best measured between vertical lines through teardrop and centre of cup.		Standardisation of x-ray films: at 10% magnification, distance between obturator line and teardrop line should be < 5mm; distance betweeen middle of sacroiliac line and vertical line through pubis < 5mm.
Moore, McAuley, Young and Engh 2006	CORR 2006:;444:176-183.	Signs of osseointegration of uncemented cups	Osseointegration shown by: 1=Absence of RL lines (>1mm, 2 or more zones) 2=Presence of superolateral buttress 3=Presence of medial stress shielding 4=Presence of radial trabeculae (zones I&II) 5=Presence of inferomedial buttress.	Surgeon assessed stability at RHR surgery	Five signs give 97% PPV (NPV 48%), 90% sensitivity if 3 or more are present.

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Müller, Gautier, Roeder, Busato 2003	JBJS Br 2003:85;31-6.	Acetabular cup failure		Acetabular failure - continuous RL in zones I to III; migration >5mm, severe protrusion, progressive tilt, # of cup or cement mantle.	Uncemented cups rely on direct mechanical stability which is subsequently reinforced with osseointegration - threatened by early instability. Cemented cups are immediately stable but no osseointegration means that they loosen over time - late failure.
Mulroy, Estok, Harris 1995	JBJS Am 1995;77:1845-52.	Cementation of femoral stem	Grade A = complete filling of proximal portion medullary canal of diaphysis (white-out), B = nearly complete filling, C1 = RL line > 50% cement-bone interface, C2 = <1mm cement mantle at any site or other mantle defect, D = gross deficiencies of cement mantle.		
Paprosky, Weeden, Bowling 2001	CORR 2001;393:181-93.	Osteolysis in cementless THA	Engh for stability and bone ingrowth of femoral components; subsidence, osteolysis and stress shielding. Acetabular components - tilt, migration, metal particle shedding.	Osteolysis - expansile if rounded of scalloped and space extended away from surface of implant; linear if radiolucent space adjacent to component and >1mm.	Importance of planning pre RHA.
Pollard, Baker, Eastaugh-Waring, Bannister 2006	JBJS Br 2006;88: 592-600.	Resurfacing - femoral component assessment	*HSA classification*: 0 = No change; 1 = Pedestal sign but no migration (a= sclerotic line confined to tip of stem, b = sclerotic line confined to distal 1 cm of shaft, c = sclerotic line with/without lucent lines beyond distal 1cm of shaft); 2 = Migration, usually varus with lucent lines; 3 = displaced fracture.		
Ranawat, Dorr and Inglis 1980	JBJS Am 1980;62:1059-65.	Determination of anatomically correct centre of rotation.	A perpendicular line with 20% of pelvic height is plotted on the vertical of the Köhler line connecting teardrops. A line of the same length is then drawn laterally from top of previous line; end points of two lines connected to form triangle representing anatomically correct acetabular region (see Perka et al 2004 for diagrammatic illustration).		

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Schmalzried & Harris 1993	JBJS Br 1993;75:608-15.	Cement grading, hybrid THA	Grade A = complete filling >2mm width and beyond tip, white-out; B = some diaphyseal trabecular bone not filled; C = focal deficiencies including bubbles; D = multiple focal deficiencies or extensive mantle defects or lucencies.		Gaps behind uncemented cups post-op may be filled by bone subsequently. Calcar resorption common but partly stress shielding and partly lysis.
Shaver, Brown, Hillis, Callaghan 1997	JBJS Am 1997;79:693-700.	Wear measurement - cemented cups (serial x-rays)	Digital edge-detection measurement of polyethylene wear (maximum decrease in distance between femoral head ellipse and acetabular cup ellipse generated from grey-scale changes)		Maximum decrease in distance used rather than point of femoral head penetration as in Livermore; more accurate than template method for smaller wear measurements; only 3-4 minutes per measurement. Emphasises importance of early detection of high wear rates.
Sutherland, Wilde, Borden, Marks 1982	JBJS Am 1982:64:970-82.	Migration of stem and cup in THA	Cup migration measured with reference to Kohler's line (medial) and interteardrop line (vertical). Femoral subsidence measured with reference to centre of femoral head and tip of greater trochanter; also measured from inferior tip of prosthesis collar to superior margin of lesser trochanter.		
Toom, Fischer, Martson, Rips, Haviko 2005	Int Orthopaedics 2005;29:156-9.	Heterotopic ossification	Classification of heterotopic ossification: A0=absent, A1=isolated ossification <1cm length; B= >1cm distance between pelvis and femur B1=isolated ossification 1cm length, B2=marginal ossification; C= < 1cm between pelvis and femur C1=isolated >1cm C2=marginal ossifications, C3=ankylosis.		

AUTHORS	JOURNAL	SUBJECT	REVIEW METHODS X-RAY	DEFINITION FAILURE	KEY FINDINGS
Zicat , Engh, Gokcen 1995	JBJS Am 1995; 432-9.	Retrospective review of radiographs for pattern of osteolysis	RL lines by zones and width. Linear osteolysis = RL space adjacent to component with relatively uniform width >1mm (extent measured by width and zones as per Cornell and Ranawat). Expansile osteolysis = sharply demarcated space with rounded or scalloped appearance, extending away from implant (extent measured by maximum width and length and then approximate with formula for ellipse). Osteolytic lesions also defined as periarticular or remote from joint.	Unstable acetabular component = >1mm width RL line in all three zones. Stem fixed if calcar resorption, absence of demarcation lines along prosthesis surface, new endosteal bone in areas of previous gaps.	Effective joint space = path in periprosthetic region for passage of particulate debris away from articulating surfaces. Can be limited by circumferential porous coating that extends into diaphyseal cortical bone. Less osteolysis (linear pattern) around cemented acetabular components is probably due to fibrous tissue layer commonly found around them which provides path of least resistance to particulate debris. Limitation to debris is provided by the halo of sclerotic bone peripheral to cup. (Linear) osteolysis of cemented cups often leads to symptomatic loosening but (expansile) osteolysis of uncemented cups is asymptomatic and radiographs are needed for diagnosis.

Key: RL = radiolucency, PPV = positive predictive value, NPV = negative predictive value

Appendix IX

Publications and presentations

Publications and presentations

Ahmad, R., Gillespie, G., Barakat, M.J., Ahmed, S.M., Annamalai, S., Smith, L.K., Spencer, R.F. (2009) Leg length and offset following hip resurfacing and hip replacement. *Hip International*, 19(2):136-40.

Ahmad, R., Gillespie, G., Barakat, M.J., Ahmed, S.M., Annamalai, S., Smith, L.K., Spencer, R.F. (2008) Compariosn of biomechanics of total hip replacement and hip resurfacing. In: *Proceedings of the British Hip Society Annual Scientific Meeting, Norwich 27-29 February 2008.*

Parry, M.C., Whitehouse, M.R., Mehendale, S.A., Smith, L.K., Webb, J.C., Spencer, R.F., Blom, A.W. (2010) A comparison of the validity and reliability of established bone stock loss classification systems and the proposal of a novel classification system. *Hip International,* 20(1): 50-55.

Smith, L.K. (2011) Patient related outcome measures and radiological changes. In press: *British Hip Society Annual Scientific Meeting, Torquay 2-4 March 2011.*

Smith, L.K., Cramp, F., Palmer, S., Coghill, N. and Spencer, R.F. (2010) Use of morphometry to quantify osteolysis after total hip arthroplasty. *Clinical Orthopaedics & Related Research,* 468, 3077-3083.

Smith, L.K., Parry, M.C., Barakat, M.J., Spencer, R.F. (2009) Use of morphometry to quantify radiological osteolysis following total hip arthroplasty. In: *Proceedings of the British Hip Society Annual Scientific Meeting, Manchester 11-13 March 2009.*

Smith, L.K., Langkamer, V.G., Shannon, M.N., Dixon, J.H., Case, R., Mahajan, A.J. and Spencer, R.F. (2008) Performance of the Cenator and EPF-Plus acetabular components at three years. *Journal of Bone and Joint Surgery - British Volume*, 90(Supplement 2), 310.

Smith, L.K., Williams, D.H., Langkamer, V.G. (2007) Post-operative blood salvage with autologous retransfusion in primary total hip replacement. *Journal of Bone and Joint Surgery - British Volume*, 89(8): 1092-7.

Stevenson, A.J., Smith, L.K., Spencer, R.F. (2009) The effect of femoral stem geometry and the use of a proximal femoral centraliser on the restoration of lateral femoral offset in cemented hip arthroplasty. In: *Proceedings of the European Federation of National Associations of Orthopaedics and Traumatology, Vienna, Austria 3-6 June 2009.*

Utting, M.R., Lankester, B.J., Smith, L., Spencer, R.F. (2006) Proactive compliance with national guidelines (NICE) on total hip replacement. *Journal of Bone and Joint Surgery - British Volume*, 88(Supplement 2), 244.

Podium presentation, British Hip Society 2011

Patient related outcome measures and radiological changes

Authors: L K Smith, F Cramp, S Palmer, N Coghill, R F Spencer

Institutions: Weston General Hospital and the University of the West of England

(Abstract)

Introduction
Routine post surgery surveillance of total hip arthroplasty (THA) is widely recommended but is difficult to achieve in the current economic climate. Further evidence is needed to identify the most effective tools and time intervals for review to support established orthopaedic opinion.

Methods
A cohort of 201 hip replacements (101 cemented, 100 hybrid) were reviewed at three years and 154 (65 cemented, 89 hybrid) were available for mid-term review (average 7.5 years) when radiographic signs of deterioration commonly appear. Participants completed a patient reported outcome measure (Oxford Hip Score) at each review and x-rays were obtained. The data were explored using multiple regression analysis for associations between changes over time in the Oxford Hip Score (OHS), radiographic status, age, general health and comorbidities.

Results
The changes in OHS could be partially predicted by general health but not by age or comorbidities and there was no association between the OHS and radiographic changes. There was no statistically significant difference between participants grouped by age and the magnitude of the radiographic changes or the linear wear rate of polyethylene.

Discussion
Radiographic changes were evident in the majority of the THA but the magnitude of changes was not reflected in the change in individual OHS. We suggest that this provides evidence that it is not sufficient to use a joint specific patient reported outcome measure alone at mid-term review and that any surveillance service must include radiographic review. The age of the patient did not allow prediction of the magnitude of changes implying that selection for review by age must be with caution.

Copy of paper published November 2010

Reprinted with permission of Springer, license no. 2643110717917.

BASIC RESEARCH

Use of Morphometry to Quantify Osteolysis after Total Hip Arthroplasty

**Lindsay K. Smith MSc, MCSP, Fiona Cramp PhD, MCSP,
Shea Palmer PhD, MCSP, Nikki Coghill PhD,
Robert F. Spencer MD, FRCS**

Abstract

Background Progressive osteolysis threatens the longevity of hip arthroplasties and radiographic review is recommended. Measurement of osteolytic lesions in a clinical setting has not been achieved easily in the past. Other radiologic investigations provide accurate information but cost and risk to the patient prohibit their use in routine review.

Questions/purposes We developed a simple, reliable tool to be used in hip arthroplasty review to quantify osteolytic changes seen on plain film radiographs.

Methods A morphometric grid was developed and tested on simulated and actual osteolytic lesions. Four health professionals measured lesions on each of two occasions. Intraclass correlation coefficients (ICC) for interobserver and intraobserver reliabilities were calculated and Bland-Altman plots were constructed for graphic analysis.

Results The ICCs for interobserver reliability on the simulated and actual osteolytic lesions were in the range 0.90 to 0.96. The values for intraobserver (test-retest) reliability were 0.97 to 0.98. The Bland-Altman plots confirmed agreement and in each case, proximity of the mean to zero indicated no significant bias.

Conclusions The data show a morphometric grid is reliable for measuring osteolytic changes after hip arthroplasty.

Clinical Relevance This tool has potential to improve monitoring processes for hip arthroplasty and to be useful in future research studies. Additional work is needed to test for validity and clinical importance of the measurements obtained.

L. K. Smith (✉), R. F. Spencer
Department of Trauma and Orthopaedic Surgery,
Weston General Hospital, Grange Road, Uphill,
Weston-s-Mare BS23 4TQ, England
e-mail: smith@unforgettable.com

F. Cramp, S. Palmer, N. Coghill
School of Health and Social Care, Faculty of Health & Life Sciences, University of the West of England, Bristol, England

Introduction

There currently are more than 64,000 primary hip arthroplasties conducted annually in England and Wales, with as many as 10% subsequently requiring revision [26]. In the United States, the number of revision hip arthroplasties is expected to continue to increase in future decades [3]. Early causes of failure of the prosthetic joints include infection and recurrent dislocation, whereas late failure often is attributable to aseptic loosening and periprosthetic osteolysis [20, 43].

Osteolysis is a process of bone destruction in response to particulate wear debris [9, 16, 18, 24]. It may develop silently [1, 22], but the associated deterioration of the host bone predisposes the patient to periprosthetic fracture or aseptic loosening and may compromise the success of a revision procedure [19, 43]. Consequently, periodic review of patients is recommended and should include radiographs to record progressive changes [4, 22, 28]. Failure to do so can have major consequences [27].

Assessment of substantial bone loss attributable to osteolysis often is made from radiographs using a categorical scale (eg, classification of Della Valle and Paprosky [7] or Saleh et al. [35]), which provides an indication of management at revision surgery. However, it has been suggested that assessment of earlier osteolytic changes should include such information as the shape, size, and location of lesions [9]. This requires more precise measurement to quantify progressive changes. The area of a lesion has been measured in numerous ways, including geometric formulae and computer software [15, 17, 31], but these techniques are not readily applied in routine clinical practice.

Use of plain radiographs in the assessment of pelvic osteolysis sometimes is debated, because accurate detection of a lesion in some locations is more difficult than with CT and MRI—51.7% detected on radiographs compared with 74.7% by CT and 95.4% by MRI [41]. The use of CT with metal artifact suppression also has been shown to be more accurate for estimating the volume of osteolytic lesions [17, 22, 37]. However, the associated cost of CT or MRI and increased radiographic exposure for the patient (CT) preclude their use in routine screening. The addition of Judet oblique views to plain AP pelvic radiographs will improve sensitivity of detection from 41% to 73% and enable the observer to more precisely define the location of a lesion [5, 12, 40]. Consequently, it still is recommended that radiographic assessment be used in arthroplasty review [22, 39, 40].

Our purposes were (1) to develop the concept of a morphometric grid to provide a simple tool for measuring the area of osteolytic lesions and (2) to test the reliability of this tool to measure osteolytic lesions on radiographic films when used by a representative range of observers.

Materials and Methods

This observational study used the principle of cross intersect counting to develop a tool to measure the area of osteolytic lesions on radiographs. A morphometric grid is superimposed on the object of interest and the number of test points that fall within a defined area is counted.

Cross intersect counting has been used in biologic analysis for many years and it has been recommended that grid dimensions be chosen according to the object to be measured [10]. In radiographic assessment of hip arthroplasty, 2 mm is commonly used as a critical measurement of width of radiolucency [29]. Numerous authors consider linear radiolucencies greater than 2 mm a definite indication of loosening of the prosthesis [6, 8, 17, 34]. Therefore, a grid was developed and printed on transparency film with 2 mm between each point, which was a

small cross at the intersection of grid lines. The uniformity of the spacing was checked by measuring fixed distances on each of three grids with electronic calipers. The mean of 15 measurements between six crosses (10 mm) was 9.98 mm (SD, 0.04 mm). This was considered sufficient accuracy for the purposes of the study.

Interobserver and intraobserver (test-retest) reliability of the morphometric grid were tested using various observers, as recommended in the development of a new tool [11, 21, 25, 42]. Four health professionals were selected to represent the range and levels of experience involved in arthroplasty review. They included a senior orthopaedic surgeon in hip arthroplasty (RFS, Observer 2), an orthopaedic registrar (MCP, Observer 4), a senior physiotherapist with more than 8 years' experience in reviewing hip arthroplasty but no experience in quantifying osteolysis (LP, Observer 3), and the primary author (LKS, Observer 1) who is a researcher and trained physiotherapist who currently does long-term surveillance of hip and knee arthroplasties.

The sample size was determined by taking into account the number of observers. It has been suggested a sample size of 20 to 50 observations is needed to obtain an intraclass correlation coefficient (ICC) with a standard error less than 0.05 [38]. Others recommend, at 80% power and $\alpha = 0.05$, four raters would need to perform at least 30 observations each for a hypothesis test against a true reliability coefficient of 0.90 [11]. Consequently, a pragmatic decision was made to use 20 observations initially, as has been reported elsewhere [2, 33], and 35 observations at the second stage. For test-retest reliability, the observers reviewed the set of lesions twice with the recommended interval of 2 weeks [32, 38].

Initially, the reliability of the grid to measure area was tested using 20 simulated osteolytic lesions of varying size and shape (expansile and linear), represented as line drawings on white paper (Fig. 1). The images were reproduced in four booklets with a randomized order in each to reduce bias resulting from any learning effect. A simple description was provided to place the grid over the lesion and count the total number of crosses in the delineated area including any on the line. If no defined edge was provided, as with the simulation of some linear lesions, the individual observer decided the limit to mimic a clinical situation.

Additional booklets were produced with images in a new randomized order for repeat observations by two of the observers (RFS, LKS). There were no missing data and results from the measurements of the simulated lesions were analyzed before progressing to measurements on radiographs.

The second stage of testing was conducted on radiographs of joint arthroplasties with osteolytic lesions. There were 35 lesions around 27 joint arthroplasties, of which 11

Fig. 1 An example is shown of a simulated expansile lesion (used for initial testing of the morphometric grid) with maximum dimension of 2.0 × 0.9 cm. Observer ratings of the number of crosses on the morphometric grid were 29, 27, 29, and 32.
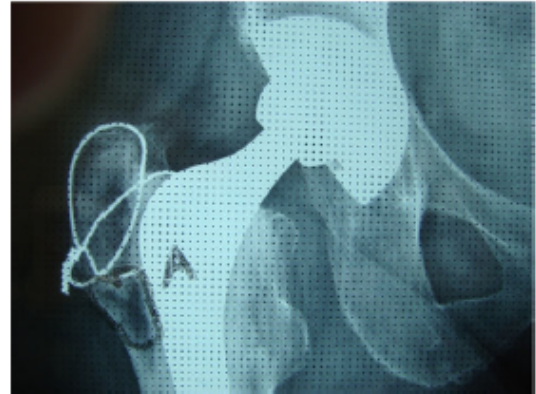

Fig. 2 The morphometric grid when superimposed on a radiograph for measurement of a femoral osteolytic lesion is shown. Observer ratings of the number of crosses for this lesion were 23, 21, 28 and 22.

were cemented and 16 were hybrid (cemented femoral stem, uncemented acetabular cup). The location of the osteolysis was in the pelvis in eight (23%) and the femur in 27 (77%). Twenty-eight (80%) of the lesions were expansile and seven (20%) were linear; 20 (57%) of them were associated with an uncemented component, and the remainder with cemented components. The joint arthroplasties had been in situ for 4 to 16 years (average, 8 years). The size of the lesions ranged from 0.04 cm^2 to 2.2 cm^2.

The method of testing the interobserver reliability of the morphometric grid was for each health professional to measure preidentified lesions. The lead author marked the general area but did not delineate the lesion precisely to allow individual interpretation of the size (Fig. 2). When more than one lesion existed on a radiograph, the lesion of interest was clearly identified. Not all lesions were included on any one radiograph as this study did not attempt to attribute clinical importance to the location or size of selected lesions, only to assess the reliability of the measurement tool.

Instructions were provided for placement of the grid, as before. An osteolytic lesion was defined as a clearly darkened area in which no trabeculae were visible compared with the adjacent bone or prosthesis. When composite shadow complicated the picture, observers were instructed to use the edge of the lesion as they estimated it to be regardless of shading. The observers were blinded to the identity of the patient and had no access to older images. Each observer recorded the number of crosses counted over each lesion and there were no missing data. The method of application of the morphometric grid is illustrated (Fig. 2).

Interobserver reliability was tested using an ICC, two-way random, absolute agreement method as the observers were representative of a larger population of potential users of the morphometric grid. Absolute agreement was used to

provide a measure of consistency of the tool and the interchangeability of observers [36, 38]. Intraobserver reliability (test-retest) was tested with an ICC using a one-way random effects model; the people effects are random, but the same sample set is measured on each occasion.

The Bland-Altman method also was used, as it has been suggested the use of two methods provides a better estimation of agreement [25, 32, 33, 36]. The Bland-Altman method usually is applied to two sets of results at a time. A graphic plot is produced to show the differences between each pair of measurements in relation to the mean of each pair. Limits of agreement (mean difference ± 2 standard deviations from the mean) are superimposed on this plot. It allows good visual representation of the reliability between two sets of observations [2]. The pairs of observers were chosen to reflect differences in profession and differences in experience in professional groups. We used SPSS® Version 13.0 software (SPSS Inc, Chicago, IL) for the statistical analysis.

Results

The morphometric grid was first tested by the four observers for measurement of simulated lesions. The results of the intraclass correlation coefficients for interobserver and test-retest reliability are summarized in Table 1. The graphic plots of these results show the mean close to zero in each case (Figs. 3–5). They also show a wider dispersion of the points with increasing lesion size. This indicates a tendency for the difference between observers to be greater for the larger lesions. It can be seen that the test-retest plot for Observer 1 (Fig. 4) shows a greater tendency for this dispersion than the plot for

Table 1. Results for interobserver and test-retest reliability on simulated lesions

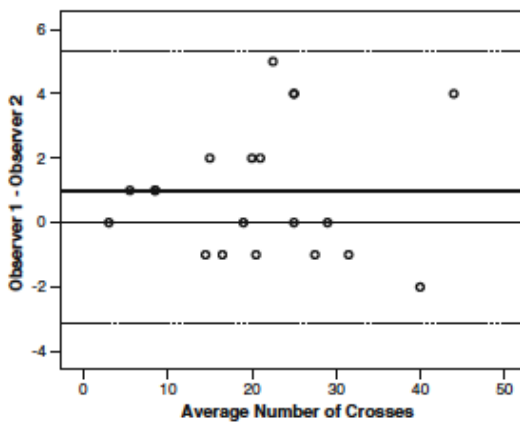| Reliability	Intraclass correlation coefficient		Bland-Altman				
	Coefficient	95% confidence interval	Range	Mean difference	95% confidence interval for mean difference	SD	95% limits of agreement
Interobserver	0.93	0.75–0.98	3–55	0.95	0–1.9	2.04	−3.13–5.03
Test-retest: Observer 1	0.97	0.93–0.99	3–55	−1.05	−2.29–0.19	2.65	−6.35–4.25
Test-retest: Observer 2	0.98	0.95–0.99	3–46	−0.45	−1.49–0.59	2.21	−4.87–3.97

Fig. 3 A distribution plot from a Bland and Altman test for interobserver reliability of the morphometric grid when used with 20 simulated lesions is shown. Mean of differences is shown as a solid line; 95% limits of agreement are shown as dotted lines. (There are only 18 points shown on the graph, as there were two sets of identical results.)
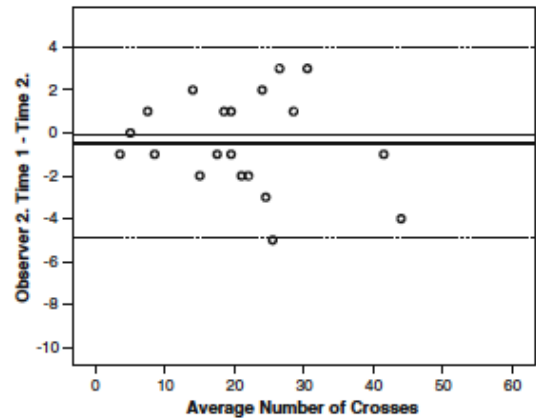

Fig. 5 A distribution plot from a Bland and Altman test for test-retest reliability of the morphometric grid when used with 20 simulated lesions (Observer 2) is shown. Mean of differences is shown as a solid line; 95% limits of agreement are shown as dotted lines.

Fig. 4 A distribution plot from a Bland and Altman test for test-retest reliability of the morphometric grid when used with 20 simulated lesions (Observer 1) is shown. Mean of differences is shown as a solid line; 95% limits of agreement are shown as dotted lines.
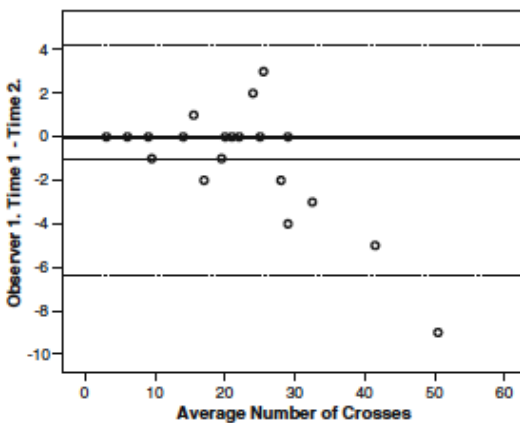
Table 2. Intraclass correlation coefficient results for interobserver reliability on radiographs

Interobserver reliability	Intraclass correlation coefficient	95% confidence interval
Time 1	0.96	0.93–0.98
Time 2	0.90	0.83–0.95

Observer 2 (Fig. 5), the latter being the more experienced observer of the two.

The reliability of the morphometric grid was subsequently tested on plain radiographs of hip replacements with osteolytic lesions in the surrounding bone. The four observers each recorded the size of 35 lesions and repeated the measurements 2 weeks later. The intraclass correlation coefficients between observers were 0.90 or greater on both occasions (Table 2). The Bland Altman method was used to assess reliability between four different pairs of observers (Observers 1 & 2, 3 & 4, 1 & 3, and 2 & 4;

Table 3. Bland-Altman results for interobserver reliability on radiographs at Time 1

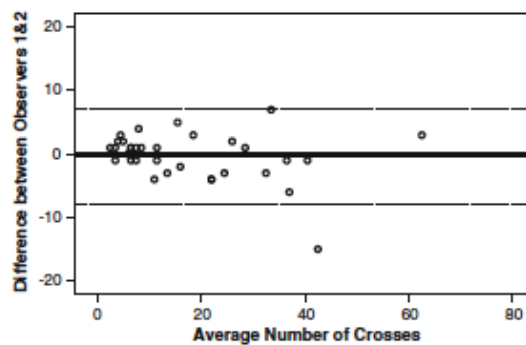| Observers	Bland-Altman				
	Range	Mean difference	95% confidence interval for mean difference	SD	95% limits of agreement
1 and 2	2–64	−0.31	−1.60–0.98	3.76	−7.83–7.21
3 and 4	1–71	−0.60	−2.30–1.10	4.95	−10.50–9.30
1 and 3 (physiotherapists)	2–63	1.31	0.02–2.60	3.76	−6.21–8.83
2 and 4 (surgeons)	1–71	1.03	−0.83–2.89	5.40	−9.77–11.83

Fig. 6 A distribution plot from a Bland and Altman test of interobserver reliability of the morphometric grid when used to obtain measurements on 35 radiographs of hip arthroplasties (Observers 1 and 2) is shown. Mean of differences is shown as a solid line; 95% limits of agreement are shown as dotted lines.

Table 3) and an illustration of one of the resulting plots is provided (Fig. 6). As before, the mean was close to zero in each case but there was a greater difference between observers and between repeat observations for larger lesions.

Discussion

Progressive osteolysis is a threat to the longevity of a hip arthroplasty and regular followup of patients with serial radiographs ensures changes are monitored even in the absence of symptoms [4, 19, 39, 43]. The measurement of the area of osteolysis on radiographs has not been achieved easily [9, 12]. Our aims were (1) to develop the concept of a morphometric grid to provide a simple tool for measuring the area of osteolytic lesions and (2) to test the reliability of this tool to measure osteolytic lesions seen on radiographic films when used by a representative range of observers.

We acknowledge limitations of our study. We used a two-dimensional tool to estimate the extent of a three-dimensional lesion. In other studies of the relationship between area and volume, one study reported no direct relationship [5], but another reported a correlation,

especially when a lesion was greater than 10 mL in volume [22]. We did not attempt to relate volume to area and additional testing would be required to investigate the association using the morphometric grid.

Additionally, we did not relate the size and location of lesions to clinical importance. Additional study would be required including relating the number of crosses on the grid to the actual area of the lesion being measured. Moreover, a study quantifying the crosses and location of lesions around hip replacements immediately before a revision would be useful in validating the grid for use in clinical diagnosis.

Our results suggest this simple tool can be used reliably to measure the area of osteolytic lesions in hip arthroplasty on radiographic film. A coefficient of reliability is a dimensionless property of an instrument applicable when the tool is used in a specified way. The value of the coefficient is context based. Values of 0.75 to 0.90 have been described as "good to excellent reliability" [36], but elsewhere it is suggested the value should be 0.90 or more for clinical use [38]. Our ICCs were all 0.90 or greater for a range of observers which indicates high reliability and is comparable with other orthopaedic radiologic studies [13, 14]. This is an acceptable level of agreement for use of the measurement obtained as one part of the decision-making process for a patient when recommending invasive treatment such as revision surgery. A higher coefficient of reliability would be required if this tool was the sole source of information on which such a decision was based.

Although the Bland-Altman plots indicated a greater difference between observers and between repeat observations for larger lesions, this may be explained by the greater number of crosses to be counted. It could be improved by providing a grid with increased spacing (> 2 mm) for larger lesions to optimize measurement by reducing the number of points to be included. It is of interest that the more experienced observer was more consistent in the measurement of size.

Although CT or MRI can provide three-dimensional assessment of the lesions, the cost, the risk to the patient, and the availability are barriers to their use in routine surveillance of patients with hip arthroplasties [22, 37, 40].

This simple, reliable tool requires no mathematical or software expertise and may be advantageous in a busy clinical situation when estimation of size is commonly made by visual analysis [31].

The interobserver reliability of this tool when used to measure osteolytic lesions identified on plain radiographs was 0.90 or greater (ICC), and on Bland-Altman plots, the mean was close to zero in each case, indicating no significant bias. The range of lesions measured represented commonly seen sizes, shapes (linear and expansile), and locations (pelvic and femoral). These results support the interchangeability of its use by any member of an orthopaedic team conducting routine surveillance. The purpose of surveillance is to facilitate timely intervention to improve the outcome for patients and potentially reduce the cost of surgery [23, 39, 43]. Additionally, reliable measurements obtained with the morphometric grid may provide evidence of the prevalence of osteolysis for specific acetabular components to assist a surgeon in his or her choice of prosthesis [39]. An essential component of this process is measurement of change over time with use of serial radiographs. Differences in magnification between radiographic films potentially could distort measurement of lesions made with the grid. However, the known diameter of the spherical femoral head provides a fixed measurement for calibration of films and the widely used digitized system for radiographs allows resizing of an image in the software provided. Transfer of the morphometric grid to a digital platform is being explored and it is recognized this must include consideration of measurement integrity when images are manipulated. The sensitivity of the grid to changes with time for use with progressive osteolysis must be investigated [30, 32], but should retrospective comparisons with plain radiographs be required, a formula (based on the size of the femoral head) can be used to compare the number of crosses observed (Appendix 1).

It is widely recognized there is a need for routine surveillance of patients who have had hip arthroplasties to monitor changes including development of asymptomatic osteolysis as a response to particulate wear debris. The use of radiographic review with AP and Judet oblique views allows assessment of pelvic osteolytic lesions, although estimates of volume usually require a CT scan or MR image. The routine assessment of area of osteolytic lesions on radiographic film often has been by visual analysis and the addition of a simple tool to quantify this assessment, developed from a morphometric grid, improves the reliability. This tool was reliable when used by a range of orthopaedic observers (medical and nonmedical health professionals) involved with hip arthroplasty review. It has potential to be incorporated in the assessment process to improve monitoring for patients who have had hip arthroplasties and may be a useful tool in future research studies.

Appendix 1

Formula to Adjust for Films of Differing Magnifications

If the morphometric grid is to be used on sequential plain radiographs of differing magnification, a simple formula can be used to adjust the number of crosses observed. For Film 1, the femoral head diameter = F1 and the size of the lesion on the grid = N1. For Film 2, the femoral head diameter = F2 and the size of the lesion on the grid = N2. The number of crosses for N1 then can be adjusted using the following formula: true number for N1 = $(F2/F1)^2 \times$ N1. This true number for N1 can be compared with the number of crosses recorded for the other lesion (N2) to assess any change in size.

References

1. Barrack RL, Folgueras A, Munn B, Tvetden D, Sharkey P. Pelvic lysis and polyethylene wear at 5–8 years in an uncemented total hip. *Clin Orthop Relat Res.* 1997;335:211–217.
2. Bland MJ, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1:307–310.
3. Bozic KJ, Kurtz SM, Lau E, Ong K, Vail TP, Berry DJ. The epidemiology of revision total hip arthroplasty in the United States. *J Bone Joint Surg Am.* 2009;91:128–133.
4. British Orthopaedic Association. Primary total hip replacement: a guide to good practice. 2006. Available at: http://boa.ac.uk. Accessed October 7, 2009.
5. Claus AM, Engh CA Jr, Sychterz CJ, Xenos JS, Orishimo KF, Engh CA Sr. Radiographic definition of pelvic osteolysis following total hip arthroplasty. *J Bone Joint Surg Am.* 2003;85:1519–1526.
6. Dall DM, Learmonth ID, Solomon M. A scoring system to determine radiological loosening in cemented total hip arthroplasty. *Int Orthop.* 1992;16:339–343.
7. Della Valle CJ, Paprosky WG. The femur in revision total hip arthroplasty evaluation and classification. *Clin Orthop Relat Res.* 2004;420:55–62.
8. Duffy PJ, Masri BA, Garbuz DS, Duncan CP. Evaluation of patients with pain following total hip replacement. *J Bone Joint Surg Am.* 2005;87:2566–2575.
9. Dumbleton JH, Manley MT, Edidin AA. A literature review of the association between wear rate and osteolysis in total hip arthroplasty. *J Arthroplasty.* 2002;17:649–661.
10. Elias H, Hennig A, Schwartz DE. Stereology: applications to biomedical research. *Physiol Rev.* 1971;51:158–200.
11. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Phys Ther.* 1994;74:777–788.
12. Geerdink CH, Grimm B, Vencken W, Heyligers IC, Tonino AJ. Cross-linked compared with historical polyethylene in THA: an 8-year clinical study. *Clin Orthop Relat Res.* 2009;467:979–984.

13. Ghelman B, Kepler CK, Lyman S, Della Valle AG. CT outperforms radiography for determination of acetabular cup version after THA. *Clin Orthop Relat Res.* 2009;467:2362–2370.

14. Gosvig KK, Jacobsen S, Palm H, Sonne-Holm S, Magnusson E. A new radiological index for assessing asphericity of the femoral head in cam impingement. *J Bone Joint Surg Br.* 2007;89:1309–1316.

15. Han CD, Choe WS, Yoo JH. Effect of polyethylene wear on osteolysis in cementless primary total hip arthroplasty: minimal 5-year follow-up study. *J Arthroplasty.* 1999;14:714–723.

16. Harris WH. Conquest of a worldwide human disease: particle-induced periprosthetic osteolysis. *Clin Orthop Relat Res.* 2004; 429:39–42.

17. Hernigou P, Zilber S, Filippini P, Poignard A. Ceramic-ceramic bearing decreases osteolysis: a 20-year study versus ceramic-polyethylene on the contralateral hip. *Clin Orthop Relat Res.* 2009;467:2274–2280.

18. Holt G, Murnaghan C, Reilly J, Meek RM. The biology of aseptic osteolysis. *Clin Orthop Relat Res.* 2007;460:240–252.

19. Hozack WJ, Mesa JJ, Carey C, Rothman RH. Relationship between polyethylene wear, pelvic osteolysis, and clinical symptomatology in patients with cementless acetabular components: a framework for decision making. *J Arthroplasty.* 1996;11: 769–772.

20. Jacobs CA, Christensen CP, Greenwald AS, McKellop H. Clinical performance of highly cross-linked polyethylenes in total hip arthroplasty. *J Bone Joint Surg Am.* 2007;89:2779–2786.

21. Karanicolas PJ, Bhandari M, Kreder H, Moroni A, Richardson M, Walter SD, Norman GR, Guyatt GH; Collaboration for Outcome Assessment in Surgical Trials (COAST) Musculoskeletal Group. Evaluating agreement: conducting a reliability study. *J Bone Joint Surg Am.* 2009;91(suppl 3):99–106.

22. Kitamura N, Pappedemos PC, Duffy PR III, Stepniewski AS, Hopper RH Jr, Engh CA Jr, Engh CA. The value of anteroposterior pelvic radiographs for evaluating pelvic osteolysis. *Clin Orthop Relat Res.* 2006;453:239–245.

23. Lavernia CJ. Cost-effectiveness of early surgical intervention in silent osteolysis. *J Arthroplasty* 1998;13:277–279.

24. Matthews JB, Besong AA, Green TR, Stone MH, Wroblewski BM, Fisher J, Ingham E. Evaluation of the response of primary human peripheral blood mononuclear phagocytes to challenge with in vitro generated clinically relevant UHMWPE particles of known size and dose. *J Biomed Mater Res.* 2000;52:296–307.

25. Morris RW. Assessing the reliability of clinical measurement. In: Pynsent PB, Fairbank JCT, Carr AJ, eds. *Assessment Methodology in Orthopaedics.* Oxford, UK: Butterworth-Heinemann; 1997:1–18.

26. National Joint Registry. 6th Annual report. 2009. National Joint Registry for England and Wales. Available at: http://www.njr centre.org.uk. Accessed October 7, 2009.

27. Needham J, Burns T, Gerlinger T. Catastrophic failure of ceramic-polyethylene bearing total hip arthroplasty. *J Arthroplasty.* 2008;23:627–630.

28. Nixon M, Taylor G, Sheldon P, Iqbal SJ, Harper W. Does bone quality predict loosening of cemented total hip replacements? *J Bone Joint Surg Br.* 2007;89:1303–1308.

29. Paterson D. The International Documentation and Evaluation System (IDES). *Orthopedics.* 1993;16:11–14.

30. Peat JK, Mellis C, Williams K, Xuan W. *Health Science Research: A Handbook of Quantitative Methods.* London, UK: Sage Publications; 2002.

31. Prevrhal S, Klifa CS, Shepherd JA, Ries MD, Genant HK. Computer-assisted technique to delineate osteolytic lesions around the femoral component in total hip arthroplasty. *J Arthroplasty.* 2008;23:833–838.

32. Pynsent PB, Fairbank JCT, Carr AJ. *Outcome Measures in Orthopaedics and Orthopaedic Trauma.* Ed 2. London, UK: Arnold; 2004.

33. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil.* 1998;12:187–199.

34. Roder C, Eggli S, Aebi M, Busato A. The validity of clinical examination in the diagnosis of loosening of components in total hip arthroplasty. *J Bone Joint Surg Br.* 2003;85:37–44.

35. Saleh KJ, Holtzman J, Gafni A, Saleh L, Davis A, Resig S, Gross AE. Reliability and intraoperative validity of preoperative assessment of standardized plain radiographs in predicting bone loss at revision hip surgery. *J Bone Joint Surg Am.* 2001;83: 1040–1046.

36. Sim J, Wright C. *Research in Health Care: Concepts, Designs and Methods.* Cheltenham, UK: Nelson Thornes; 2000.

37. Stamenkov R, Howie D, Taylor J, Findlay D, McGee M, Kourlis G, Carbone A, Burwell M. Measurement of bone defects adjacent to acetabular components of hip replacement. *Clin Orthop Relat Res.* 2003;412:117–124.

38. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use.* Ed 3. Oxford, UK: Oxford University Press; 2003.

39. Stulberg SD, Wixson RL, Adams AD, Hendrix RW, Bernfield JB. Monitoring pelvic osteolysis following total hip replacement surgery: an algorithm for surveillance. *J Bone Joint Surg Am.* 2002;84(suppl 2):116–122.

40. Thomas A, Epstein NJ, Stevens K, Goodman SB. Utility of Judet oblique x-rays in preoperative assessment of acetabular periprosthetic osteolysis: a preliminary study. *Am J Orthop (Belle Mead NJ).* 2007;36:E107–E110.

41. Walde TA, Weiland DE, Leung SB, Kitamura N, Sychterz CJ, Engh CA Jr, Claus AM, Potter HG, Engh CA Sr. Comparison of CT, MRI, and radiographs in assessing pelvic osteolysis: a cadaveric study. *Clin Orthop Relat Res.* 2005;437:138–144.

42. Wright JG, Feinstein AR. Improving the reliability of orthopaedic measurements. *J Bone Joint Surg Br.* 1992;74:287–291.

43. Wroblewski BM, Siney PD, Fleming PA. Charnley low-friction arthroplasty: survival patterns to 38 years. *J Bone Joint Surg Br.* 2007;89:1015–1018.

Appendix X

Oxford Hip Score

Problems with your hip

Tick (✓) <u>one</u> box for <u>every</u> question.

1. During the past 4 weeks...

How would you describe the pain you <u>usually</u> have from your hip?

None	Very mild	Mild	Moderate	Severe
☐	☐	☐	☐	☐

2. During the past 4 weeks...

Have you had any trouble with washing and drying yourself
(all over) <u>because of your hip</u>?

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐	☐	☐	☐	☐

3. During the past 4 weeks...

Have you had any trouble getting in and out of a car or using public
transport <u>because of your hip</u>? (whichever you tend to use)

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐	☐	☐	☐	☐

4. During the past 4 weeks...

Have you been able to put on a pair of socks, stockings or tights?

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐	☐	☐	☐	☐

5. During the past 4 weeks...

Could you do the household shopping <u>on your own</u>?

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐	☐	☐	☐	☐

6. During the past 4 weeks...

For how long have you been able to walk before <u>pain from your hip</u>
becomes severe? (with or without a stick)

No pain/More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house only	Not at all/pain severe on walking
☐	☐	☐	☐	☐

7. During the past 4 weeks...

Have you been able to climb a flight of stairs?

Yes, easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, impossible
☐	☐	☐	☐	☐

8. During the past 4 weeks...

After a meal (sat at a table), how painful has it been for you to stand up from a chair <u>because of your hip</u>?

Not at all painful	Slightly painful	Moderately painful	Very painful	Unbearable
☐	☐	☐	☐	☐

9. During the past 4 weeks...

Have you been limping when walking, <u>because of your hip</u>?

Rarely/ never	Sometimes, or just at first	Often, not just at first	Most of the time	All of the time
☐	☐	☐	☐	☐

10. During the past 4 weeks...

Have you had any sudden, severe pain - 'shooting', 'stabbing' or 'spasms' - <u>from the affected hip</u>?

No days	Only 1 or 2 days	Some days	Most days	Every day
☐	☐	☐	☐	☐

11. During the past 4 weeks...

How much has <u>pain from your hip</u> interfered with your usual work (including housework)?

Not at all	A little bit	Moderately	Greatly	Totally
☐	☐	☐	☐	☐

12. During the past 4 weeks...

Have you been troubled by <u>pain from your hip</u> in bed at night?

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
☐	☐	☐	☐	☐

Finally, please check back that you have answered each question.

Thank you very much.