

# **Recent advances in musculoskeletal imaging research**

Dr Andoni Paul Toms

A thesis submitted for the degree of  
**PhD by Publication**

School of Medicine  
University of East Anglia  
Norwich, England

June 2011

©This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

## **Abstract**

This thesis describes 10 years of research in musculoskeletal radiology focussing on radiological-anatomical studies, reliability studies, perfusion imaging and the imaging of joint replacements.

Digital imaging archives in modern radiology departments have allowed new approaches to anatomical population studies. Using these techniques the limits of normal vertebral angulation and acetabular morphology have been defined with implications for the detection of osteoporotic fractures, risk factors for osteoarthritis and planning for hip surgery.

Studies of radiological reliability are being published more frequently as the importance of reliability as part of the diagnostic performance of a test has become increasingly recognised. This thesis describes outcomes from reliability studies in conventional radiography, ultrasound, computed tomography and magnetic resonance imaging as well as in segmentation studies and the development of new grading systems.

Dynamic contrast-enhanced magnetic resonance imaging is an established technique for imaging the perfusion of tumours but there have been difficulties applying it in the musculoskeletal system. The neoangiogenesis stimulated by tumour growth can be confused with neoangiogenesis associated with normal repair mechanisms following successful chemotherapy, possibly because bone tumours occupy a space within a rigid structure, and the timing of these dynamic studies during the course of treatment is likely to be critical.

Orthopaedic prostheses have traditionally been contra-indications to magnetic resonance examinations because the susceptibility artefact associated with the magnetisation of the prosthesis overwhelms signal from the adjacent tissue. A newly defined histology disease called Aseptic Lymphocytic Vasculitis Associated Lesions required a new approach to imaging of hip replacements because conventional radiographs are typically



normal in this condition. Phantom studies have defined optimal imaging parameters for successful metal artefact reduction allowing the first radiological description of this disease, correlation with clinical and serological findings and the development of a grading system for categorising severity.

# Contents

<b>List of illustrations .....</b>	<b>7</b>
<b>List of tables .....</b>	<b>10</b>
<b>Acknowledgments .....</b>	<b>11</b>
<b>Introduction .....</b>	<b>13</b>
<b>1.0 Chapter 1: Anatomical studies .....</b>	<b>15</b>
<b>1.1 Background.....</b>	<b>15</b>
1.11 Anatomical studies.....	15
1.12 PACS .....	16
<b>1.20 Vertebral wedging .....</b>	<b>18</b>
1.21 Background.....	18
1.22 Materials and methods .....	19
1.23 Results .....	20
<b>1.30 Defining reference ranges for acetabular anatomy .....</b>	<b>25</b>
1.31 Background.....	25
1.32 Materials and Methods .....	27
1.33 Results .....	28
<b>1.40 Summary .....</b>	<b>32</b>
<b>2.0 Chapter 2: Reliability studies.....</b>	<b>34</b>
<b>2.1 Historical Perspective .....</b>	<b>34</b>
2.11 Definition .....	35
2.12 Statistical techniques .....	36
<b>2.2 Imaging of the femoral sulcus.....</b>	<b>38</b>
2.21 Background.....	38
2.22 Reliability.....	41
<b>2.3 Developing an MR grading system.....</b>	<b>45</b>
2.31 Background.....	45
2.32 Results .....	49
<b>2.4 Segmentation techniques.....</b>	<b>50</b>

2.41 Orthopaedic implant phantom study .....	52
2.42 Metatarsophalangeal joint study. ....	54
<b>2.5 PACS studies .....</b>	<b>56</b>
2.51 Conventional computerised radiography .....	57
2.52 Computed tomographic radiographs .....	61
<b>2.6 Summary .....</b>	<b>64</b>
<b>3.0 Chapter 3: Angiogenesis imaging .....</b>	<b>65</b>
<b>4.1 Background.....</b>	<b>65</b>
3.11 Angiogenesis .....	65
3.12 Pharmacokinetic modelling .....	65
3.13 Bone tumour imaging.....	68
<b>3.2 Limitations of Single Slice DCE-MRI .....</b>	<b>70</b>
3.21 Methods .....	70
3.22 Results .....	71
<b>3.3 Summary .....</b>	<b>77</b>
<b>4.0 Chapter 4: Imaging of joint replacements .....</b>	<b>78</b>
<b>4.1 Historical perspective .....</b>	<b>78</b>
4.11 Evolution of joint replacements .....	78
4.12 Mechanisms of joint replacement failure .....	79
4.13 Conventional imaging of joint replacements .....	80
<b>4.2 CT and MRI of joint replacements.....</b>	<b>82</b>
4.21 Computed tomography .....	82
4.22 Magnetic resonance imaging .....	85
<b>4.3 Phantom studies.....</b>	<b>87</b>
4.31 Computed tomography .....	87
4.32 Magnetic resonance imaging .....	90
<b>4.4 Clinical Studies.....</b>	<b>93</b>
4.41 Aseptic Lymphocytic Vasculitis-Associated Lesions (ALVAL) .....	93
4.42 First description of MR appearances of ALVAL .....	95
4.43 Characteristics of ALVAL in other MOM prostheses .....	99
4.44 Clinical outcomes.....	102
4.45 Asymptomatic patients and volunteers .....	103
4.46 Grading of severity of ALVAL with MR .....	105

<b>4.5 Summary .....</b>	<b>108</b>
<b>Conclusion.....</b>	<b>109</b>
<b>Appendix 1.....</b>	<b>110</b>
<b>Publications .....</b>	<b>110</b>
Chapter 1 .....	110
Chapter 2 .....	110
Chapter 3 .....	111
Chapter 4 .....	111
<b>Appendix 2.....</b>	<b>114</b>
<b>Published Abstracts.....</b>	<b>114</b>
<b>Scientific presentations .....</b>	<b>115</b>
Oral .....	115
Posters.....	116
<b>Definitions.....</b>	<b>117</b>
<b>Glossary.....</b>	<b>120</b>
<b>References.....</b>	<b>130</b>
<b>Index.....</b>	<b>145</b>

## List of illustrations

Figure 1. Lateral radiograph of the lumbar spine and a lateral CT radiograph from an abdominal CT examination. ....	17
Figure 2. Line diagram illustrating the method for defining the line of the superior vertebral endplate. ....	20
Figure 3. Diagrams illustrating geometric differences between a CT radiograph and a conventional radiograph of the lumbar spine. ....	21
Figure 4. Trigonometric analysis of parallax error in plain radiographic projections of the lumbar spine. ....	22
Figure 5. Extremes of acetabular morphology predisposing to OA. ....	26
Figure 6. Technique for measuring acetabular inclination and centre-edge-angle. ....	27
Figure 7. Diagrammatic representation of the sacro-coccygeal to pubis and foramen obturator index measurements. ....	28
Figure 8. Frequency histograms of acetabular inclination. ....	29
Figure 9. Diagrammatic representation of pelvic tilt in the sagittal plane measured on a lateral radiographic projection. ....	32
Figure 10. Frequency histogram of publication rates of inter-rater reliability studies in the five most highly cited general radiology journals ....	34
Figure 11. Diagrammatic representation of an axial section through the distal femur demonstrating the trochlea sulcus articulating with the patella. ....	40
Figure 12. Dot-plot illustrating of observer variability for observations of femoral sulcus angle measurements taken from hyaline cartilage. ....	44

Figure 13. Diagram illustrating the principle of growing regions of interest, starting at a seed point.....	51
Figure 14. Screenshot of thresholds defining background signal intensities and segmentation of “normal” background signal from phantom.....	53
Figure 15. Scatterplot illustrating less than perfect but strong correlation between the two observers measuring metal artefact. ....	54
Figure 16. Scatterplot demonstrating correlation of volume measurements in metatarsophalangeal joints of asymptomatic volunteers .....	55
Figure 17. Histogram of volumes of fluid in each metatarsophalangeal joint of nine volunteers.....	56
Figure 18. The inter-tear drop line, acetabular inclination and leg length .....	57
Figure 19. Lateral offset and stem angle. ....	58
Figure 20. Centre of rotation for a THR. ....	58
Figure 21. Scatterplot comparing measurements of acetabular inclination from two observers.....	59
Figure 22. A selection of cadaveric dry bones and the corresponding lateral CT radiograph. ....	62
Figure 23. Bland-Altman plot comparing vertebral endplate angulation obtained from CT and direct from cadaveric dry bones. ....	63
Figure 24. Diagrammatic representation of the two compartment pharmacokinetic model commonly used in DCE-MRI.....	66
Figure 25. Staging MR through the long axis of an osteosarcoma arising from the distal femur.....	69
Figure 26. Graph of the mean arterial slope and parametric maps of the distribution of $A$ , $K^{trans}$ and the EES after chemotherapy .....	71

Figure 27. Histogram of mean arterial slope (A) before and after chemotherapy. ....	75
Figure 28. Sagittal gradient echo T1W images taken from peak enhancement during the DCE-MRI acquisition.....	76
Figure 29. Line drawings representing the radiographic silhouette of three MOM arthroplasties.....	80
Figure 30. Series of axial images from a femoral stem phantom imaged with CT demonstrating “streak artefact”. ....	84
Figure 31. Line diagram demonstrating magnetisation of a femoral prosthesis.....	86
Figure 32. Line drawing representing a patient undergoing a CT examination with the gantry set at 0° and at -15°. ....	88
Figure 33. Area under the curves for standard deviations in background attenuation of a TKR with different gantry angles... ..	89
Figure 34. MR THR with metal artefact reduction.....	92
Figure 35. Intra-operative findings in ALVAL. ....	95
Figure 36. MR through a left MOM THR demonstrating a large fluid filled cavity, typical of the cystic mass found in ALVAL.. ....	97
Figure 37. Line drawing of the radiographic silhouettes of three second generation large bearing MOM THRs.. ....	99
Figure 38. MR of Birmingham MOM THR demonstrating a markedly hypointense nodule in the left gluteal compartment.....	100
Figure 39. MR images of metal particles in lymph nodes .....	101
Figure 40. Diagram of MR grading system for ALVAL.....	107

## List of Tables

Table 1. Table comparing published reference ranges for acetabular and centre-edge-angle with the results of this study .....	30
Table 2. Table for interpreting kappa correlation coefficients .....	37
Table 3: Table demonstrating the estimated variance from each potential source of variability for measurements of femoral sulcus angles.. .....	42
Table 4. Summary of the grading system for scoring MR findings of ALVAL .....	46
Table 5. 2x2 contingency tables for pairings of the three observers.....	48
Table 6. Summary of proportion of absolute agreement between observers for each category. ....	49
Table 7. A summary of the descriptive statistics of THR positional measurements.. .....	60
Table 8. Summary of the reliability of anatomical angular measurements derived from CT radiographs. ....	64
Table 9. Table summarising the differences in pharmacokinetic endpoint measures, from before and after treatment, for the two prognostic groups. .	72
Table 10. Criteria for grading ALVAL using MR.....	106



## Acknowledgments

The bulk of the published work on which this thesis is based is my own. My contribution to each of the publications is detailed in Appendix 1.

There are however notable contributions to the body of work from a number of friends and colleagues. This body of work started in 2001 at the Department of Medical Imaging in Mount Sinai Hospital, Toronto where Professor Larry White provided me with the design of the study into angiogenesis imaging in spindle sarcomas and then supervised me through the institutional review board approval, conduct, analysis and preparation of the final paper. Larry introduced me to musculoskeletal radiology research and continues to be a friend providing help and advice on subsequent research particularly in the imaging of orthopaedic prostheses in which he is an authority.

In all but one of the publications the statistical analysis of radiological data is my own. The exception is the analysis of Generalizability coefficients for evaluating the reliability of imaging of the femoral sulcus. This was performed by Dr Louise Swift at the University of East Anglia. I am grateful for her patience in the face of my ignorance.

The success of the publications that form the basis for chapter 4, Imaging of joint replacements, is in part due to enthusiastic support from the orthopaedic surgeons at the Norfolk & Norwich University Hospital, in particular Mr John Nolan, Mr Jim Wimhurst and Professor Simon Donell. They have provided the clinical context for the developments described in this thesis. They were also the senior authors on two clinical papers, on which I was a co-author, and which I have included in this thesis (Donell ST 2010, Wynn-Jones H 2011) because my contribution to these papers included substantial analysis of radiological data.

As supervisors Professor Alex MacGregor and Professor Lee Shepstone have provided enthusiasm, encouragement and support. Lee in particular has been a critical friend both as a co-author on two of the papers cited here, a teacher of statistics and as a steadfast guide through the construction of this thesis.

My friends, colleagues and co-authors Dr John Cahir and Dr Paul Malcolm never fail to help, without reservation, whenever asked, with even the most menial of research activities. Sharing the frustration of peer-review rejection and the satisfaction of successful publication with these two has made it fun and made it worth starting the next paper.

I am grateful to all of my colleagues in the Radiology department at the Norfolk & Norwich University Hospital and at the Radiology Academy for humouring, encouraging and supporting my research interests.

The final word, as always, goes to Katie, Cameron, Emily, Dougal and Angus without whom there is no point.

## Introduction

This thesis reflects some of the key developments that have taken place in imaging of the musculoskeletal system in the first decade of the 21<sup>st</sup> century. During this time radiology, in the UK, has completed its transition from an analogue to a digital discipline. With the widespread adoption of digital image file formats have come developments in storage, retrieval, networking and image analysis. These lie at the heart of much of the research described in this thesis; research which was started in 2001 and continued through to 2011.

The Norfolk & Norwich University Hospital (NNUH) was the largest film-free digital hospital in Europe when it opened in 2001 and has been archiving millions of clinical radiological examinations on computer servers ever since. This archive of clinical data forms a powerful database for cohort studies of both disease and normal radiological anatomy. However there are limitations when attempting to define normal anatomical profiles in populations using these archives. The clinical indications for the test in the first place introduce selection bias that means that by definition these patients do not represent a random sample of the local population. Careful methodological approaches can minimise this bias.

Increasing computing power has allowed complex mathematical models to be applied to dynamic datasets from magnetic resonance (MR) examinations each with over half a million data points. The results have exposed limitations in our understanding of some of the pathophysiological processes that control tumour growth, death and subsequent healing.

With the explosion of clinical radiology in terms of numbers of studies performed and the complexity of new imaging techniques has come a considerable increase in cost. With this increase in cost has come more scrutiny of radiology as an intervention in a patient pathway and the process

of assessing diagnostic performance has become more clearly defined. While the sensitivity and specificity of a radiological test have long been recognised as important measures of diagnostic accuracy the importance of the reliability of a test has only received substantial attention in the last ten years. A number of approaches to assessing reproducibility, in a variety of radiological modalities, in the musculoskeletal system will be described.

The impact of the ubiquitous application of computing is felt in other areas of medicine outside radiology. Computer numerical control has been part of the development new total hip replacements with extremely high machine tolerances that are aimed at reducing friction and increasing the longevity of the prostheses. However these prostheses are associated with the production of new atomic, ionic and particulate forms of metal with toxic effects on local soft tissue that have called for a new approach to imaging of total hip replacements. Where once a hip replacement was a contra-indication to examination with MR imaging new protocols have now been developed to minimise metal artefact and clinical indications have been clearly defined.

## **1.0 Chapter 1: Anatomical studies**

### **1.1 Background**

#### **1.11 Anatomical studies**

Anatomical studies are the foundation of radiological interpretation. Conventional radiographic studies in the musculoskeletal system have, in the past, typically consisted of analysis of radiographs of cadaveric specimens. In the case of the lumbar spine this might be a radiograph of a sagittal section through a cadaveric vertebral column which would produce an image similar to a lateral radiograph (1). The radiograph and cadaveric specimen could then be directly compared side-by-side. With the evolution of cross-sectional imaging procedures have changed. Cadaveric specimens are now imaged first and then dissected to obtain radiological-anatomical correlation (2).

One limitation of anatomical-radiological studies has been the difficulty in defining normal anatomical ranges of morphology. This is also true of pure anatomy studies which rely on cadavers that are almost always elderly and often have signs of disease. The risks associated with ionising radiation prohibit large scale x-ray studies of normal volunteers and therefore normal anatomical ranges of morphology are often derived from clinical data (3); in other words from patients who have presented with symptoms in the anatomical area of interest. This raises the real possibility of selection bias in that patients with a particular range of anatomical parameters may be more at risk of disease, such as degenerative disease, and therefore would be more likely to present for an imaging investigation, and as a result would be over-represented in any ranges of “normal” anatomy described by these cohorts. While ultrasound (US) and MR do not carry a risk of ionising

radiation they are relatively expensive and time consuming for routine use in large scale studies.

Cross-sectional imaging with computed tomography (CT) and MR offer new opportunities for the study of normal anatomical ranges. Although the clinical indication for a particular CT or MR study will usually be targeted to a specific organ or biological system other organs and systems are necessarily included in the study without being pertinent to the investigation. If the assumption holds that this information is incidental; that it is independent of the presenting complaint, then this provides an opportunity for large scale radiological-anatomical studies.

## 1.12 PACS

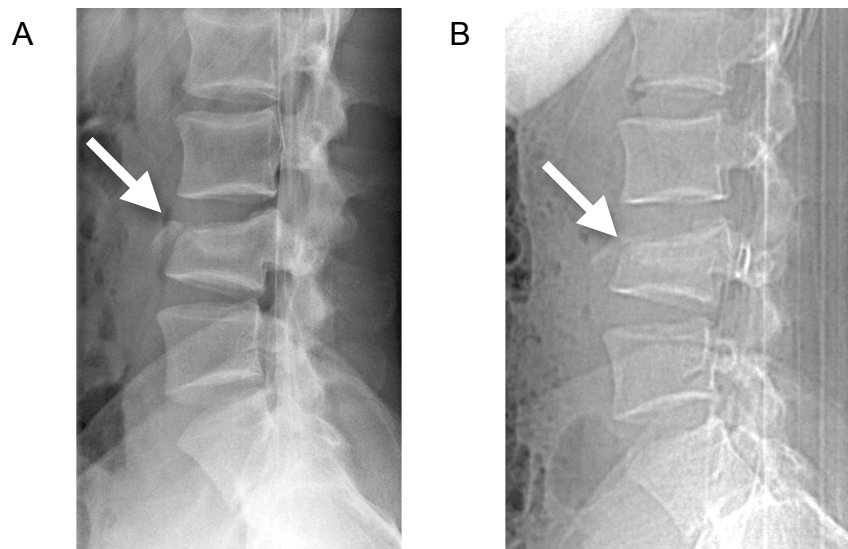
Picture Archiving and Communication Systems (PACS) are electronic archiving networks that store medical images in a standard image file format (DICOM<sup>1</sup>). These images can then be retrieved and reviewed simultaneously from workstations attached to that network (4). By 2008 127 hospital trusts in UK had had PACS successfully installed with measureable cost savings and improvements in clinical care (5). The NNUH was the first Trust in Europe to have a complete hospital wide PACS which was installed in 2001. A decade later the NNUH PACS archive contains over 2 million radiological studies. As a research resource PACS offers direct and indirect research opportunities. The main direct opportunity is the ability to easily construct datasets for pathological-anatomical research studies. Radiological examinations can be readily organised by age, sex and date from any standard PACS workstation.

CT examinations start with the acquisition of one or two preliminary low resolution images in a frontal and lateral projection. These images have

---

<sup>1</sup> Digital Imaging and Communications in Medicine is a standard that allows the creation, storage and transmission of medical images that allows communication between different manufacturers.

been described variously as scout views, tomograms or scanograms and are used to define the limits and orientation of the slices to be acquired. The term “scout view” is acceptable, the term “tomogram” is incorrect (the image is not that of a slice) and the term “scanogram” is meaningless. The images are a form of computed radiograph, which will be explained later, and will be referred to in this thesis as a CT radiograph. The images look similar to conventional radiographs, are not used as part of the normal radiological report, but come for free with the CT study (Figure 1). It is these images on which the following anatomical studies are based.



**Figure 1. Lateral radiograph of the lumbar spine (a) and a lateral CT radiograph from an abdominal CT examination (b).**

**Both demonstrate an unstable L3 vertebral fracture. While the trabecular detail is not as well demonstrated with the CT radiograph the cortical outlines are clearly demarcated.**

## **1.20 Vertebral wedging (6)**

### **1.21 Background**

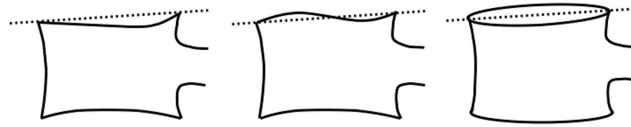
It is well recognised that as the lordosis of the lumbar spine transforms into the kyphosis of the thoracic spine, at the thoraco-lumbar junction, it is normal to find vertebral bodies that are “wedged” anteriorly. In particular it is normal to find, in asymptomatic individuals, that the anterior vertebral height of T12 and L1 is less than it is posteriorly (7). Unfortunately the thoraco-lumbar junction is also the most common site for osteoporotic (and traumatic) fractures of the spine (7,8) and therefore differentiating normal vertebral wedging from mild grade 1 osteoporotic fractures (9) can be difficult with conventional radiographs. Having a reference range for the normal morphology of the T12 and L1 vertebral bodies would be a useful tool with which to inform the reporting radiologist. There is published data describing the thoraco-lumbar vertebral morphology but it has a number of limitations. Early data was derived from small numbers of symptomatic patients or from cadaveric specimens, neither of which can be considered normal populations. In these publications the data were usually presented as ranges of anterior and posterior vertebral body height which are not easy to remember (7,10-13). There has been renewed interest in defining normal angular morphology of the spine and pelvis in the 2000s with several publications describing normal ranges for angles of lordosis of the lumbar spine and kyphosis of the thoracic spine and their relationship to sacral angulation (14-16). One group has also measured mean segmental vertebral angles from T9 to S1 but not published normal reference ranges for these (17). All of these studies have relied on recruiting asymptomatic volunteers with ages that have ranges from 18 to 70 and sample sizes ranging from n=34 to n=300. In one sample the majority of the recruits were healthcare workers and in others the sampling methods were not described.



The aim of this study was to define a reference range for normal vertebral angulation using the angle of the endplate on a lateral projection radiograph as the endpoint measure. The rationale being that the angle would be easier to measure than a ratio of anterior and posterior body heights and that referrals to CT might represent a better population sample than previous studies.

## 1.22 Materials and methods

200 consecutive CT studies on patients aged 25 (n=100) and 35 (n=100) (114 men and 86 women) were included in the study following exclusion of segmentation anomalies, trauma, diseases of the retroperitoneum or spine, and poor quality images. The lateral CT radiograph localiser was enlarged on a PACS workstation to approximately life-size. Two observers independently measured the endplate angle of the T12 and L1 vertebral bodies using the “Cobb angle” tool and the following landmarks. The superior “endplate” was defined as a line between the antero-superior and postero-superior corners of the vertebral body. The inferior endplate was similarly drawn between the inferior corners of the vertebra. The shape of the intervening endplate, be it flat, concave or S-shaped was ignored (Figure 2). The angle between the two lines was taken as the *endplate angle*. Measurements were repeated if the difference between the two observers was greater than 3°. This was to ensure that any differences were due to inter-rater variability and not because one or other observer had miscounted the vertebral levels. For the purpose of this study the data was assumed to conform to a normal distribution and therefore the normal range was defined as two standard deviations (SD) above and below the mean and so included 95% of the data.



**Figure 2.** Line diagram illustrating the method for defining the line of the superior endplate by connecting the antero-superior and postero-superior corners of the vertebral body. Three line drawings represent different projected outlines of a vertebral body as projected on the lateral CT radiograph. In each case the intervening course of the true endplate is ignored.

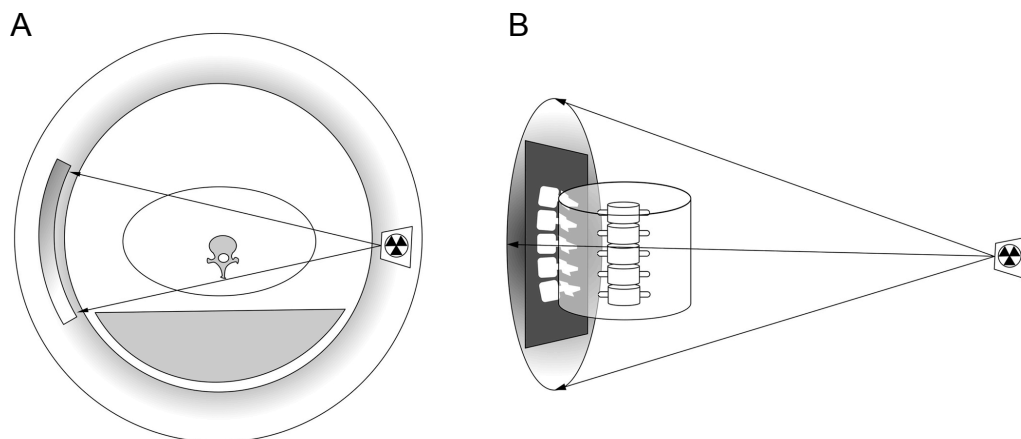
### 1.23 Results

The mean endplate angle for T12 was  $4.34^\circ$  (2 SD  $4.5^\circ$ ) and for L1 was  $4.48^\circ$  (2 SD  $4.26^\circ$ ). There was no statistically significant difference between the T12 and L1 levels, between the genders nor between the 25 and 35 year olds. Therefore reference ranges, defined by two standard deviations from the mean, were calculated as  $-0.2^\circ$  to  $8.8^\circ$  for T12 and  $0.2^\circ$  to  $8.7^\circ$  for L1.

The mean endplate angle of  $4.34^\circ$  was similar to the mean segmental angle for T12 previously published by Vialle et al but the endplate angle for L1 of  $4.48^\circ$  was substantially different to Vialle's  $3.8^\circ$ . Vialle did not publish standard deviations for these segmental angles for comparison.

In clinical practice it is the upper limit of vertebral endplate angulation that is the most important and, for practical purposes, this can be rounded up to  $9^\circ$ . In other words a T12 or L1 vertebral body with an endplate angle of more than  $9^\circ$  is either fractured or has a 1:40 chance of being normal.

However before  $9^\circ$  can be considered as the 95% upper limit of normal for routine reporting practice the differences between CT radiographs and conventional radiographs need to be evaluated. The geometry of the two techniques is very different. The CT radiograph uses an x-ray fan beam from a point source from which data is collected from a curved array of detectors approximately 70cm (industry standard gantry diameter) away. The spine is positioned midway between the two. Conventional radiographs use a cone beam positioned somewhere between 80-100cm from a radiographic plate with the spine positioned as close to the plate as possible (half the width of the patient for a lateral view). Clearly differences in parallax may mean that vertebral endplate angles may not be represented the same with the two techniques (Figure 3).

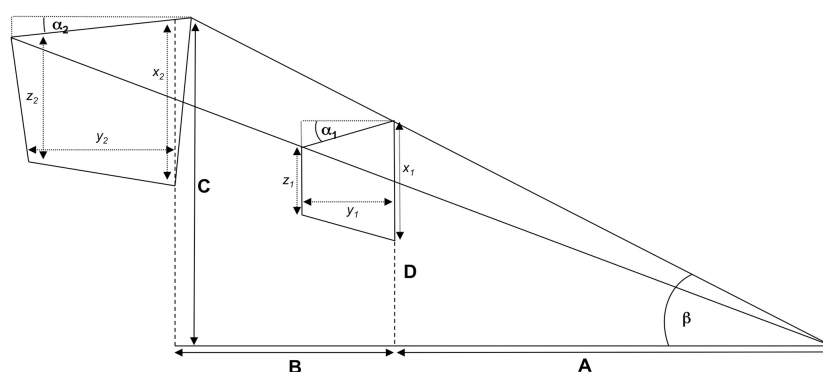


**Figure 3. Diagrams illustrating the geometric differences between a CT lateral radiograph (A) and a conventional lateral radiograph (B) of the lumbar spine. To acquire a lateral CT radiograph the x-ray tube is kept in a fixed position in the CT gantry and produces a fan beam from which data is collected by a detector array on the opposite side of the gantry. This arrangement is susceptible to parallax effects in one plane only (and these are corrected by the computed reconstruction of the image). The image is acquired by keeping the tube and detectors fixed and moving the patient through the gantry thereby collecting sequential lines of data to fill the image. In conventional radiography a**

**cone beam is produced by the x-ray tube and projected on to a radiographic plate producing magnification of the subject and parallax effects in three dimensions around the centre of the cone.**

The effect of parallax, in clinical practice, on the projection of the endplates is most marked at the thoraco-lumbar junction because radiographs are most frequently performed as either lumbar or thoracic views centred on L3 or T6 respectively and therefore the T12 and L1 vertebra are projected at the edge of the radiographic plate, where parallax is greatest, on both of these views.

To estimate the likely maximal effect of this parallax error on endplate angle measurements a trigonometric exercise was performed. The change in the angle of the endplate for a hypothetical vertebral body measuring 4.5cm antero-posterior and 3.5cm cranio-caudal, with an endplate angle of  $9^\circ$ , positioned 20 cm from a 42x30 cm radiographic plate and projected on to the margins of the plate by a cone beam with a film-focus distance of 100cm gives rise to a *reduction* in endplate angle of approximately  $0.25^\circ$  (Figure 4). It appears that any angular distortion due to parallax is minimal compared to changes due to magnification and therefore  $9^\circ$  seems like a reasonable upper limit for the 95% reference range.



**Figure 4. Diagrammatic representation of the basis for the trigonometric estimate of the effect of magnification and parallax distortion of a thoracolumbar vertebra projected onto the edge of a radiographic plate 42cm (2C) in vertical height. With a film-focus**

distance of 100 cm ( $A + B$ ), the x-ray beam at edge of film subtends angle of  $11.9^\circ$  ( $\beta$ ) to horizontal [ $\tan^{-1} (21/100)$ ]. A vertebral body with endplate angle of  $9^\circ$  ( $2\alpha_1$ ), measuring 3.5 cm in posterior body height ( $x_1$ ) and 4.5 cm in anteroposterior width ( $y_1$ ) is positioned 20 cm ( $A$ ) from the film. From this the following can be estimated using simple trigonometry. Position of the postero-superior corner of the vertebra from central beam ( $D$ ) is 16.8 cm, the anterior body height ( $z_1$ ) equals 3.15 cm, and superior and inferior endplate angles ( $\alpha_1$ ) measure  $4.5^\circ$ . The posterior vertebral body height ( $x_2$ ) of the projected image increases to 4.4 cm, anterior body height increases to 3.5cm, superior and inferior endplate angles ( $\alpha_2$ ) decrease to  $4.35^\circ$  and  $4.39^\circ$ , respectively.

A limitation of the published results was that the standard error of the mean, for the vertebral endplate angles at T12 and L1, was not included and therefore readers would not have been able to assess the precision of the sample as a representation of the population from which it was derived. In fact the 95% confidence levels for the means were  $\pm 0.47^\circ$  at T12 and  $\pm 0.45^\circ$  at L1 for women, and  $\pm 0.43^\circ$  and  $\pm 0.40^\circ$  respectively for men. This suggests that the true mean for endplate angles is likely to lie within  $0.9^\circ$  of the published results.

The confidence intervals are close to  $\pm 10\%$  of the mean. A sample size calculation for descriptive statistics, based on the mean and standard deviation for T12, indicates that for a statistical power of 0.95 and 10% confidence intervals a sample size of  $N=77$  is required. Reducing the confidence intervals to  $\pm 5\%$  would increase the required sample size to  $N=451$  (18).

Although this provides some reason to be confident that the sample is a good representation of the population the suitability of the population then needs to be addressed. The patients undergoing CT are clearly not a random sample of the general population. Each patient has a clinical indication for a CT and therefore either has a set of symptoms or a known

disease that warrant investigation. For the results of this study to be generally applicable the assumption has to be made that the indications for CT were unconnected to the vertebral morphology of the thoraco-lumbar spine and therefore the spread of data from a sample of patients presenting for CT would represent a sample of the general population. Apart from designing the exclusion criteria to ensure that disease of the vertebral column was not included in the sample there is no definite way of confirming these assumptions.

There is also the potential for selection bias in that a number of the CT examinations were excluded because the images were of too poor quality to identify the necessary anatomical landmarks clearly. This is likely, but not confirmed, to select against those patients with the largest abdominal girth. If the assumption is that vertebral endplate angles are independent of body mass index, and in young patients that is probably the case (although this is not known), then this should not matter and the results can be considered generally applicable.

The application of the results in everyday clinical radiological practice needs to be considered. The paper describes the results as defining a *normal* range. Although this is statistically acceptable terminology it does define the range of vertebral endplate angles that should be reported as normal. One in 40 patients will have a vertebral endplate angle of more than 9° and for them this will be normal and not indicate an underlying disease process. A patient may have an osteoporotic fracture, such as a central endplate depression or crush fracture, that does not change the endplate angle and therefore would be in the *normal* range. The term *upper limit reference* may have been a more suitable description for the 9° measure. The reporting radiologist may then use this to raise the suspicion of an osteoporotic fracture, in the appropriate clinical setting, however with 1 in 20 patients falling outside this reference range this could result in two or three patients being unnecessarily investigated following each plain film reporting list. It may also have been more useful to publish the 3SD (three standard

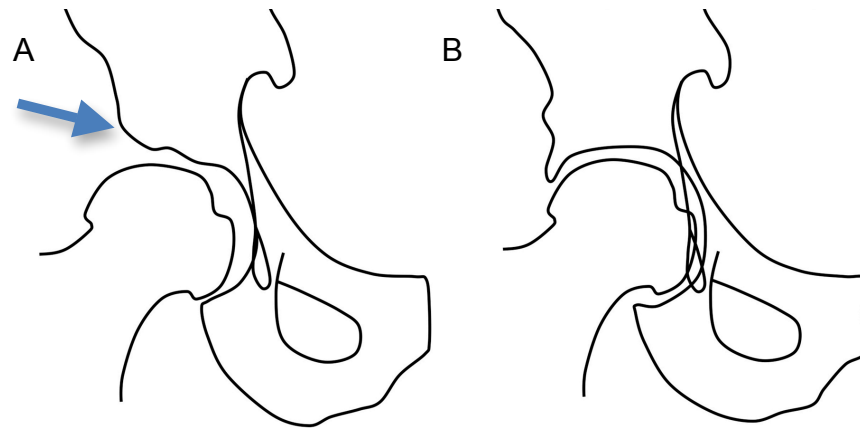
deviation) upper of  $11.3^{\circ}$  above which only 1 in 400 patients might be considered to be otherwise normal.

### ***1.30 Defining reference ranges for acetabular anatomy (19)***

#### **1.31 Background**

It is clear that acetabular morphology is one of the factors that can contribute to premature osteoarthritis (OA). This association was first reported in adults who had had a slipped upper femoral epiphysis (SUFE) during childhood. During a SUFE there is no direct damage to the articular cartilage but the resulting incongruence in the joint and reactive modelling abnormalities in the acetabulum result in suboptimal loading of the cartilage. This manifests as a patient presenting in their 20s or 30s with premature OA (20).

Developmental dysplasia of the hip (DDH) is also associated with premature OA. The most common form is a mild isolated hypoplastic acetabular roof with limited lateral coverage of the femoral head. This typically presents with symptoms, due to degenerative changes, in women in their 40s and 50s (21). There are also two main forms of femoro-acetabular impingement (FAI) that can cause early hip OA which are caused by cam-type and pincer-type developmental deformities. The cam-type is a developmental deformity of the femoral head, and is more common in men, and the pincer type is a developmental deformity of the medial-to-lateral position of the femoral head resulting in increased lateral coverage, and is more common in women (22,23)(Figure 5). Grading of FAI can be used to predict the likelihood of subsequent OA (24-26).



**Figure 5. Line drawing illustrating extremes of acetabular morphology that predispose to early OA; acetabular dysplasia (a) with lateral uncovering of the femoral head and protrusion acetabuli resulting in a pincer-type over hanging supero-lateral acetabular margin (b).**

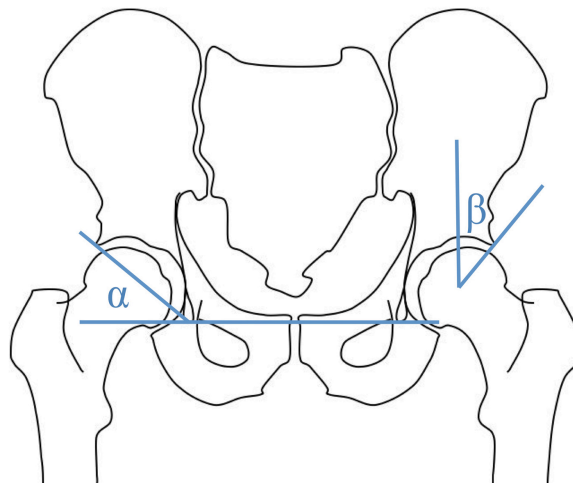
Despite this clear association between acetabular morphology and the risk of OA the normal range of radiographic measurements that are commonly used to describe the acetabulum has been imperfectly known. The restrictions around irradiating the pelvis of young asymptomatic volunteers are clear. Therefore reference data has come from cadaveric (27-29) and synthetic pelvic (30) studies and clinical radiographic material predominantly of males over the age of 40 (31,32).

The two most commonly used radiographic measures of acetabular morphology have been acetabular inclination (AI) and the centre-edge-angle (CEA). “Normal” ranges for these values have been quoted variously as ranging from 33 to 42 for AI (31,32), and 20 to 40 for CEA (31) but the descriptive statistics are unclear, not uniform and sometimes contradictory.

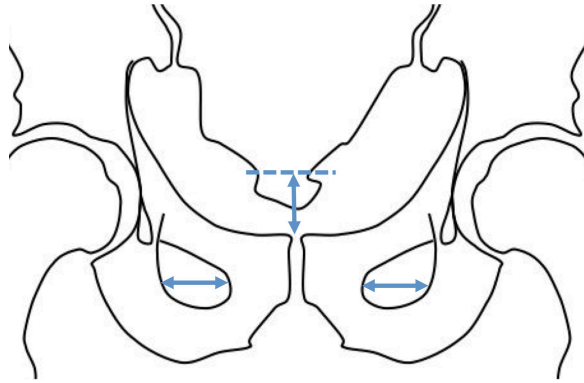


### 1.32 Materials and Methods

To derive new reference ranges for AI and CEA 100 AP CT radiographs of the pelvis (50 men and 50 women: mean age 26 years) were included in this study. For each CT radiograph two observers independently measured the AI and CEA of each hip (Figure 6). The vertical distance between the sacro-coccygeal joint and the pubic symphysis (SCP) was recorded as a measure of pelvic tilt in the sagittal plane, and the foramen obturator index (FOI) was recorded as a measure of pelvic tilt in the axial plane (Figure 7).



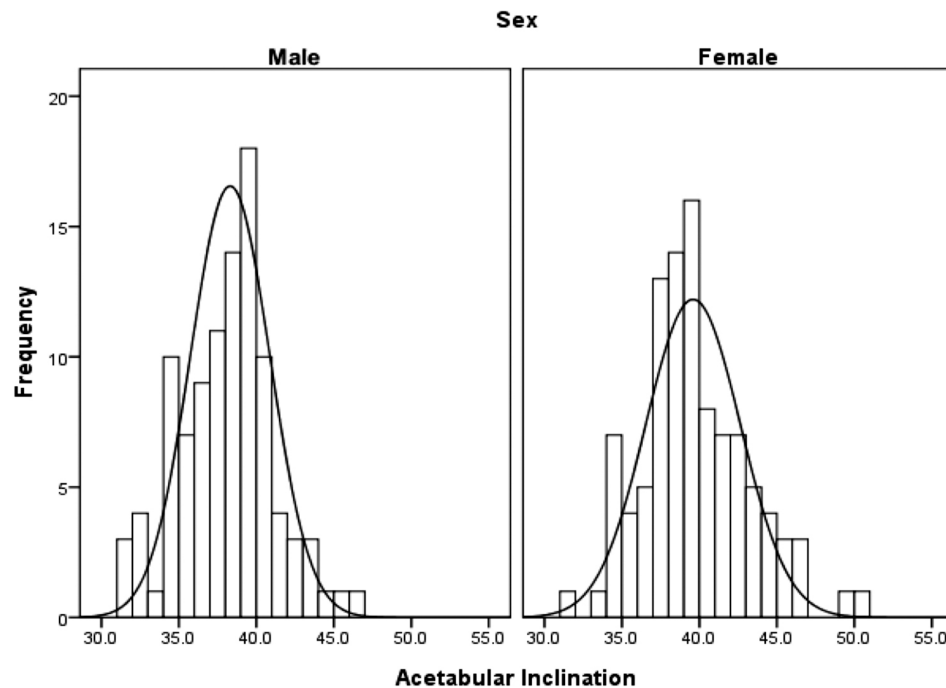
**Figure 6.** Diagrammatic illustration of the technique for measuring (a) acetabular inclination ( $\alpha$ ) and centre-edge-angle ( $\beta$ ). The AI is taken as the angle subtended by the inter-tear drop line and a line drawn from the supero-lateral margin of the acetabulum and the point where the inferior margin of the tear-drop meets the inter-tear drop line. The CEA is taken as the angle subtended by a vertical line drawn from the centre of the femoral head and a second drawn from the centre of the femoral head to the supero-lateral margin of the acetabulum.



**Figure 7. Diagrammatic representation of the sacro-coccygeal to pubis (SCP) and foramen obturator index (FOI) measurements. FOI is calculated as the ratio of the maximal transverse diameters of the two obturator foramina.**

### 1.33 Results

The mean AI was  $38.0^{\circ}$  (2 SD  $31.8^{\circ}$  -  $44.1^{\circ}$ ) for men and  $39.6^{\circ}$  (2 SD  $32.7^{\circ}$  -  $46.8^{\circ}$ ) for women. The mean CEA was  $37.7^{\circ}$  (2 SD  $26.9^{\circ}$  -  $48.5^{\circ}$ ) for men and  $34.9^{\circ}$  (2 SD  $23.5^{\circ}$  -  $46.3^{\circ}$ ) for women. The differences between men and women were statistically significant for both measures confirming that women tend to have a shallower acetabular roof and less lateral acetabular cover than men. This is also reflected by the upper and lower limits of the normal statistical ranges (Figure 8).



**Figure 8. Frequency histograms of acetabular inclination (AI) in men and women.**

Pelvic tilt in the sagittal plane also differed between men and women with men having more vertically positioned pelvises. Again the range of measures demonstrated substantial overlap between men and women. The data on sagittal pelvic tilt was limited to 48 of the 100 samples because the AP CT radiograph was not deemed adequate to identify the necessary landmarks by either or both of the observers. This is a substantial loss of data that means that larger patients may have been under-represented by these results. A Pearson correlation coefficient demonstrated almost no linear correlation between sagittal pelvic tilt and either AI or CEA.

FOI was a more robust measure in that it was confidently obtained in 83 out of 100 of the AP CT radiographs. These demonstrated that for men and women there was no systematic bias to a right or left sided pelvic tilt (a theoretical source of bias if patients always get on to the CT table from the same side) and that the variance of FOI was the same for both sexes. Once again there was no linear correlation between FOI and either AI or CEA.

The results of this study demonstrate that the normal reference range is apparently broader than previously suggested with higher upper limits of both CEA and AI (Table 1).

Source	Acetabular Inclination	Centre-edge-angle
Historical	25°-40° <sup>1</sup>	20°-40° <sup>2</sup>
	33°-38° <sup>3</sup>	
New <sup>4</sup>	32.1°-45.5°	24.9°-47.8°

<sup>1</sup> Stulberg *et al* 1974 (32)

<sup>2</sup> Armbuster *et al* 1978 (31)

<sup>3</sup> Sharp 1961 (33)

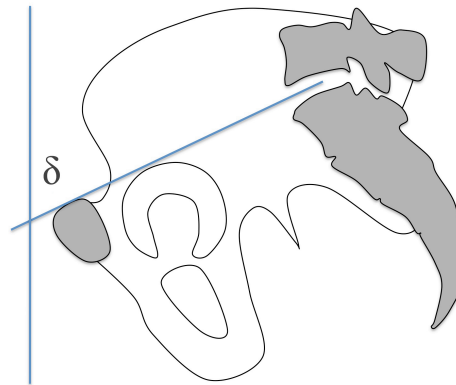
<sup>4</sup> Fowkes *et al* 2011 (19)

**Table 1. Table comparing published reference ranges for acetabular and centre-edge-angle with the results of this study.**

There are two broad reasons why the reference ranges from this study may differ from previously published data. The first set of reasons relate to the design of the original studies including sample selection, measurement technique and the format of descriptive statistics employed. The datasets used were derived from older patients, some with known degenerative disease. The measurements were performed using chinagraf pens, rulers and hand-held goniometers or protractors and the simple ranges of data were quoted. The second main reason is that the population of patients attending NNUH radiology for CT examinations may have different acetabular anatomy to other populations with different socioeconomic, racial and genetic backgrounds. If this is the case then this data may not be widely applicable but the technique of harvesting CT radiographic metrics described in this paper would allow samples to be easily obtained in different communities. The reliability measurements were *very good to excellent* for all the results described (See page 37 for interpretation of reliability). This

was somewhat of a surprise because it might be predicted that the CEA measurements would be less reliable than AI measurements because the anatomical landmarks that define the measure are less discrete.

One area of concern is the limited number of measurements obtained for pelvic tilt. Both the SCP and FOI measurements were secondary endpoint measures and were not considered carefully when selecting the 100 CT examinations. The rationale for this was that there is some evidence to suggest that the position of the pelvis in a supine patient is a good representation of the position of the pelvis when the patient is upright (34). In retrospect it would probably have been more robust to exclude all AP CT radiographs that were not considered suitable at the beginning of the study. There is currently plenty of interest in standardising anatomical approaches to the measurement of acetabular inclination and version, particularly with CT in the orthopaedic literature (35,36), and therefore more accurate measurements of pelvic tilt would have allowed us to defend or reject the validity of the supine pelvic position in CT for this study. In fact the SCP is probably not the optimal measure of sagittal pelvic tilt; it should probably be performed as an angular measure of a line drawn from the lumbo-sacral junction to the pubis on a lateral view (37,38)(Figure 9). This was attempted in this study but abandoned because the low dose lateral CT radiograph simply did not allow adequate visualisation of these anatomical landmarks.



**Figure 9. Diagrammatic representation of pelvic tilt angle ( $\delta$ ) in the sagittal plane measured, on a lateral radiographic projection, as the angle between a line drawn from the lumbo-sacral junction to the superior margin of the pubis and a vertical baseline.**

For many modern hip prostheses the position of the cup is critical to its long term survival. In recent years an enormous amount of effort has gone in to developing sophisticated surgical techniques for spatial orientation of the acetabular cup which are designed by manufacturers to operate in a known “normal” range of positions. Yet these known “normal” ranges are derived from data that is 30-50 years old, that is not derived from “normal” populations and certainly may not be applicable to other populations worldwide. This study provides arguably the best (although not without its limitations) reference range published so far and illustrates a mechanism for generating reference ranges in other populations.

### **1.40 Summary**

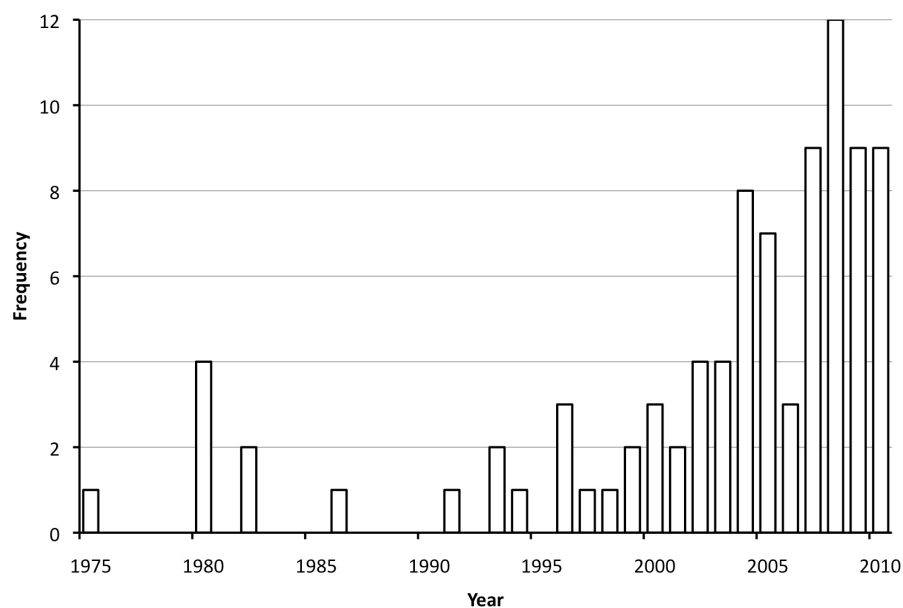
The two radiological-anatomical papers described in this chapter illustrate, for the first time, how PACS can be used as a research resource for population based anatomical studies, using imaging data that is incidental to

the primary diagnostic study. A reference range for normal anterior wedging of the thoraco-lumbar vertebrae has been defined. This is the first reference range that is simple to use in clinical practice with modern DICOM workstations and will help inform the analysis and reporting of osteoporotic vertebral fractures. Decades old radiographic reference ranges for acetabular inclination and centre-edge-angle, that define acetabular morphology, have been updated. Not only have the reference ranges been revised, with a particular increase in the upper limit of both measurements, but the sample size and selection result in an increase in confidence in the accuracy of this data compared to previously published results.

## 2.0 Chapter 2: Reliability studies

### 2.1 Historical Perspective

The concept and definition of the reliability of a test has been well established in the psychological sciences since the 1960s. It has only been appreciated as a concept in the biomedical sciences, in general, relatively recently and in radiological research in the last decade (Figure 10). Until the 2000s the term *reliability* was often used to report the validity of a test and, unlike the psychological sciences literature, a variety of terms still confuse the descriptions of this statistical concept, the most common being *reproducibility* (39).



**Figure 10. Frequency histogram demonstrating the rise in publication rates of inter-rater reliability studies in the five most highly cited general radiology journals (PubMed: American Journal of Roentgenology, British Journal of Radiology, Radiology, European Radiology, Clinical Radiology)**



## 2.11 Definition

Reliability is defined, in statistics, as the consistency of a test. The test consisting of repeated measurements that require an observation often from an instrument. The calculations of reliability described below all extend from Classical Test Theory which is based on the simple assumption that any measurement comprises two parts: the true measurement (which can never be known) and the error. Both the observer and the instrument can contribute to random, and systematic, errors to the measurements obtained. These errors are inversely proportional to the reliability of the test. For a test to be reliable the variance (or error) of the observations needs to be sufficiently small in comparison to the variability in the subject being measured that variance between subjects is not overwhelmed by the variance of the measurements. Therefore the reliability ( $R$ ) of a test is expressed as a proportion of variance in the subject population ( $\sigma_s^2$ ) and the variance of the error in the recorded measurements ( $\sigma_e^2$ ) (39).

$$R = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

As the definition of reliability has been consolidated in the radiological literature so the role of reliability studies in the evaluation of radiological tests has been more clearly described. In the hierarchy of diagnostic tests they are grouped with other tests of diagnostic performance such as sensitivity, specificity, accuracy, positive and negative predictive values and receiver operator characteristics (40,41). The standards for publishing the results of these tests are now clearly defined by the STARD<sup>2</sup> criteria so that reported data can be compared with other published results (42).

---

<sup>2</sup> Standards for the Reporting of Diagnostic Accuracy Studies

## 2.12 Statistical techniques

The three main categories of reliability studies are inter-rater, intra-rater (test-retest) and inter-method estimates. Inter-rater estimates test the reliability between observers using the same method to acquire measurements. Intra-rater reliability refers to the consistency of repeated measurements on the same subject taken by one observer and inter-method reliability is the comparison of consistency of observations taken from the same subject using different techniques.

The standard statistic for reliability (for continuous data from two or more observers measuring the same subject) is the *intraclass* correlation (43,44) so called to distinguish it from *interclass* correlation of which the most commonly used calculation is the Pearson product-moment correlation. The Pearson correlation measures the degree of linear association between two different measurements and has the potential to over-estimate an association between two sets of observations on the same data. Intraclass correlation uses pooled, rather than separate, measures of variance for both datasets giving a theoretically more accurate, if not always practical, measure of reliability. There are a number of different types of intraclass correlation coefficients (ICC) (45). For radiologists there are two important subtypes; those that estimate *consistency* and those that measure *absolute agreement*. A measure of *consistency* will provide a measure of reliability for the observations in that particular study setting i.e. for those two or more observers. This will provide a measure of support for a particular set of research findings. A measure of *absolute agreement* will calculate the reliability of a test for any observer and therefore this would be appropriate for methodologies that are being proposed as universal techniques such as grading systems. Predictably correlation coefficients for *absolute agreement* are always lower than for *consistency*.

Categorical data is handled differently. Levels of agreement can be measured using contingency tables but, particularly with binomial datasets, if

the proportion of any one outcome is high or low the level of agreement will be high by chance alone. Cohen's kappa coefficient of agreement resolves this by including a factor that reflects the distribution of data and this is the standard tool for measuring reliability in categorical data (46). However when the categorical data is ordinal disagreements between observers can vary in magnitude and a standard kappa correlation will tend to underestimate reliability. In these situations a modified *weighted* version of the kappa coefficient takes account of the magnitude of disagreement (47).

Interpreting the significance of resulting correlation coefficients is not without controversy. Landis and Koch proposed a guide for interpreting kappa correlation coefficients which has been criticised for not being supported by evidence (48). However this, and variants of this table, are commonly used to interpret the significance of kappa and ICC statistics (48-50) (Table 2).

K	Interpretation of level of agreement		
	Landis & Koch (48)	Altman (49)	Shrout (50)
0.0 – 0.20	Slight	Poor	Virtually none
0.21 – 0.40	Fair	Fair	Slight
0.41 – 0.60	Moderate	Moderate	Fair
0.61 – 0.80	Substantial	Good	Moderate
0.81 – 1.00	Almost perfect	Very good	Substantial

**Table 2. Table for interpreting kappa correlation coefficients (48-50)**

There are two main limitations of reliability tests. The first is that they are defined by the population that they are calculated from ( $\sigma_s$ ) and therefore are not necessarily applicable to other populations. The second is that there is no clear consensus on how good a correlation or agreement is good enough

although some authorities would consider anything much lower than 0.5 as poor (49) and anything over 0.8 as good for basic science research (51). A second approach is to calculate a *repeatability coefficient* ( $2s_d$ ) which can be defined as the *limits of agreement* between two sets of observations. By assuming a normal distribution of differences about a mean ( $d$ ) upper and lower limits that include 95% of all differences can be calculated and presented in a Bland-Altman plot (52). Although this is a common approach to the problem of repeatability it is by no means universally accepted as the only approach (39).

The calculations of reliability described so far are based on an assumption of a single source of error. However this assumption does not allow for all possible sources of variance. These might include variance between the left and right side of the body or the time of day that the observation was made. Generalizability theory, also known as G theory (GT), is an approach that aims to combine all possible sources of variance in a study which then allows the investigator to examine individual relationships between subjects and all sources of error (39,53). The resulting generalizability or G coefficient can range from 0 to 1 and be interpreted in the same way as correlation coefficients.

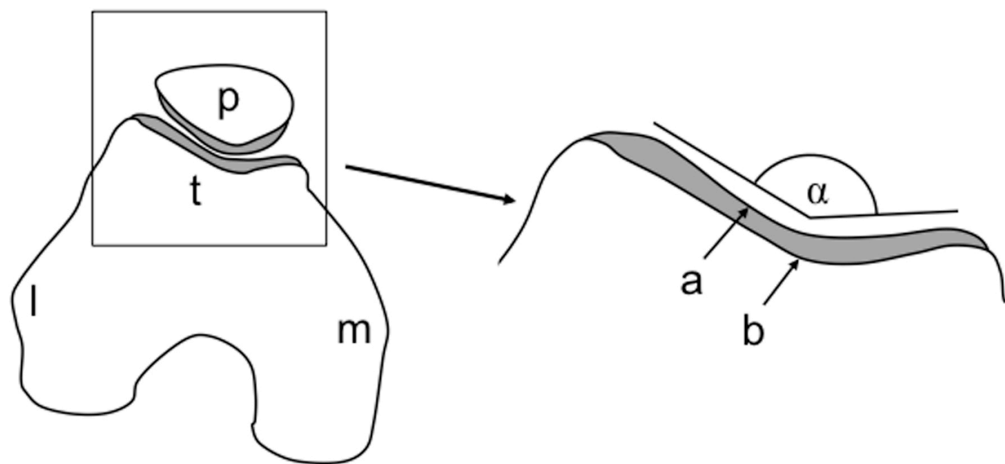
## ***2.2 Imaging of the femoral sulcus***

### **2.21 Background**

Trochlear dysplasia is a developmental condition where the femoral sulcus (trochlea) fails to develop normally. The normal sulcus should develop between the medial and lateral condyles as a V-shaped groove that forms an angle of between approximately 120° and 160° at its base (54). Failure of normal development may result in a small (hypoplastic) or absent (aplastic) medial condyle and flattening or convexity of the sulcus (dysplasia). All of these may be associated with patellofemoral disorders, in particular patella instability.

There are two broad suggestions of how the dysplasia may develop. The first is that the final morphology of the distal femur is entirely genetically pre-determined. The second is that a more minor genetically coded anatomical variant, such as a lateralised insertion of the patella tendon on the tibia, may secondarily influence the development of the sulcus. If the modelling of a normal femoral sulcus is dependent on the position that a patella articulates with the distal femur then any congenital abnormality that causes the patella to track an abnormal course may lead to trochlear dysplasia. There is some evidence that correcting the patella tracking before puberty will influence modelling of the trochlea (55,56).

In order to investigate these hypotheses, and the effect of intervention, a reliable method for measuring the angle of the femoral sulcus is needed. Conventional radiography can demonstrate the femoral sulcus angle with a *skyline* view (Figure 11) which is projected along the cranio-caudal axis of the patellofemoral joint with the knee in 20° of flexion. The angle of knee flexion is critical because the trochlea angle varies continuously along its cranio-caudal length. Historically this technique has been considered to be too unreliable for serial studies because the variability in patient positioning has such a significant influence on the trochlea angle measurement. As a result computed tomography (CT) has become the standard for assessing trochlea morphology for more than 20 years. It is easier to localise the cross-sectional anatomy and standardise patient positioning and as a result is considered a reliable tool for measuring the trochlea (57-59). CT has a limited ability to demonstrate articular cartilage and carries a small radiation burden. Both of these make it less than ideal for longitudinal studies in children and adolescents where articular cartilage morphology may be more important than bone morphology (60).



**Figure 11. Diagrammatic representation of an axial section through the distal femur at the level of the medial (m) and lateral (l) epicondyles demonstrating the trochlea (t) sulcus articulating with the patella (p). The femoral sulcus angle ( $\alpha$ ) was measured at both articular surface (a) and at the bone (b) cartilage interface.**

US and MR both have the advantage of superior cartilage imaging and neither carry a radiation burden. MRI should be as reliable as CT when it comes to errors relating to the position of the patient and identifying anatomical landmarks but this is not supported by any published literature to date. Children may not tolerate an MRI as easily as CT, US or conventional radiography, because the acquisition time is relatively long and the bore of the magnet can cause claustrophobia. US has been used to demonstrate trochlea cartilage with a high degree of accuracy (61-64) but there have been no assessments of patient related errors which might be expected to be similar to those encountered in conventional radiography.

## 2.22 Reliability

Twenty four patients presenting to a specialist patella instability outpatient clinic were recruited in to the study and each underwent a CT (the standard of care) of both knees as well as MR and US. US was performed twice at the initial visit by two independent operators in order to measure inter-rater reliability. The patients returned for a third US examination performed by one of the original observers in order to assess intra-rater reliability. Similarly the femoral sulcus angle was measured from the recorded CT, MR and US images twice by one observer and once by the second observer.

The assessment of reliability was performed using classical intra-class correlation coefficients but because there were clearly a number of potential sources of error generalizability coefficients (G) were also calculated. The *facets* (term used to describe elements associated with error in GT) that were considered were the patient, observer, time, and the laterality of the knee being examined with G being calculated as

$$G = \frac{Pf + Kf}{\Sigma f}$$

where *Pf* is the patient facet, *Kf* is the knee facet and  $\Sigma f$  is the sum of all facets (65).

The results indicated that the most reliable technique for measuring the trochlea angle was CT when the angle was measured from subchondral bone (Table 3). In other words the standard of care was the most reliable with G and ICC for inter-observer reliability of 0.87. However MR and US were very similar with *substantial* to *almost perfect* G and ICC and therefore there was no clear advantage of one technique over another. As discussed earlier, the osseous femoral sulcus angle may be very different to the cartilaginous sulcus angle. In which case the osseous sulcus angle may not

be a mechanically relevant measurement, and therefore the reliability of this measure may not be relevant.

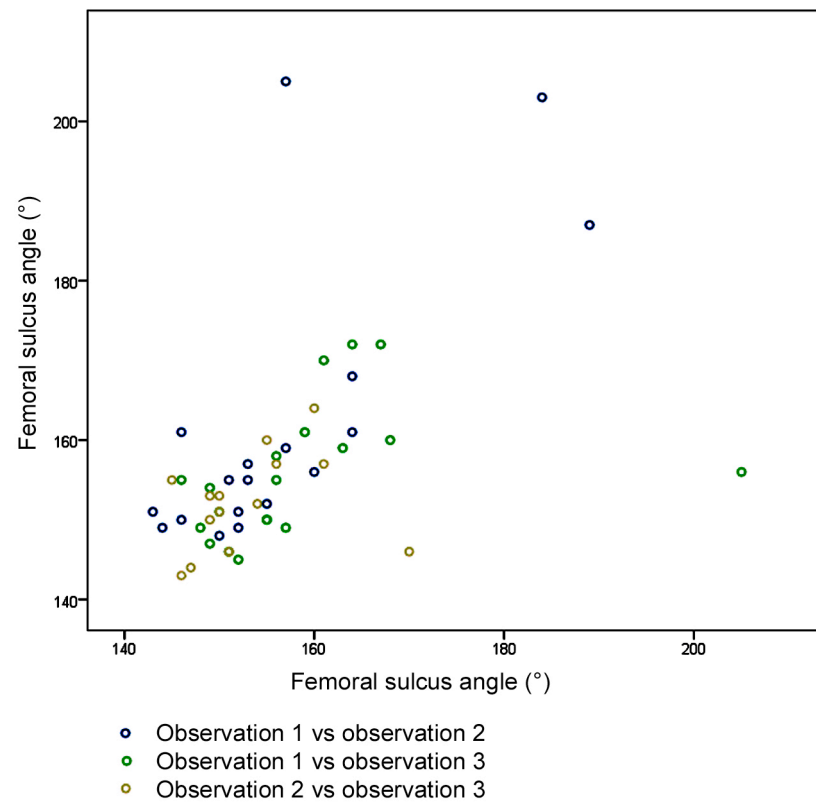
	CT		MR		US	
	Bone	Cartilage	Bone	Cartilage	Bone	Cartilage
Components						
Patient	294.0	262.4	213.1	223.7	151.6	0
Knee	84.5	41.6	47.8	40.3	15.5	9.2
Time	0	0	0	0	0	0
Observer	0.6	0.3	1.0	0	0	0
Patient x time	0	2.6	14.6	24.6	3.7	75.3
Patient x observer	13.7	29.9	0	11.0	3.5	57.4
Knee x time	0	13.0	6.1	3.6	0	0
Knee x observer	6.3	1.0	42.7	23.0	0	10.7
Time x observer	0	0.14	0	0	0	5.6
Error	36.5	47.7	19.5	36.0	31.9	12.1
Coefficients						
Inter-observer	0.87	0.80	0.82	0.81	0.83	0.50
Intra-observer	0.92	0.84	0.88	0.82	0.83	0.45
Generalizability coefficient	0.87	0.76	0.76	0.73	0.81	0.05

**Table 3: demonstrates the estimated variance from each potential source of variability for CT, MR and US measurements of femoral sulcus angles. The variance of US measurements from cartilage calculated using GT for *patient x time* and *patient x observer* is particularly high and accounts for the overall poor reliability indicated in the bottom three rows. The ICC were calculated in a similar manner to G. The denominator for each was the sum of all facets. The numerator was the same but without observer related facets for inter-observer ICC and without time related facets for intra-observer ICC.**



When the femoral sulcus measurements from articular cartilage (arguably the more important measure) were compared MR and CT produced almost identical G and ICC in the region of 0.8. However US had a *moderate* ICC for inter-rater reliability but almost no generalizable reliability with  $G = 0.05$  (Table 3). When estimates of variance were performed to identify the source of error in the ultrasound measurements from cartilage the error was most strongly associated with the facets *patient x time* and *patient x observer*. In other words the patient effect on error varies substantially from time to time and from observer to observer making it an unreliable tool.

The answer to the research question seems to be clear. Despite previous reports that US is a reliable tool for assessing the articular cartilage of the femoral sulcus this does not appear to be the case in patients with trochlea dysplasia. Why US should prove to be reliable when measuring angles from subchondral bone and not cartilage is unclear. A standardised protocol for positioning of the US probe perpendicular to the distal femur at the level of the femoral epicondyles was used and the operators appear to have applied this technique reasonably well if the substantial ICC for osseous sulcus angle measurements are to be believed. The results did suggest that there might be an association between the trochlea angle and inter-rater error although this was not formally calculated (Figure 12). It seems likely that the surface morphology of the articular cartilage is more susceptible to small differences in probe position than that of the subchondral bone but we do not understand the geometric reasons for this.



**Figure 12. Dot-plot illustrating of observer variability for three sets of observations of femoral sulcus angle measurements taken from hyaline cartilage with US (blue and yellow circles represent inter-observer comparisons and the green circles represent intra-observer comparisons). The largest discrepancies appear to be associated with the few largest measurements of the sulcus angle.**

While the results seem clear there are two main limitations in the general applicability of these findings. The first is the nature of the measurement itself. A single angular measurement may only represent the shape of the sulcus in one plane. Craniad or caudad there may be morphological changes, that contribute to patella instability, that do not register their importance. Mathematical surface modelling may in the future provide a more relevant, albeit more complex, measure of mechanical function than a simple sulcus angle measurement. The second limitation arises from the rapid technological progress in CT and MR (although this is a limitation in the

relevance of the results rather than study design). The CT protocol in this study comprised a limited set of axial slices; the minimum required to image the necessary anatomical landmarks and keep the radiation dose as low as possible. The protocol broadly mirrored the protocol that had been the standard for 20 years and therefore was appropriate for the study. Recently there have been significant advances in dose-reduction technologies that may allow us to exploit multi-slice CT in order to produce surface shaded 3-dimensional models of the sulcus. Improvements in MR machines and receiver coils combined with new cartilage specific sequences offer a wealth of new approaches to imaging the sulcus including dynamic MR of the patellofemoral joint.

## **2.3 Developing an MR grading system (66)**

### **2.31 Background**

There are many image based grading systems that allow observers to score the severity of diseases. These are most common in the area of oncological imaging (67) where the primary data type is usually categorical or ordinal, although there may be elements of continuous measurements, such as lymph node diameters, that determine which category the observer scores for a particular case.

After describing the MR features of a newly defined soft tissue disease (ALVAL<sup>3</sup>) associated with metal-on-metal hip replacements (described in detail in Chapter IV) a number of population studies were performed. These required an assessment of the severity of MR findings which could then be correlated with other clinical and radiological scores, and haematological and histological markers of disease. To do this we devised an MR grading system which was based on surgical decisions to treat at NNUH. The

---

<sup>3</sup> ALVAL is characterised by peri-prosthetic necrosis, tendon avulsion, cavity formation, marrow replacement and myopathy.

surgeons would consider observing patients if they were asymptomatic, or had mild symptoms related to their MOM THR, if MR demonstrated disease confined to the immediate peri-prosthetic soft tissues with no other complicating features. If the disease was of a greater volume with involvement of the adjacent muscles, such as the gluteals or ilio-psoas, they would consider revising the hip electively. If the disease extended in to the subcutaneous fat or there were tendon avulsions or obvious bone involvement this would merit an urgent revision. The MR correlates of these three treatment options were categorised as mild, moderate or severe features of ALVAL. These grades also had to be distinguished from normal post-operative appearances and infection (the most common other complication of hip replacement surgery seen with MR).

Grade	Severity
A	Normal
B	Infection
C1	Mild ALVAL
C2	Moderate ALVAL
C3	Severe ALVAL

**Table 4. Summary of the grading system for scoring MR findings of ALVAL** (See page 106)

To do this 73 MR examinations were selected from databases by a single researcher with the aim of placing roughly equal numbers of cases in to each of the five categories. The full DICOM datasets were anonymised using software built in-house that strips out all patient identifiable data from the file header (68). These were then replaced by a randomised study number (Random number sequence from Randomnumber.org (69)). Three observers (Observers 1 and 2 were experienced at reporting MR examinations of total hip replacements (THR) and observer 3 had almost no experience but was a senior musculoskeletal radiologist) were then asked to

score each of the MR studies independently, blinded to history, previous radiological studies and outcomes, using the scoring system described in Table 4 and Table 10 (page 106). Two way contingency tables for each pairing of observers were drawn up and weighted kappa coefficients of agreement were calculated (Table 5).

	Observer 1					
Observer 2	A	B	C1	C2	C3	Total
A	19	5	0	0	0	24
B	0	1	1	0	0	2
C1	1	0	3	2	1	7
C2	0	0	1	15	6	22
C3	0	3	0	1	14	18
Total	20	9	5	18	21	73
	Observer 2					
Observer 3	A	B	C1	C2	C3	Total
A	16	0	0	1	0	17
B	1	1	1	1	0	4
C1	5	0	5	4	1	15
C2	2	0	1	11	2	16
C3	0	1	0	5	15	21
Total	24	2	7	22	18	73
	Observer 1					
Observer 3	A	B	C1	C2	C3	Total
A	16	0	1	0	0	17
B	1	0	2	1	0	4
C1	3	4	1	5	2	15
C2	0	3	1	10	2	16
C3	0	2	0	2	17	21
Total	20	9	5	18	21	73

**Table 5. 2x2 contingency tables for pairings of the three observers.**

		Complete agreement between observers	
		Observers 1 and 2	All 3 observers
Grade	Frequency (%)	Proportion (%)	Proportion (%)
A	28	93	74
B	6	23	0
C1	12	33	11
C2	26	80	59
C3	28	77	63

**Table 6. Summary of proportion of absolute agreement between observers for each category.**

### 2.32 Results

Weighted Kappa coefficients of agreement varied from  $\kappa=0.66 - 0.78$  (95% confidence intervals: 0.54-0.88) with the strongest agreement ( $\kappa=0.78$ ) occurring between the two most experienced observers (1 and 2); each having seen or reported at least 300 cases. Agreement between these and the least experienced (observer 3) was less strong but still suggested substantial agreement (48). The contingency tables indicate that agreement was best for normal studies, and in patients with moderate or severe ALVAL. Absolute agreement for infection or mild ALVAL was limited (Tables 5 and 6).

The first limitation of this study becomes apparent when examining the contingency and summary tables. The ideal design would include roughly equal numbers of cases in each category whereas the proportion of cases in category B and C1 were relatively small. At the stage that this study was

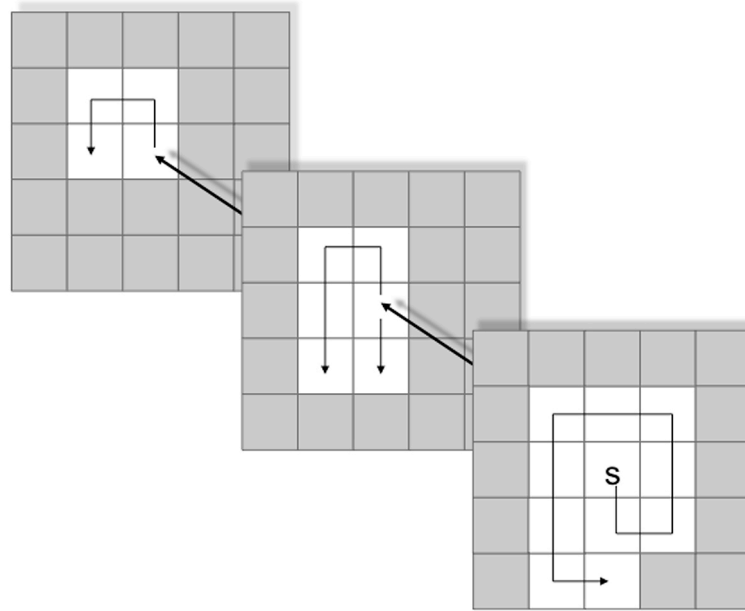
performed this was unavoidable because we had few examples of hip MR that we could be confident were normal (these were obtained from another study of asymptomatic hip replacements (70)). We also had few examples of infections of hip replacements because MR is not usually required in the diagnosis or management of these cases. Overall the study probably had reasonably good precision because the 95% confidence intervals for each observer pairing was narrow. However it could be argued that the weighted kappa coefficient used has over-estimated reliability of this grading system. The categorical scores A, C1, C2, and C3 are ordinal but the category B is nominal and therefore the data is mixed. It may have been more realistic to present kappa coefficients as well, which would presumably have demonstrated lower correlations, and report that the likely reliability of the grading system lies somewhere between these two sets of correlation statistics.

## **2.4 Segmentation techniques (71,72)**

In computerised analysis of radiological images *segmentation* describes the process of partitioning an image into multiple segments with the aim of simplifying and then quantifying properties of those segments. The most common technique for segmenting an image is to apply a seed point to a representative part of the image somewhere close to the “centre” of the required segment. This is done by an observer subjectively choosing an appropriate area and then clicking a cursor which selects a single pixel. The numerical value of that pixel e.g. Hounsfield units in CT or signal intensity in MR becomes the starting point (seed) from which the algorithm “grows” a region of interest (ROI). The operator will select two threshold values, one above and one below the numerical value of the seed point. The computer program will then examine each neighbouring pixel and if the numerical value of that pixel falls within the threshold range it is included in the segment. The process is repeated with neighbouring pixels around the newly formed segment until it reaches pixels that fall outside the threshold range and the segment stops growing. This process can be performed



within the two dimensions of the images and propagated through adjacent slices and is then known as semi-automated volume segmentation (Figure 13).

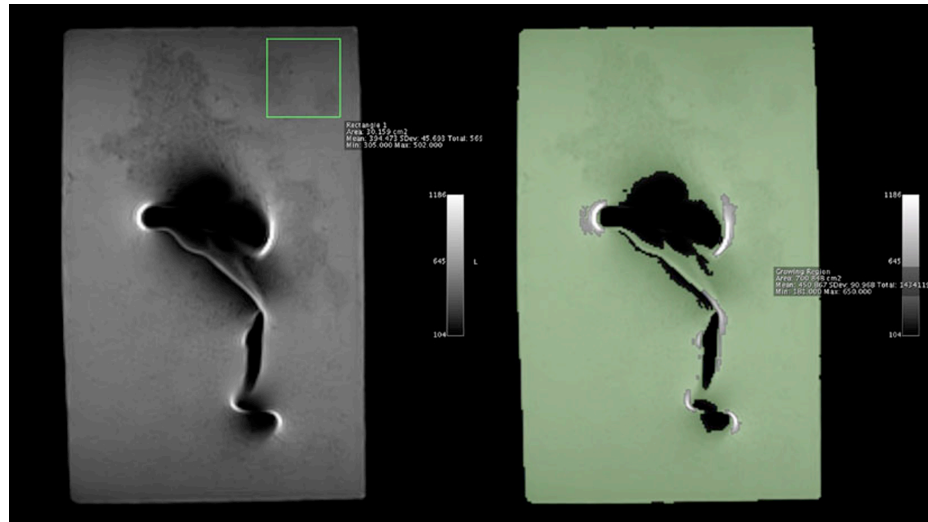


**Figure 13. Diagram illustrating the principle of growing regions of interest, starting at a seed point (s), from which is propagated within the slice, and between slices, an algorithm that assigns all pixels falling within prescribed thresholds to the final segment.**

Semi-automated segmentation is being used increasingly for measuring cross-sectional areas and volumes of target lesions, with some degree of validation, but no published reports of the factors that affect the reliability of this type of approach. There are two main sources of potential difference between observations when segmenting volumes. These are variations in selection of the seed point and the setting of threshold levels. Other geometric parameters, such as voxel size, slice thickness and inter-slice gaps, that define the image probably have little influence.

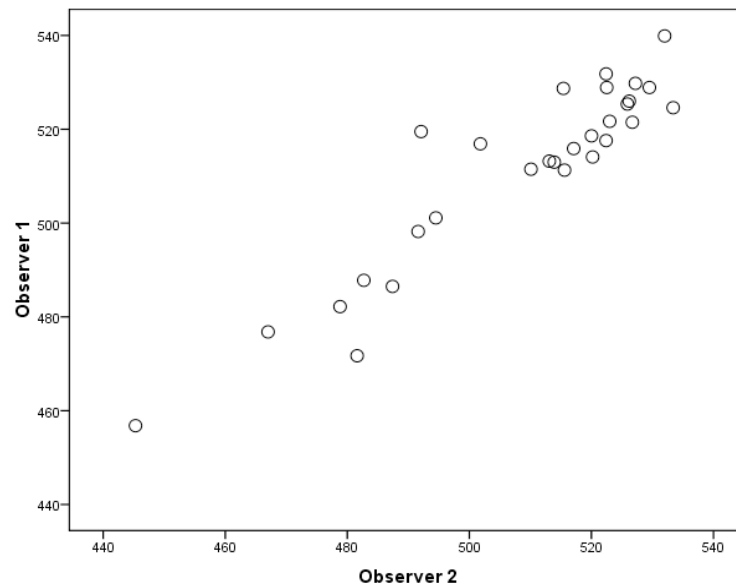
### 2.41 Orthopaedic implant phantom study (71)

Segmentation was used in a study designed to evaluate the optimal MR imaging parameters for reducing metal artefact from THR which is discussed further in Chapter IV. A phantom was built from a THR embedded in solid cooking fat. A number of different media can be used for building radiological phantoms including water, oil, gelatine and parts of butchered animals. Water, oil and gelatine have the advantage of producing very homogeneous background signal and animal parts can be used to mimic the range of signals, from muscle, bone marrow and fat that might be found in a human limb. Solid cooking fat was used for two reasons; the first was a concern that a magnetised THR might move within a gelatine or liquid based medium and the second was that inserting the prosthesis into an animal bone might introduce air, which itself produces a signal void, might produce artefacts unrelated to the THR itself and so complicate the analysis. The phantom was then imaged with MR to investigate the influence of voxel size and receiver bandwidth on the severity of metal artefact associated with the THR. The metal artefact manifests predominantly as a signal void (black areas on the image where signal is absent intensity). This is caused by magnetisation of the THR whose own magnetic field causes rapid dephasing of the protons in its vicinity. At some points the magnetic field of the THR can summate with the  $B_0$  magnetic field and result in areas of hyperintense artefact (Figure 14). To measure the amount of artefact produced by each MR sequence the background signal was segmented to exclude the signal voids and peaks, produced by the artefact, and subtracted from the total cross-sectional area of the phantom. The thresholds for defining the “normal” background signal were defined by selecting an ROI in one corner of the phantom not affected by metal artefact and using the mean and standard deviation of signal intensities within that ROI to determine thresholds.



**Figure 14. Screenshot from workstation demonstrating selection of thresholds using a square ROI to define the normal range of background signal intensities (left hand image) and subsequent segmentation of “normal” background signal (fat) from the phantom (shaded green in right hand image) containing a THR.**

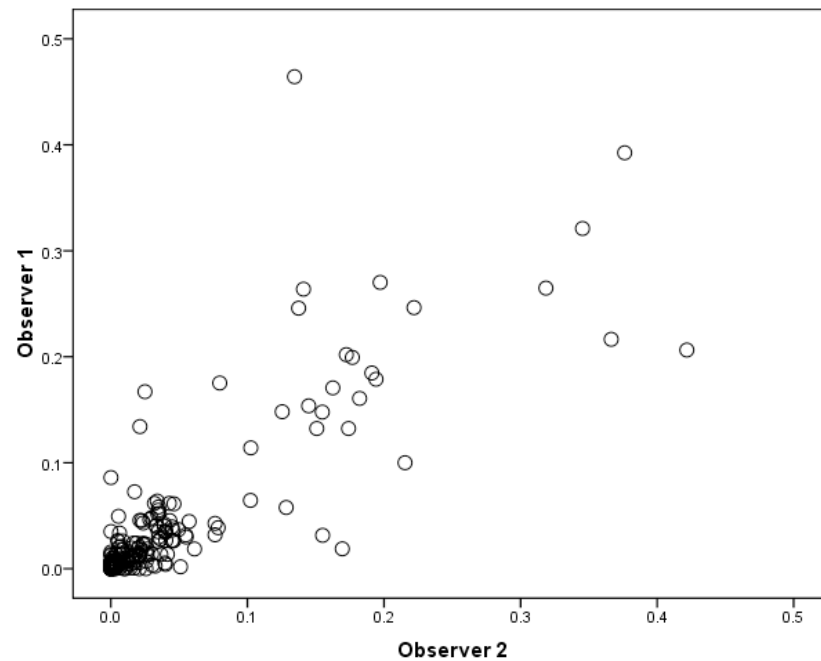
This was performed by two observers independently and again segmented ROIs were correlated using ICCs resulting in  $r=0.96$  (95% CI: 0.92–0.98). This correlation coefficient is extremely high but not quite perfect (Figure 15). The source of the inter-rater variability is the slight heterogeneity caused by imperfections in the way the fat set which, although small, means that there are small variations in the pixel value of the seed point. This is likely to be the cause of the less than perfect ICC.



**Figure 15. Scatterplot illustrating less than perfect but strong linear correlation ( $r=0.96$ , 95% CI: 0.92-0.98) between the two observers measuring metal artefact.**

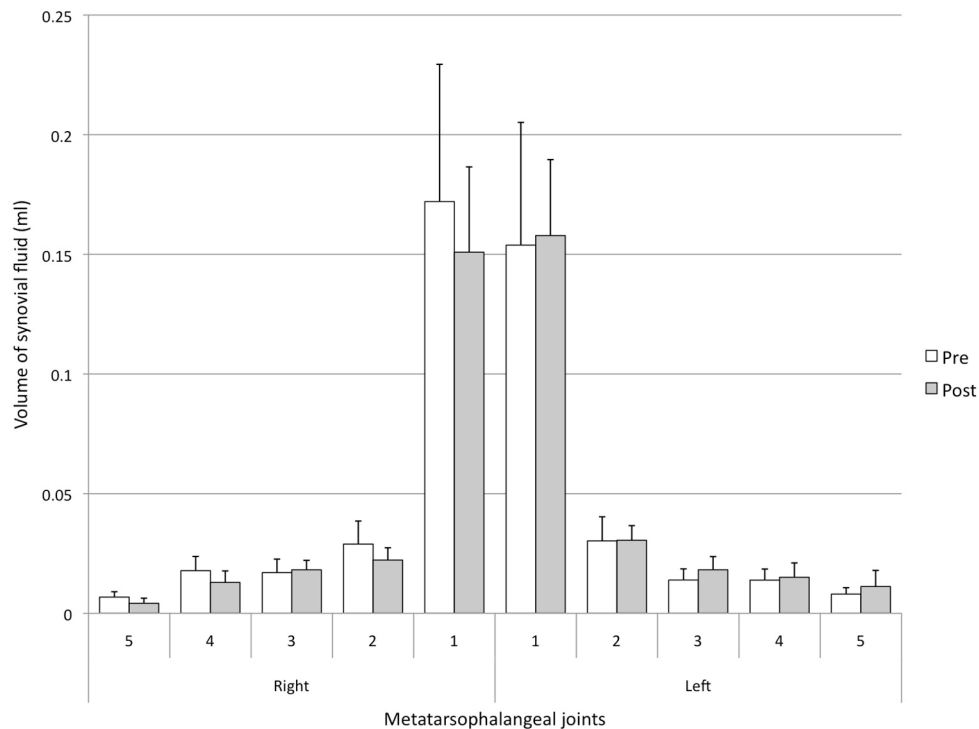
## **2.42 Metatarsophalangeal joint study (72).**

In this study segmentation was used to measure the volume of synovial fluid in the metatarsophalangeal joints (MTPJs) in athletes at rest and after running. The aim of the study was determine whether or not load bearing through the joint with stretching of the synovium was associated with increased volumes of synovial fluid. High resolution T2 weighted sequences were obtained to produce MR images that were fluid sensitive (fluid is represented as high signal intensity) in order to facilitate segmentation. Two observers independently selected seed point and grew one or more ROI and propagated them to segment the volume of fluid in each joint. When observations were compared with ICC the reliability was significantly lower than in the phantom study (Figure 16).



**Figure 16. Scatterplot demonstrating correlation of volume measurements (ml) in metatarsophalangeal joints of asymptomatic volunteers ( $r=0.70$ , 95%CI: 0.60-0.78)**

In particular there was a direct relationship between the magnitude of the differences in observations and the volume of joint fluid. This will have happened for the following reasons. The larger joints, which were typically the first MTPJ (Figure 17), were often complex shapes requiring two or three ROIs to segment completely. The joint fluid was also heterogeneous making seed point selection and thresholding more variable. There was also a tendency for mean signal from fluid to drift from slice to slice; this was a result of magnetic field inhomogeneities caused by the volunteer's body in the MR machine; a problem not encountered with phantom studies.



**Figure 17. Histogram of mean volumes of fluid in each metatarsophalangeal joint of nine volunteers (n=9 of each joint) with standard errors of the mean for each joint.**

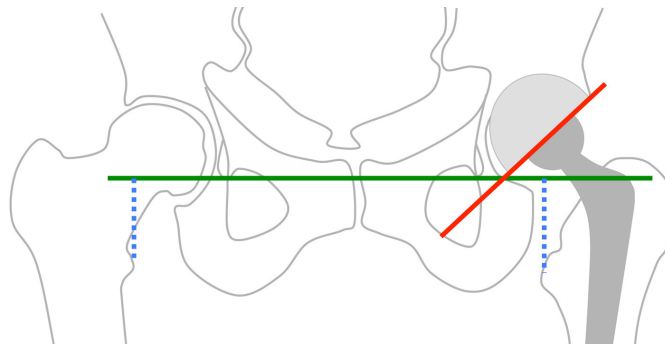
### **2.5 PACS studies (6,19,73)**

As described in chapter 1 the development of PACS has come with direct and indirect research opportunities. The standardisation of a digital image file format (DICOM) that would allow the distribution of image data across networks has also provided a standardised platform for developing image analysis tools. Although some software applications for analysing DICOM data are complex (see Chapter 3) other simple electronic tools can now typically be found on every networked workstation. These electronic tools have now replaced the wax pencils, rulers, protractors and goniometers that were used to measure features on conventional hardcopy radiographs. Although linear and angular measurements from conventional radiographs have been commonplace in clinical and research practice for decades there have been limited attempts to assess the reliability of these techniques and

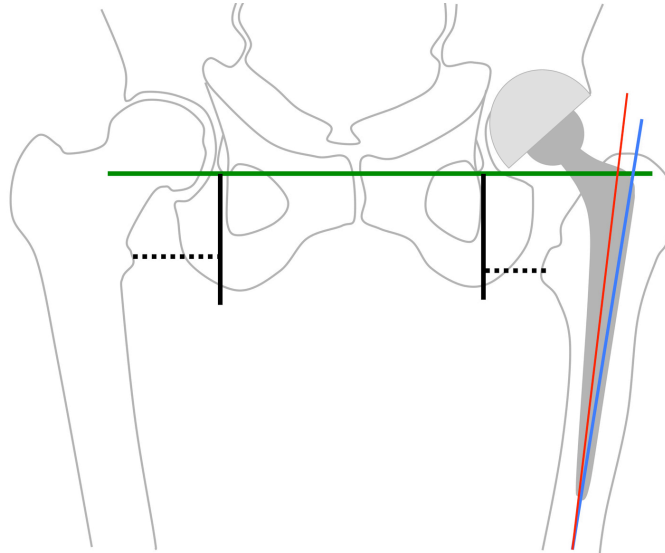
those that have been attempted arguably did not use optimal statistical methods (75). Certainly the reliability of measurements obtained from electronic callipers and goniometers on DICOM workstations had not until now been established.

### 2.51 Conventional computerised radiography (73)

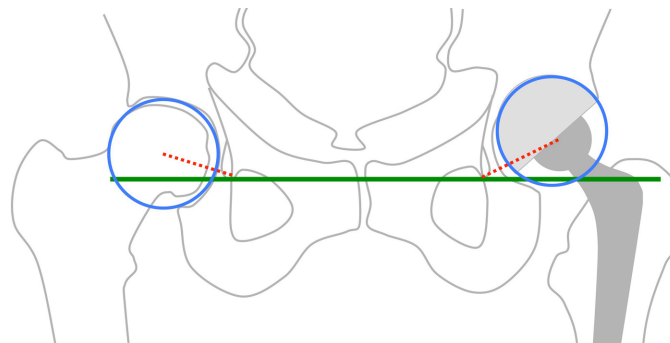
To measure the reliability of a set of standard radiographic measurements that define the anatomical position of a THR on a post-operative radiograph 50 post-operative pelvic radiographs were selected sequentially from PACS (a correlation sample size calculation assuming that  $r=0.7$  yielded a sample size of  $n=34$ ). Two independent observers obtained the following measurements: acetabular inclination, leg length, lateral offset, centre of rotation and stem angle working from a pre-defined method (Figures 18 to 20).



**Figure 18.** The inter-tear drop line (green) is taken as a base line against which acetabular inclination (angle subtended by red line) and leg length (difference in lengths of dotted blue lines) are measured



**Figure 19.** Lateral offset is measured as the difference in the length of a line drawn from a perpendicular dropped from the tear drop to the lesser trochanters (dotted black line). Stem angle is the angle between a line drawn along the centre of the femoral stem and a line drawn midway between the two femoral cortices of the proximal femoral diaphysis.

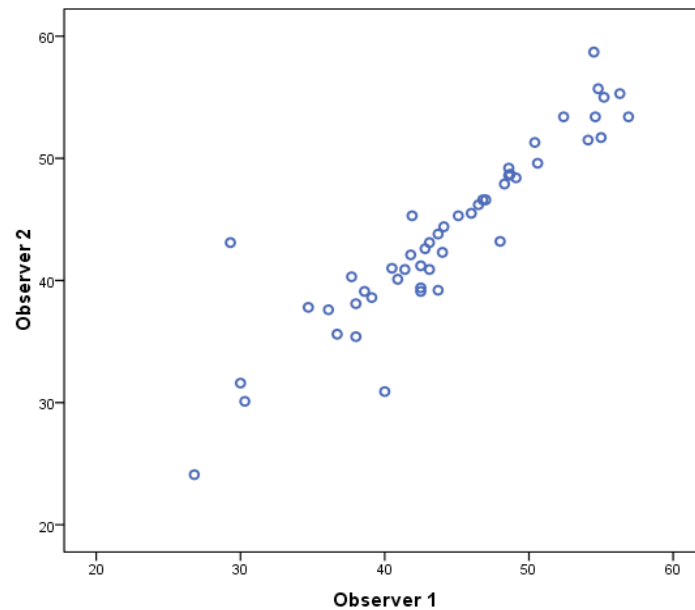


**Figure 20.** Centre of rotation for the THR was defined by a line drawn from the tear drop to the centre of the “bearing”. This line (red dotted) was defined in terms of its length and angle subtended to the baseline.

ICCs were calculated for inter and intra-rater reliability for each measurement.

The results are summarised in table 7 where for most measures the ICC was very high ranging from 0.73 to 0.95 (Figure 21)





**Figure 21. Scatterplot comparing measurements of acetabular inclination from two observers demonstrating closing linear correlation ( $r=0.95$ ).**

The lowest ICC was for the lateral stem angle which measured 0.68 and this was probably because it is difficult to determine the centre of the femoral canal on the lateral view of the proximal femur because of the normal posterior curvature (recurvatum). The standard deviation for the differences in stem angle measurements was in the order of the standard deviation for the normal range and therefore a reduction in reliability is predictable.

Measurement	Mean (SD)	Mean difference in observations (SD)	Intraclass correlation coefficient (95% CI)
Acetabular inclination (°)	44 ± 7.1	1.8 ± 2.4	0.95 (0.92-0.97)
Leg length (mm)	3.9 ± 11.1	4.2 ± 4.8	0.91 (0.85-0.95)
Lateral offset (mm)	-3.0 ± 13.2	3.6 ± 4.1	0.95 (0.92-0.97)
Centre of rotation (distance: mm)	44 ± 5.0	2.7 ± 2.5	0.86 (0.75-0.92)
Centre of rotation (angle °)	30.9 ± 5.3	4.6 ± 3.5	0.73 (0.53-0.85)
Stem angle AP (°)	-1.4 ± 2.0	1.1 ± 1.0	0.85 (0.74-0.91)
Stem angle Lateral (°)	-1.5 ± 3.6	2.5 ± 3.2	0.68 (0.43-0.81)

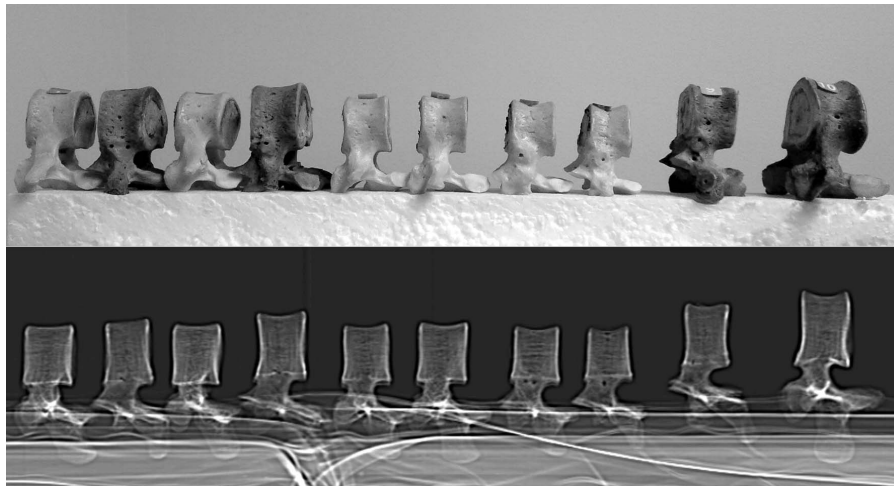
**Table 7. A summary of the descriptive statistics of THR positional measurements, differences between observers and the reliability as measured by ICC.**

The results contra-indicate previous studies that have suggested that conventional radiographic measurements of THR position are unreliable. This may be for several reasons. The first is that it is easier to use electronic tools to replicate measurements on radiographs. Certainly one would imagine that the process of using a digital goniometer would be less prone to error than the three step process of drawing lines with a chinagraph pencil and ruler and then measuring angles with a hand held protractor. The second reason is that the THR itself may influence the reliability of the measurements. The contour of the radiographic image of the THR varies considerably from manufacturer to manufacturer and it may be that be easier to reliably reproduce the selection of a line of measurement with some than others. For example selecting the outer margins of an acetabular prosthesis may be easier with a metal backed cup than with a pure polyethylene cup. Therefore the results of this study may not be applicable to all prostheses. The third main reason may be methodological. Previous study designs for

reliability testing have included using large numbers of observers on a relative small sample of cases using kappa statistics to measure levels of agreement (75). The multiple observer approach may yield a reliability result that is closer to general clinical practice whereas the result of our work could be argued to be the best that can be produced in the controlled context of a research study. Kappa statistics on continuous numerical data are likely to underestimate reliability because it is a measure of agreement and not correlation.

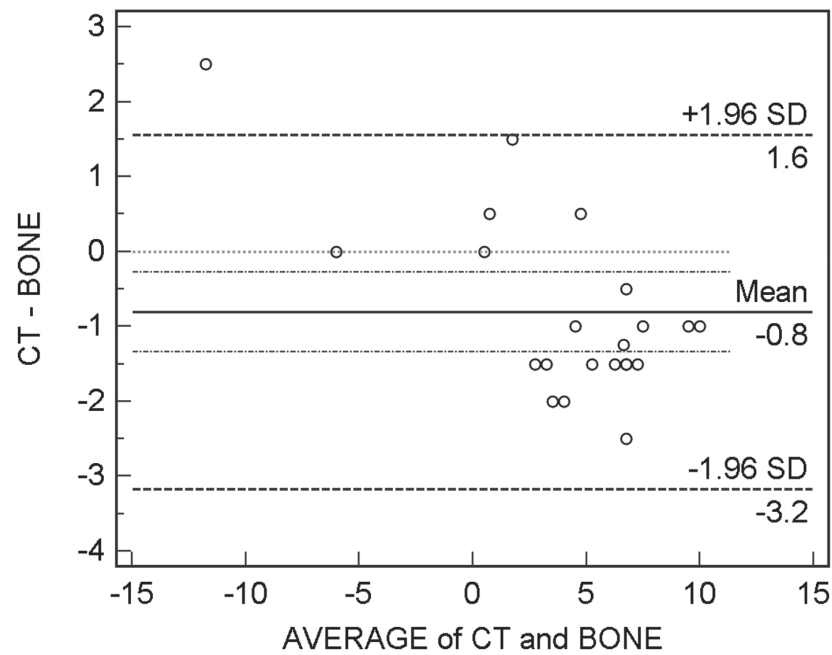
### **2.52 Computed tomographic radiographs (6)**

As described in Chapter 1 CT radiographs obtained during routine planning of CT examinations (Figure 22) can be a useful source of incidental anatomical data particularly for defining normal population profiles. However there are two questions that arise from using this data. The first is the reliability because the images are created from low dose protocols which are relatively low resolution and have high signal to noise levels which might adversely influence the observers' ability to identify key anatomical landmarks. The second question is of validity. CT manufacturers do not disclose the mathematical algorithms that are used to generate these planar images and the geometric validity of a CT radiograph has not been previously assessed.



**Figure 22. A selection of cadaveric dry bones (upper row) and the corresponding lateral CT radiograph.**

This geometric validity of the lateral CT radiograph was validated in the spine by obtaining a lateral CT radiograph of 22 dry cadaveric vertebrae and comparing endplate angles measured on CT with those measured directly from the dry bones with a goniometer (in this case considered the gold standard). A Pearson correlation coefficient measured ( $r = 0.961$ , 95% CI: 0.96-0.99) and a Bland-Altman plot (Figure 23) demonstrated a small  $0.8^\circ$  mean difference in the measurements (95% CI in difference of the means:  $0.27^\circ$ - $1.34^\circ$ , 95% limits of agreement:  $-3.2^\circ$  to  $1.6^\circ$ )



**Figure 23. Bland-Altman plot comparing vertebral endplate angulation obtained from CT and direct from cadaveric dry bones. The mean difference between the two datasets was  $-0.8^\circ$  (cadaveric bone measurements were  $0.8^\circ$  less than CT measurements), the dot-dash line above and below the mean difference represents the 95% CI ( $0.3^\circ$  to  $1.3^\circ$ ) for the difference in the means, while the dashed lines represent the limits of agreement ( $-3.2^\circ$  to  $1.6^\circ$ ).**

The reliability of three angular measurements of “normal” anatomy obtained from CT radiographs described in Chapter 1 (vertebral endplate angulation, acetabular inclination and centre-edge-angle of the acetabulum) was measured using ICCs (Table 8). The inter-observer reliability for vertebral endplate angles could be classified as “almost perfect agreement” by the Landis and Koch criteria (48). The ICC for inter-observer acetabular inclination measurements could be classified as “substantial” and that of centre-edge-angle measurements again as “almost perfect”.

Measurement	ICC	95% confidence intervals
T12 endplate angle	0.86	0.79-0.92
L1 endplate angle	0.81	0.72-0.89
Acetabular inclination	0.67	0.57-0.75
Centre-edge-angle	0.83	0.77-0.87

**Table 8. Summary of the reliability of anatomical angular measurements derived from CT radiographs.**

## **2.6 Summary**

Over the past decade there has been a substantial increase in interest in the study of reliability in radiological research. CT and MR are reliable tools for performing repeated measurements of skeletal anatomy such as the femoral sulcus but in the context of clinical studies of pathological morphology US is not. The soft tissue complications that can occur with metal-on-metal hip replacements can be reliably graded using an MR scoring system but there are limitations to the reliability when disease volume is low. Segmentation techniques are becoming increasingly common as a method of obtaining volume measurement but the reliability of the technique is dependent on the heterogeneity and complexity of the target tissue. While in principle the technique can be almost perfectly reproducible in phantom studies, in clinical practice it is reduced to “moderate” or “substantial” agreement. Simple electronic calipers found on all modern DICOM workstations can be used to measure anatomical angles in the spine and hip with excellent reliability.

## 3.0 Chapter 3: Angiogenesis imaging

### 4.1 Background

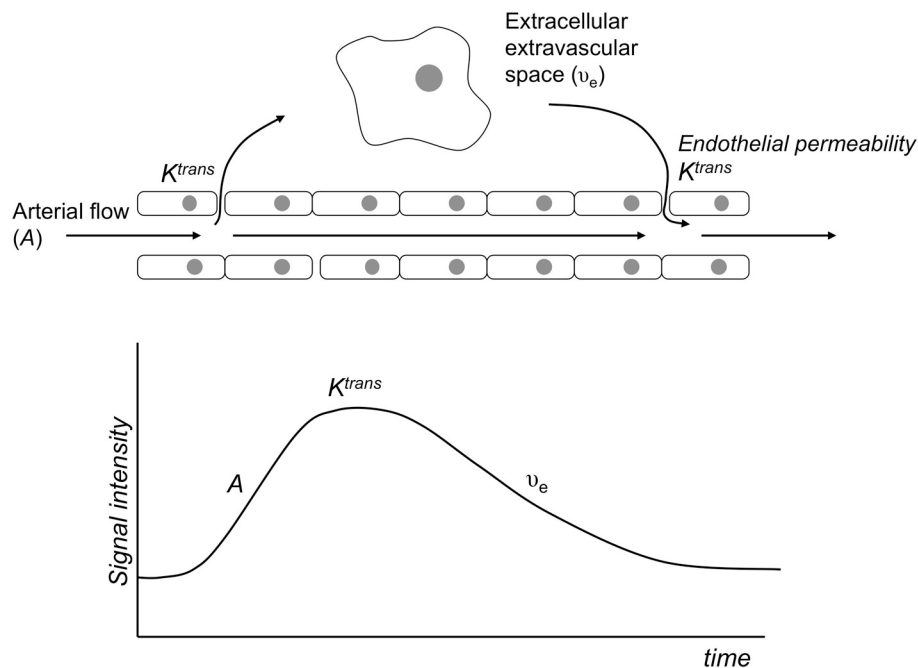
#### 3.11 Angiogenesis

Angiogenesis is the term most frequently used to describe the proliferation of new blood vessels and in particular refers to the formation of thin walled endothelial lined vessels. New blood vessels proliferate by sprouting or splitting off from existing vessels (intussusceptive) as a normal physiological process during growth and then later as a normal response to trauma such as in wound healing. Angiogenesis is also a feature of malignant tumours. Without angiogenesis tumours are unable to grow above a size of approximately  $2\text{mm}^3$  and therefore the tumour cells secrete a number of growth factors of which the most important is probably Vascular Endothelial Growth Factor which results in an often chaotic proliferation of immature capillaries that meet the oxygen requirements of the tumour. Angiogenesis has become the focus of intense clinical research as a mechanism for arresting tumour growth. For example drugs with anti-angiogenic properties such as Thalidomide (which wreaked its havoc in the late 1950s by interrupting vascular proliferation in the developing limbs of fetuses in the first trimester) are now being used to treat multiple myeloma and inflammatory conditions such as Behçet's disease (76).

#### 3.12 Pharmacokinetic modelling

Physiologically-based pharmacokinetic modelling (PBPK) is a technique that allows the absorption, distribution and excretion of compounds in the body to be described in mathematical terms. PBPK was combined with MR imaging by modelling the distribution of gadolinium chelates (injected intravascular

superparamagnetic contrast medium) in a technique known as dynamic contrast enhanced magnetic resonance imaging (DCE-MRI). DCE-MRI has become a standard research and clinical tool for describing angiogenesis in tumours. The most frequently used PBPK model is a two compartment model comprising the vascular and extra vascular extracellular spaces (EES) (77,78). In this model, as the injected intravenous contrast medium reaches the tumour on its first pass, there is an increase in signal intensity as the concentration of gadolinium in the vessels rises. This rise in signal intensity is typically linear and the slope of this ( $A$ ) is taken as a measure of vascularity (Figure 24).



**Figure 24. Diagrammatic representation of the two compartment pharmacokinetic model commonly used in DCE-MRI. Arterial flow ( $A$ ) is calculated as the slope of increase in signal intensity in the wash-in phase of the perfusion curve. The endothelial permeability coefficient ( $K^{trans}$ ) is derived from the peak of the curve, where the movement of gadolinium from the vascular to the extracellular extravascular space is in equilibrium. The extravascular extracellular space volume ( $v_e$ ) is calculated from the wash-out slope of the curve.**



Gadolinium will then leak through the endothelial lining of the tumour vessels in to the EES in the tumour. Very aggressive tumours will often have immature new vessels with highly permeable endothelial linings and therefore this transfer of gadolinium chelate from the vascular to the EES compartment will be elevated. With time the concentration of gadolinium in the vascular space will fall as it is excreted by the kidneys and molecules in the EES will diffuse back in to the vascular space. Until there is an equilibrium phase where as much gadolinium is crossing from one compartment to another as the reverse. From this equilibrium phase can be calculated a volume transfer constant between blood plasma and the EES ( $K^{trans}$ ) which is a reflection of endothelial permeability.  $K^{trans}$  (unit is  $\text{min}^{-1}$ ) is resolved from the following equation

$$\frac{dC_t}{dt} = K^{trans}(C_p - C_t / v_e)$$

where  $C_t$  is the concentration of tracer (gadolinium chelate) in the tissue,  $C_p$  is the concentration in blood plasma and  $v_e$  is the total EES volume (78).

As the concentration of gadolinium continues to fall so the net migration of gadolinium will be from the EES to the vascular space and the slope of this curve can be used to derive a measure of the EES volume ( $v_e$ ) with what is described as the *Clearance Model*. Clearance, in this instance, is defined as the volume of gadolinium that is completely cleared per unit time ( $CL_d$ ) such that

$$\frac{dC_t}{dt} = \frac{CL_d}{V_t}(C_p - C_t / v_e)$$

where  $V_t$  is the total tissue volume (78).

One of the peculiarities of MR imaging is that the absolute values in signal intensity are arbitrary. They bear no fixed relationship to tissue types or

concentration of contrast medium although the relative signal intensity of two or more tissue types, such as muscle and fat, are constant. Therefore the endpoint measures derived from PBPK of DCE-MRI are semi-quantitative in that they describe the properties of shape of the wash-in wash-out curve rather than absolute measures of perfusion.

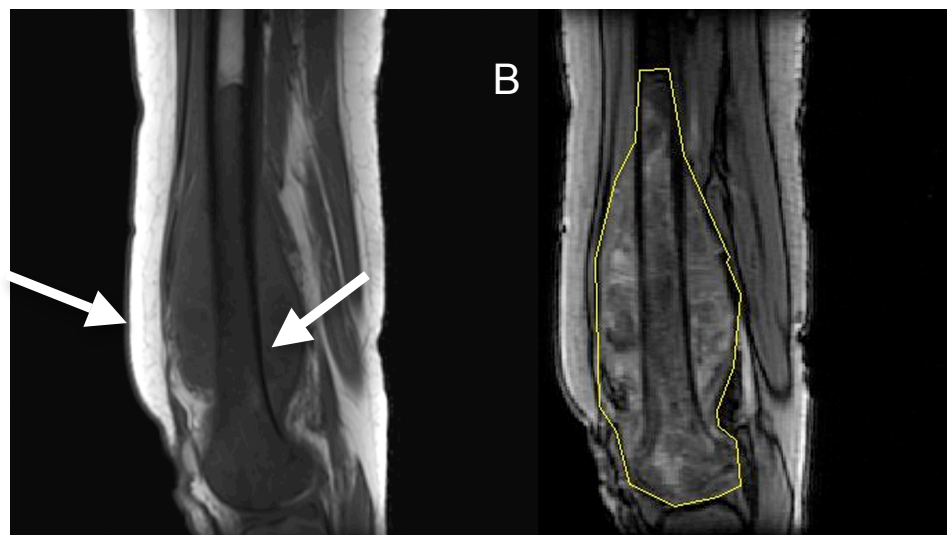
### **3.13 Bone tumour imaging**

MR imaging is a standard investigation for patients with bone sarcomas (malignant tumour arising from bone). The main purpose of the MR is local staging of the tumour; in other words defining the anatomical compartments involved by the disease and the relationship of the mass to critical neurovascular structures. This provides an effective map for determining surgical management but rarely adds to the diagnosis which is mostly determined by plain radiographic features. Neither is conventional MR particularly useful for differentiating between benign and malignant disease (79,80).

Pharmacokinetic modelling of bone tumours is similarly limited demonstrating little useful correlation with histological measures of malignancy. The main focus of research using DCE-MRI in bone tumours has been in estimating response of tumours to chemotherapy. The response to chemotherapy is often variable and at best limited. At the same time the side effects can result in severe morbidity and a recognised mortality rate. Therefore the ability to identify a poor response early on in the course of treatment may save the patient unnecessary side effects and associated healthcare costs. Conversely a good response may persuade a patient and their clinician to persist with treatment despite side effects.

Early work in this field, carried out in the 1990s, tried to characterise tumour response by examining a variety of pharmacokinetic markers at the end of treatment and prior to surgery (81-91). Strong correlations between DCE-MRI measures and histological grades were demonstrated; in particular high

grades of necrosis correlated with low measures of angiogenesis. Correlation of changes in serial measurements during treatment also appeared to be encouraging. However there was a limitation to nearly all of these studies which is determined by the nature of bone sarcomas. During the 1990s and the early part of the 21<sup>st</sup> century MR technology limited operators to acquiring a single slice dynamic MR dataset. This single slice was typically selected to cover the greatest cross-sectional area of the tumour to which the histological grades would be mapped (Figure 25). The nature of bone sarcomas is that they are typically heterogeneous. Not only are there multiple areas of necrosis but there is typically variability in the histological grade from one part of the tumour to another. In a tumour which is mostly a low grade malignancy with one small focus of aggressive high grade disease it is the high grade disease that will determine the patient's prognosis. The obvious limitation of the single slice DCE-MRI technique is that while there may be good side by side correlation with histology it may be that the most important prognostic foci within the tumour are not included in the study.



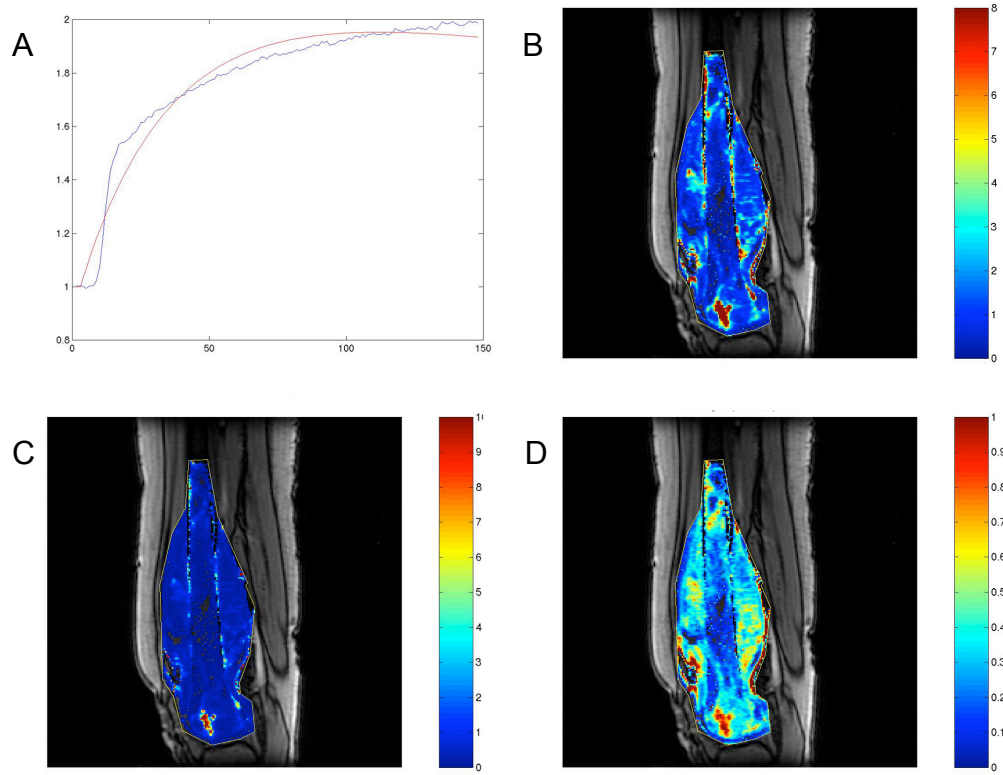
**Figure 25. (A) Sagittal T1W staging MR through the long axis of an osteosarcoma (arrows) arising from the distal femur. (B) hand drawn ROI outlining the margins of the tumour from which the perfusion data was obtained.**

## 3.2 Limitations of Single Slice DCE-MRI (86)

### 3.21 Methods

Twelve patients with appendicular spindle cell sarcomas of bone were entered in to a prospective study with the aim of determining whether or not single slice DCE-MRI could be used to distinguish patients who were found to be in a good or poor prognostic histological category following resection of their tumour (tumour necrosis of greater than 90% is considered to be a good prognostic indicator with 80% 5-year survival rates (93,94)). Each patient underwent a full staging MR followed by a DCE-MRI study prior to starting chemotherapy. Following completion of six cycles of chemotherapy the DCE-MRI study was repeated the day before resection of the tumour. An ROI was drawn around each of the dynamic MR acquisitions, using the corresponding T1W MR as a guide to the tumour margins. From this ROI A,  $K^{trans}$  and  $v_e$  were calculated for each pixel and then represented as a parametric map of the distribution of each of these values within the tumour (Figure 26). The mean, median and inter-quartile ranges of each of these endpoint measures were calculated for each ROI (most of the data failed the Jarque-Bera test for normality and was therefore considered to be non-parametric) (Table 9). The tumour was fixed and then approximately 50% of the tumour was sampled, so as to represent the total distribution of tissue throughout the tumour, and a measure of total tumour necrosis was calculated. Patients were then assigned to either a good or poor prognostic group. Pharmacokinetic endpoints prior to chemotherapy and after chemotherapy, and the change in these endpoints during chemotherapy, were compared with the allocation of patients in to one of the two prognostic groups. There was no statistically significant difference in any of the pharmacokinetic endpoints when the two prognostic groups were compared.

## 3.22 Results



**Figure 26. (a) Graph of the mean arterial slope (A) and parametric maps of the distribution of A (b), the endothelial permeability coefficient  $K^{\text{trans}}$  (c) and the EES  $v_e$  (d) in patient 7 after chemotherapy**

Variable	Necrosis <90% Median (Q1 to Q3 <sup>1</sup> )	Necrosis ≥ 90% Median (Q1 to Q3 <sup>1</sup> )	Median Difference (95% C.I.) <sup>2</sup>	p-value <sup>3</sup>
Change in Mean Arterial Slope (A)	-0.94 (-2.24 to 1.55)	-1.48 (-2.61 to - 0.58)	0.83 (-1.24 to 3.22)	0.47
Change in Median Arterial Slope (A)	-0.47 (-2.19 to 1.09)	-1.31 (-2.16 to - 0.05)	0.94 (-1.12 to 2.74)	0.58
Change in Mean K <sup>trans</sup>	-0.33 (-1.15 to 0.68)	-0.93 (0.99 to 2.36)	0.82 (-0.50 to 2.27)	0.23
Change in Mean $v_e$	-0.05 (-0.12 to 0.12)	0.07 (0.01 to 0.27)	-0.11 (-0.45 to 0.06)	0.38

1: Lower and Upper Quartiles

2: 95% Confidence Interval for difference in Medians

3: P-value from Mann-Whitney U Test

**Table 9. Table summarising the differences in pharmacokinetic endpoint measures, from before and after treatment, for the two prognostic groups.**

The result of this study did not support any of the published reports at that time. There are a number of possible reasons for this which can be broadly classified as limitations inherent in the design of the study and those resulting from the limitations of DCE-MRI in this clinical setting.

The sample size is clearly small ( $n=12$ ). The sample size had been expected to be 20 but recruitment did not reach initial estimates (A total of 8 recruited patients were excluded because of failure to complete or start chemotherapy, inadequate MR studies and death during chemotherapy). Although some authors had previously obtained statistically significant results using DCE-MRI with similar or smaller sample sizes, for example  $n=6$  (84),  $n=8$  (83),  $n=10$  (81,95) and  $n=12$  (96), the confidence intervals for this particular study were wide and therefore it was clearly underpowered. A sample size calculation using the data from this study suggests that a sample size of 40 would be required to demonstrate a statistically significant difference between the two groups. On the other hand it could be argued that any clinically useful discriminator should be evident in a small sample size. A statistically significant difference in a sample size of 40 does not necessarily equate to a cost-benefit advantage.

The second limitation to the study design relates to the inclusion criteria for tumour type. This was defined as spindle cell sarcoma of bone, a histological class of tumour based on characteristic cell shape, of which there are a number of subtypes, but all of which are treated with the same chemotherapy regime. The evidence now appears to suggest that the best results of using DCE-MRI to monitor response to chemotherapy are obtained with single tumour types such as osteosarcoma or Ewing's sarcoma.

There are potentials for selection bias with 8 out of 20 patients failing to make it through the study although it is difficult to see that this has had a real effect on the results. The test-retest reliability of DCE-MRI was not addressed in this study (nor does it appear to have been addressed in any published work) and there is plenty of scope for variability in patient positioning and selection of MR parameters. Although certain MR parameters such as relaxation and echo times, and matrix size were fixed, there will have been variability, in the position of a limb, the slice selection orientation and drawing of the ROI, from study to study. There is also evidence from animal models that tumour perfusion, measured with DCE-MRI, may vary considerably from hour to hour. The selection of the ROI

from which the pharmacokinetic measures were derived was only performed once by one observer and therefore there was no inter-rater reliability measurement. Again this is a common limitation of published DCE-MRI studies.

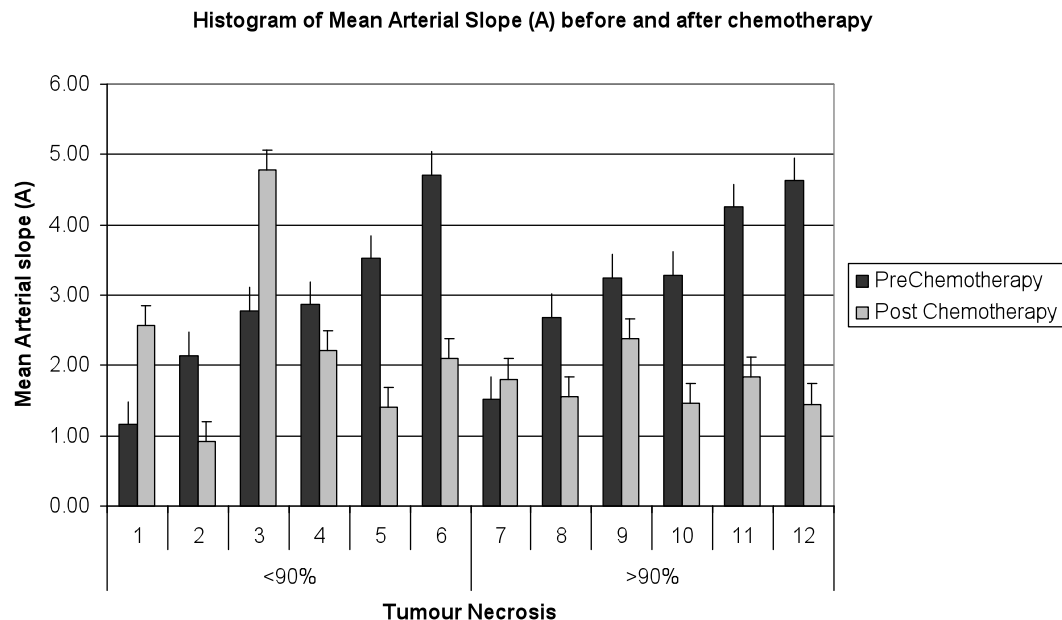
The most interesting limitation of the study arises from the timing of the DCE-MRI studies and the normal response of tumours to chemotherapy. Although tumours do induce neo-angiogenesis, through a range of mediators, neo-angiogenesis is also a normal part of any tissue repair response, mediated by the same growth factors used by tumours. When extensive necrosis is induced within a soft tissue tumour, such as a small cell lung cancer, it will often shrink and sometimes disappear as the necrotic tissue is broken down by macrophages and exported. Extensive necrosis in a bone tumour will leave a defect within bone which the body will try and repair. The repair process will be the same for any traumatic insult to bone and will involve neo-angiogenesis with new vessels providing oxygen for migrating osteoblasts that will attempt to repair the bony defect. Therefore if the second DCE-MRI is performed after successful treatment of a bone tumour the angiogenesis may be related to repair rather than tumour and therefore the PBPK endpoint measures will have no relationship to response to treatment.

This is illustrated by one of the patients in the study with near complete tumour necrosis (>98%) following chemotherapy, who was therefore clearly in the good prognostic group, who had an increase in all of her PBPK measures (Figures 27 and 28). This can only be explained by the fact that repair mediated angiogenesis was contributing to the post-chemotherapy measurements.

The timing of the second DCE-MRI was chosen because it was felt that this should take place as close to the histological assessment as possible to minimise discrepancies due to time. This was standard in most of the preceding published studies. However the aim of this study was not to correlate imaging findings with histology directly; it was to determine whether



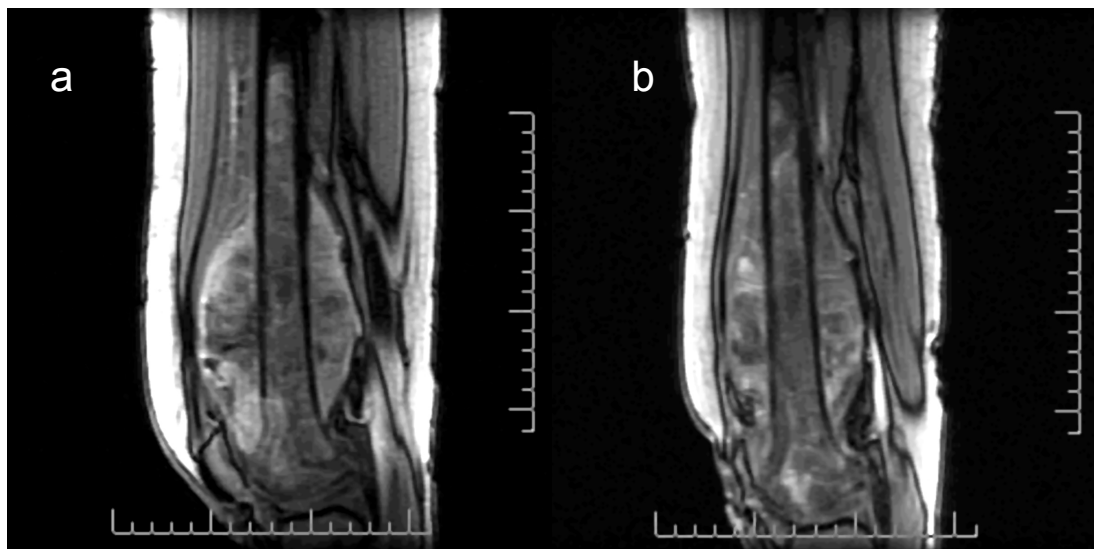
or not changes in measures of angiogenesis could be used to determine prognostic outcome from total tumour necrosis rates.



**Figure 27. Histogram of mean arterial slope (A) before and after chemotherapy for each patient divided in to the two prognostic groups. Patient 7 demonstrated an increase in all PBPK measures after >98% tumour necrosis.**

In retrospect it would probably have been better to perform the second DCE-MRI much earlier in the course of treatment; possibly after the first cycle of chemotherapy. This would have minimised any effect from repair mediated angiogenesis. However there are two other factors that need to be considered when optimising the timing of DCE-MRI studies. The first is that changes in renal function, which are not uncommon during long courses of chemotherapy, will alter the pharmacokinetics of substances, like gadolinium chelates, that are excreted by the kidneys. Therefore some changes in PBPK parameters may be due to alterations in renal function if long intervals between studies are used. The second is that it is now considered that in order to avoid nephrogenic systemic fibrosis (a fatal complication of intravenous gadolinium use in patients with renal failure) glomerular filtration

rates should be below prescribed safe levels and then doses of gadolinium should be given at least two weeks apart (97). To repeat the study to answer the original research question adequately would require a much large sample size (probably  $n=40$ ), a more homogeneous group of tumour types and an earlier second DCE-MRI study. However the original research question has to some extent been superseded by developments in MR technology whereby it is now feasible to perform multislice DCE-MRI that covers the whole tumour with a high enough temporal frequency that delivers enough datapoints for PBPK.



**Figure 28. Sagittal gradient echo T1W images taken from peak enhancement during the DCE-MRI acquisition (a) before and (b) after chemotherapy demonstrating florid perfusion despite >98% necrosis.**

### **3.3 Summary**

Single slice DCE-MRI is probably not a useful tool for predicting prognostic histological outcomes in a general population of spindle cell sarcomas.

While it might be able to differentiate statistically between the two groups in higher powered studies this would not necessarily be a useful difference in clinical practice. It is likely that repair mediated angiogenesis is a confounding factor when studying responses to treatment and therefore future study designs should consider performing DCE-MRI earlier on in the course of chemotherapy in order to isolate tumour angiogenesis responses to treatment.

## 4.0 Chapter 4: Imaging of joint replacements

### 4.1 *Historical perspective*

#### 4.11 Evolution of joint replacements

The first successful total hip replacement was performed by Philip Wiles,<sup>4</sup> at the Middlesex Hospital, in 1938. After an interruption, by the Second World War, the widespread use of total hip arthroplasty was popularised by Ken McKee<sup>5</sup> (who had trained under Wiles), in Norwich, and John Charnley<sup>6</sup> in Wrightington (98). McKee developed a number of uncemented prototypes which were to evolve, by 1960, in to the McKee-Farrar cemented total hip replacement (THR). This was the first hip prosthesis to become widely, and successfully, implanted (99,100) and comprised a cobalt-chrome metal-on-metal (MOM) articulation in which both the acetabular and femoral components were fixed to bone with polymethylmethacrylate cement. Meanwhile Charnley was pursuing a different approach (101,102). Convinced that the McKee-Farrar prosthesis would lead to unacceptable loading and shearing strains that would eventually cause loosening he designed a hip replacement comprising a small diameter metal femoral articulation and a polymer cup; the polyethylene-on-metal (POM) articulation. While both the McKee-Farrar and Charnley total hip replacements saw considerable success both had principle mechanisms of failure directly attributable to the materials used. Eventually the MOM articulation, which suffered unacceptably high failure rates and was dogged by concerns about toxicity, particularly in the cobalt-chrome prostheses, fell from favour and the POM hip replacement became the pre-eminent design from the early 1980's onwards (98).

---

<sup>4</sup> Philip Wiles FRCS 1899-1967

<sup>5</sup> George Kenneth McKee CBE FRCS 1905-1991

<sup>6</sup> Professor Sir John Charnley Kt CBE DSc FRCS FRS 1911-1982

## 4.12 Mechanisms of joint replacement failure

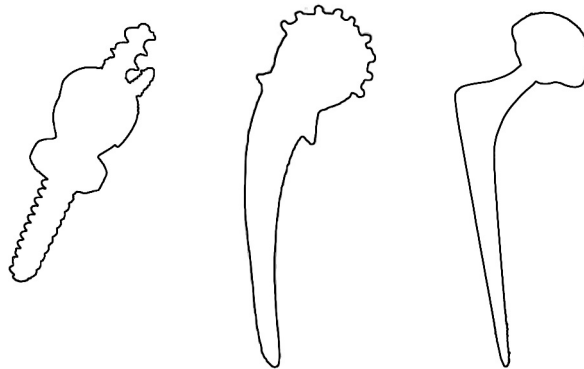
There are recognised complications in all forms of joint arthroplasty. There are generic complications, such as infection, and there are complications that are specific to a particular type of prosthesis (103,104).

The McKee-Farrar THR appear to have polarised outcomes. Either the implant survived, in which case it often outperformed the Charnley POM, or it failed early (99,102,105). Failure has been blamed on relatively poor tolerances in the manufacture of the bearings which caused excessive torsional friction resulting in mechanical loosening (106,107). The metal-on-metal bearing caused metal ions to be shed in to the effective joint space of the prosthesis. It was subsequently demonstrated that this is accompanied by elevated levels of serum cobalt and chromium ion levels (108,109) and because of this there have been concerns about carcinogenesis (110). When the congruity of articular surfaces breakdown then large amounts of macroscopic debris can be shed forming a black sludge which lines the pseudocapsule and can present with large peri-prosthetic cystic masses in a condition called metallosis (111,112).

In contrast the friction in POM THRs, albeit relatively low, produces polyethylene particulate debris. These particles are inert but are treated as foreign bodies and are ingested by macrophages. The macrophages then coalesce to form small particle granulomas, which grow and destroy adjacent bone in a process termed osteolysis. As the volume of bone loss increases so the fixation of the prosthesis reduces until it fails, and this is the principle cause of aseptic loosening in POM hip replacements (113).

By the turn of the century a second generation of MOM arthroplasties had been introduced in an attempt to find an alternative to the polyethylene osteolysis of POM THR and because it was felt that technical developments in computer numeric control (CNC) machining and pre-clinical testing had overcome the limited machine tolerances of the first generation of MOM

THRs like the McKee-Farrar (114)(Figure 29). These were primarily aimed at young patients with premature osteoarthritis (OA) who were likely to need a number of revision procedures during their lifetime if treated with POM THR, with each procedure associated with escalating risk and morbidity (115). The second generation MOM arthroplasties include the re-surfacing hip replacements which preserve the femoral neck facilitating subsequent revisions (116).



**Figure 29. Line drawings representing the radiographic silhouette of three MOM arthroplasties: the McKee from the 1950's on the left, the McKee-Farrar from the 1960's in the middle and one of the second generation MOM THRs, the DePuy Ultima® TPS, from the 2000's.**

#### **4.13 Conventional imaging of joint replacements**

For most of the twentieth century radiological evaluation of joint replacements has relied almost exclusively on conventional radiography. Radiographs in the immediate post-operative course have been used to assess the quality of the surgical procedure by assessing the position of each of the components, relative to defined bony landmarks (117,118), and the quality of prosthesis-cement-bone fixation (119). There is a recognised range of satisfactory post-operative positions, which are usually defined by variations in human morphology, as well as bio-mechanical factors determined during the design of any specific prosthesis (120-122). For example it is recognised that positioning the acetabular cup at an angle of

more than 45° increases the rate of polyethylene wear (123). Identifying this at the early post-operative stage allows the clinician to modify their surveillance of the patient.

Conventional radiography has also been used to identify and monitor the complications of THR. In particular the appearances of metallosis in MOM THR and aseptic osteolysis in POM THR are well described. Metallosis is characterised by high attenuation of the x-ray beam at the site of metallic debris within the pseudocapsule resulting in “cloud-like” opacification of the radiograph (124). Today this is most commonly seen in THRs where the polyethylene cup of the acetabular component has been breached so that the metal femoral head articulates with the metal backing.

On the other hand aseptic osteolysis, caused by small particle granulomas (125), produces lytic areas within bone that are typically well defined with a thin sclerotic margin which indicates a relatively slow and non aggressive pattern of growth. These lytic areas occur at any point within the potential joint space of the THR, in other words, wherever there is a direct communication of synovial fluid with the bearing then there is a direct route for passage of polyethylene debris that will settle and form granulomas. These usually form around the acetabulum and between the femoral stem and the cement mantle typically seven to eight years after surgery (126). As these areas of lysis enlarge and coalesce they destabilise the fixation of the components of the prosthesis causing it become loose and therefore move. Movement of the prosthesis on serial radiographs is diagnostic of a loose prosthesis (127,128).

There are however two major limitations in the use of conventional radiography for the assessment of failing hip replacements. The first is that it provides a planar two-dimensional representation of a complex three-dimensional anatomical structure. Conventional surveillance radiography in the UK comprises an antero-posterior (AP) and lateral radiograph which means that healthy bone will be superimposed on areas of lysis at a number of points and this explains the limited diagnostic sensitivity for osteolysis

(129,130). Increasing the number of radiographs taken by performing multiple oblique views improves this sensitivity but at a cost of an increased radiation burden (131). The inter-rater reliability of observers detecting osteolysis is also poor when a single radiograph only is used (132). The second limitation arises from the difficulty in reproducing exactly comparable pelvic radiographs with which to compare measurements or osteolysis over time. Variability in the positioning of the patient, the radiographic parameters and the individual characteristics of ageing cathode ray tubes all contribute to variability in serial radiographs before the reliability of the observers can be considered (133).

## **4.2 CT and MRI of joint replacements**

### **4.21 Computed tomography**

Computed tomography (CT) is an imaging technique that involves x-rays produced from a cathode ray tube which is mounted on a circular gantry opposite a set of detectors, both of which rotate around the patient. Mathematical back-projection algorithms are then applied to the received data to reconstruct sequential tomograms (image slices). The technique was invented by Hounsfield<sup>7</sup> in the early 1970s and by the 1980s was in widespread use in developed countries (134). The early CT units were single slice machines. These would acquire all the data for a single axial image before moving the patient stepwise to the next slice position. In the 1990s this technique was superseded by helical CT which acquired the data while the table moved continuously. Tomograms were then reconstructed from a helical dataset. By the early 2000s multidetector CT (MDCT) machines, that could acquire more than one slice at a time, became commonplace so that 64 slice machines are current standards with dedicated cardiac machines acquiring 256, or more, slices in one revolution. This allows isovolumetric datasets (where each volume element, or voxel, is cubic) of less than 1mm along each edge allowing multiplanar

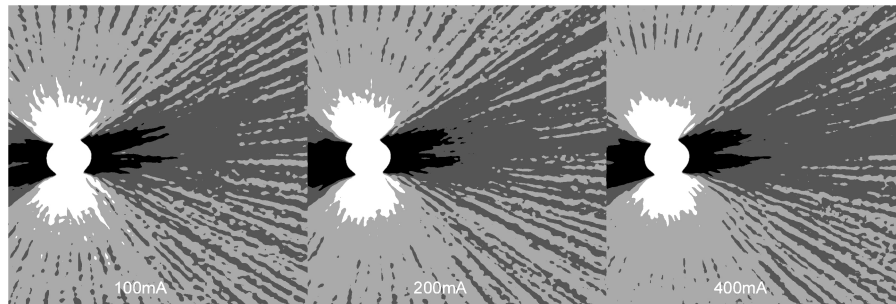
---

<sup>7</sup> Sir Godfrey Newbold Hounsfield CBE FRS, (1919 – 2004)



reconstructions. The increase in speed and computing power, that has allowed ever more complex reconstruction techniques, has now been complemented by the advent of dual energy CT. In principle two different cathodes, of different composition and density, are used to generate the x-ray fan beam which then comprises data from two sources, each with its own profile of x-ray photon energies. This means that, by measuring differential attenuation from the two sources, tissue density can be calculated, in a process similar to bone densitometry using DEXA.

CT of orthopaedic implants produces artefacts by two different mechanisms. The first is described as beam hardening artefact. This occurs because within a beam x-ray photons there is a roughly normal distribution of energies around a mean. X-ray photons of lower energies are preferentially deflected by the nuclei of the metals within the prostheses and the higher the atomic number of the metal the greater the differential deflection (ferrous alloy implants produce worse artefact than titanium). The computer that performs the back projection computation is however expecting a normal profile of energies at the detector in order to make the calculations that assign a Hounsfield unit (HU - a measure of attenuation on CT) to that pixel. The result is that incorrect measures of attenuation are calculated resulting in an alternating pattern of high and low attenuation radiating out from the prosthesis in a “streak” or “starburst” pattern (135) (Figure 30).



**Figure 30. Series of axial images, with a simplified 4 level grey scale, from a femoral stem phantom imaged with CT with tube currents of 100, 200 and 400mAs demonstrating “streak artefact” radiating from the femoral stem with only minimal improvement with increasing mAs (unpublished work).**

The second source of artefact is that of “photon starvation” or photopaenia. The reduction in the number of x-ray photons reaching the detector results in a decrease in the signal to noise ratio in the imaging system which manifests as a granular texture in the final CT image. Techniques to reduce metal artefact on CT have, until recently, been limited. The traditional teaching is that by increasing the potential difference across the cathode ray tube, and thereby increasing the mean energy (kVp) of the fan beam, that the proportion of lower energy photon being deflected is reduced and that by increasing the tube current (mAs) this increases the number of photons and reduces the noise caused by photopaenia (136-139). Newer reconstruction algorithms have come some way to addressing beam hardening artefact by compensating for the preferential attenuation of low energy photons and photon starvation (139,140). Extended CT scales and thick slice multiplanar reformats are both techniques that do not reduce the artefact but do help in the interpretation of images with metal artefact (141). In practice however beam hardening artefact can still be a significant problem when evaluating bone integrity close to a prosthesis.

## 4.22 Magnetic resonance imaging

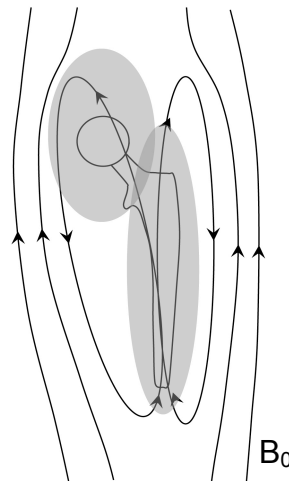
Magnetic resonance imaging (MRI) is a technique that developed from work in the 1960s and early 1970s by Lauterbur<sup>8</sup> (142) who applied nuclear magnetic resonance (NMR) to biological tissues. These were refined for use in humans by Mansfield<sup>9</sup>, who published the first image of in-vivo human tissue (a finger) in 1977 (143), and thereafter MRI became the common descriptor for clinical NMR. The first clinical MR machine was installed in the UK in Manchester in 1983 after which MR provision expanded rapidly. However the presence of metallic orthopaedic prostheses remained an absolute contra-indication to MRI, for most radiologists, until the turn of the century. From 2000 onwards there has been a steadily increasing interest in the role of MRI in the imaging of THR although use as a routine clinical tool identifying gluteal tendon rupture, psoas irritation and defining the extent of small particle disease has been very limited (103,144-150).

The source of metal artefact on MRI is the magnetisation of the prosthesis when it is placed in the magnetic field of the MR machine. The MR machine is designed so that at the iso-centre of the magnet the magnetic field lines are parallel ( $B_0$  magnetic field). When a metal prosthesis becomes magnetised within the  $B_0$  field it then generates its own magnetic field lines which either null the  $B_0$  field lines or displace them. This results in either a loss of signal (signal void), caused by dephasing of protons, or geometric distortion of the image, and are described as susceptibility artefacts (Figure 31).

---

<sup>8</sup> Dr Paul Christian Lauterbur (1929 – 2007) Nobel Laureate 2003

<sup>9</sup> Sir Peter Mansfield FRS (1933-) Nobel Laureate 2003



**Figure 31. Line diagram demonstrating magnetisation of a Charnley femoral prosthesis with magnetic field lines running counter to the  $B_0$  magnetic field causing inhomogeneity by displacement of the  $B_0$  magnetic field lines and signal voids (shaded areas) where opposing field lines meet.**

There are properties of a prosthesis and of a magnet which influence the amount of artefact that is produced which the radiologist has no control over; namely the type of metal (ferrous materials magnetise more strongly than titanium) and the strength of the magnetic field (as this increases so does the artefact). However artefact can be minimised by selecting pulse sequences that are resistant to the effect of susceptibility artefacts such as fast spin echo and inversion recovery sequences and indirectly controlling the amplitude of the frequency encoding gradient ( $G_{FE}$ ) (148,151-155). The frequency encoding gradient is a magnetic field gradient that is applied to the  $B_0$  magnetic field that means within the area of interest the protons precess (spin off-axis) at different frequencies. This difference in frequencies can be used to spatially localize the read out signal and therefore assign a signal to a point in space. The strength of  $G_{FE}$  is determined by the following equation:

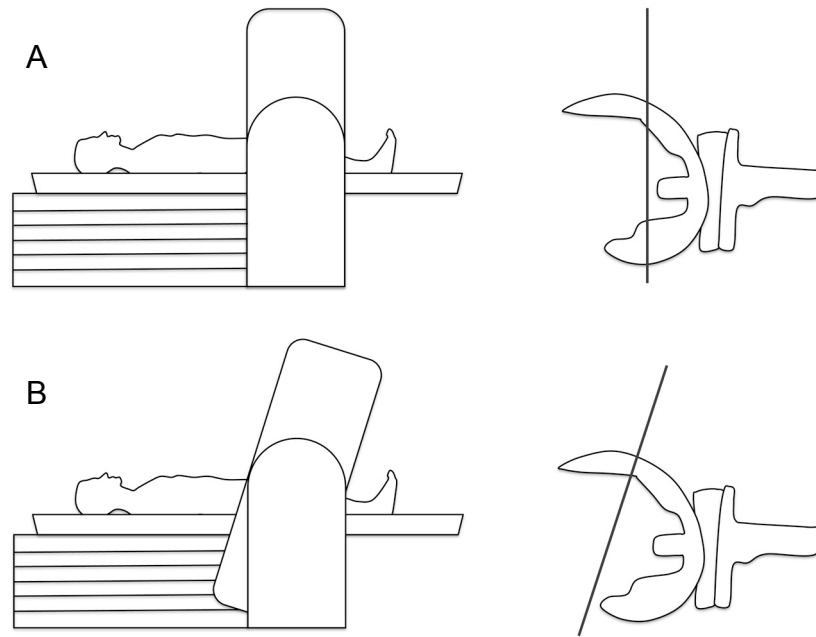
$$G_{FE} = \frac{2 \cdot \pi \cdot BW}{\gamma \cdot \Delta\chi}$$

Where BW is the receiver bandwidth,  $\gamma$  is the gyromagnetic ratio and  $\Delta\chi$  is the voxel size in the frequency encoding direction. Therefore it follows that the amplitude of  $G_{FE}$  can be increased, and metal artefact reduced, by increasing BW or decreasing  $\Delta\chi$ . However both of these adaptations come with a penalty; increasing BW decreases the signal to noise ratio (SNR) resulting in more noise in the final image, and decreasing  $\Delta\chi$  increases acquisition time linearly and longer acquisition times are more susceptible to movement artefact. What has not been understood until now is what that cost is and what the optimal combination of BW and  $\Delta\chi$  is for metal artefact reduction in the clinical setting (156).

### **4.3 Phantom studies**

#### **4.31 Computed tomography (157)**

An anecdotal observation led to the study of the effect of CT gantry angle on metal artefact. The observation was that when the axial data was reconstructed in the coronal and sagittal planes that the metal artefact appeared to be less noticeable than on the standard axial data. It was thought that the interpolation process of multiplanar reconstruction averaged out some of the artefact and that this might be exploited by acquiring non-orthogonal data and reconstructed in the axial plane as well. This technique also had the potential to allow the operator to steer the fan beam and thereby minimise the volume of metal on any single slice acquisition (Figure 32). One of the simplest ways of reducing metal artefact in an image is to reduce the amount of metal in any given slice. This can be done effectively by orientating the long axis of a screw or nail perpendicular to the CT fan beam. However other prostheses, such as the total knee replacement, are complex shapes, have high metal volume and are limited in the possibilities for positioning because of their anatomical location.

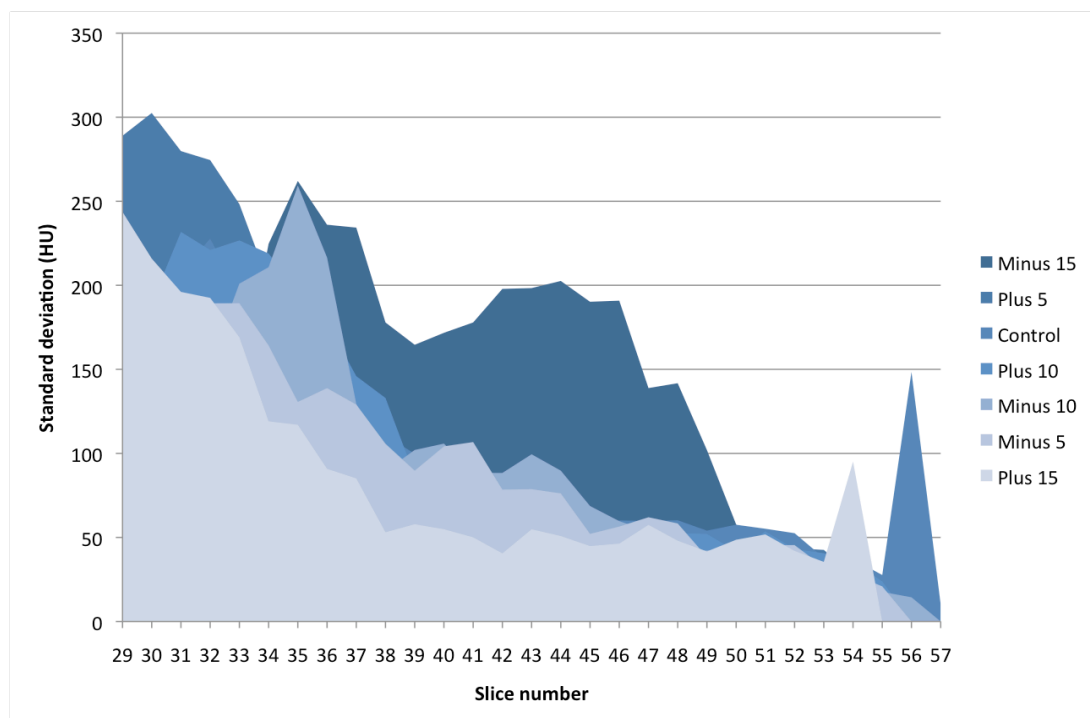


**Figure 32. Line drawing representing a patient undergoing a CT examination with the gantry set at 0° (A) and at -15° (B). The tilted gantry angle includes a smaller volume of metal per slice at the craniad end of the examination in comparison to the images acquired at 0°.**

To test this hypothesis a phantom was constructed using a cobalt-chrome total knee replacement (TKR) suspended in gelatine (containing an iodinated contrast medium that raised the attenuation of the gelatine to an average “soft tissue” value of approximately 50 HU). The phantom was imaged repeatedly on a multislice CT machine with standard orthopaedic imaging parameters with gantry angles from -15° to 15° at 5° intervals, and then each of these acquisitions was reconstructed in a true axial plane. Thereafter two observers independently drew ROIs that were propagated through each reformatted dataset and the standard deviation of the attenuation values in these regions of interest were recorded. Measuring the standard deviation (SD) of pixel values around a prosthesis is a recognised technique for assessing relative metal artefact; the wider the standard deviation, the worse the artefact. Comparison of the different gantry angles was made by

comparing the area under the curve for measures of standard deviation along the length of the prosthesis.

The results of this study demonstrated that it is possible to reduce metal artefact, in comparison with standard protocols, using this technique (Figure 33). However because of the complex shape of the TKR there is no one position that is optimal for the whole of the TKR. Minimal artefact could be achieved for the femoral component, and indeed the whole TKR overall, at a gantry angle of  $5^\circ$  but more impressive artefact reduction could be achieved for the tibial component alone at an angle of  $15^\circ$ .



**Figure 33. Graph demonstrating the area under the curves for standard deviations in background attenuation (HU) imaging the tibial component of a TKR with different gantry angles. The smallest area under the curve, and therefore the least metal artefact, occurs at  $+15^\circ$ .**

What is not clear is what the mechanism of artefact reduction really is. Is it simply the fact that the angled gantry includes less metalwork on certain single slices (there is no overall reduction in metal artefact; it is simply

spread over the longer dataset that is required to cover the prosthesis with oblique axial acquisitions) or is there a mathematical smoothing that comes from the MPR? To work this out would require a different phantom, one in which the geometric influence could be controlled in order to isolate the effect of MPR. This could be built with a metallic sphere, such as a ball bearing, set in a homogeneous background agent. A single acquisition could then be performed (any number of direct acquisitions will be geometrically identical) and reconstructed in any angle to determine the mathematical effect of MPR on noise.

### **4.32 Magnetic resonance imaging (156)**

To assess the relationship between receiver bandwidth (BW), voxel size ( $\Delta\chi$ ), metal artefact reduction and the penalties incurred by altering these imaging parameters a second phantom was constructed. This consisted of a Charnley Elite cobalt-chrome femoral stem embedded in 8 litres of solid cooking fat contained within a re-inforced polyethylene container. This was imaged using a standard clinical T1W sequence acquiring a single coronal slice through the long axis of the prosthesis. This was repeated keeping all imaging parameters; TE (Echo time), TR (Relaxation time), slice thickness and field of view (FOV) constant (there were small variations in TE at extreme BW settings but this was not felt to be significant). With each repetition the BW was increased from 150 Hz/pixel to 781 Hz/pixel in four intervals, and this was repeated with increasing matrix sizes from 125 x 125 to 768 x 768 pixels. The surface area of artefact, and the signal void from the prosthesis itself, on each image was calculated by segmenting out the background signal of the fat and subtracting this from the total known surface area. This approach was taken because the susceptibility caused by the prosthesis produced areas of both signal void and hyperintensity, which were not always connected on the image (and therefore required multiple seed points), whereas it appeared to be more reliable to segment out the fat, from a single seed point, using thresholds determined from the standard deviation of pixel values from an ROI defined at the edge of the phantom. The SNR



was calculated by using a conventional formula for calculating SNR in clinical images:

$$SNR = \frac{\text{mean tissue signal}}{SD_{\text{system}}}$$

where the mean signal in the tissue was defined as the mean of the ROI mentioned above and the SD caused by the imaging system was defined as the SD of air around the phantom.

The results of this study demonstrated that, as was already known, both BW and matrix size (the inverse of  $\Delta\chi$  for a fixed FOV) could be escalated to reduce metal artefact but that this effect was non-linear. As each of these parameters was increased there was a diminishing return in improvement in metal artefact so that 80% to 90% of achievable reduction in metal artefact occurred in the middle of the range of BW and matrix available on this clinical MR system (Figure 34).



**Figure 34. Three coronal T1W MR slices through the centre of the femoral prosthesis phantom demonstrating the most metal artefact reduction between low resolution and low BW (128x128, 150Hz/pixel) on the left and mid-range parameters (384x384, 449Hz/pixel) in the middle. The jump to high resolution and BW (768x768, 781Hz/pixel) has a relatively small effect on the amount of metal artefact but no noticeable change in SNR.**

SNR demonstrated a dramatic drop to 3% of the highest recorded SNR but interestingly, although there was an obvious granular texture caused by the drop in SNR, these images were still considered to be diagnostic by the research team. On the other hand the increase in matrix had a linear relationship with acquisition time (a function of increasing the steps in the frequency encoding gradient) with the largest matrix taking over 3 minutes to acquire a single slice. Although this is a concern because the longer acquisition times are more susceptible to movement artefact there are ways of ameliorating this in clinical practice, for instance by interleaving slice acquisitions.

The conclusion of the study is that controlling BW is the most powerful tool for reducing metal artefact on clinical MR systems. The penalties associated with increasing BW do not preclude a diagnostic study and therefore BW should be used in preference to matrix to increase the  $G_{FE}$ . However this may not apply to all clinical systems. Older machines may not have the range of BW available and in those cases a combination of BW and matrix control may be required. A second approach to imaging the phantom may reveal more useful information. A single coronal slice was assumed to be representative of the entire volume of artefact (the coronal slice should have been along the plane of symmetry) and therefore acquiring a volume dataset may have been more accurate. More importantly by acquiring a volume dataset it would have been possible to investigate a number of approaches to reducing the acquisition time for multiple slices and thereby come by a more favourable, non-linear relationship, of matrix ( $\Delta\chi$ ) to  $G_{FE}$ .

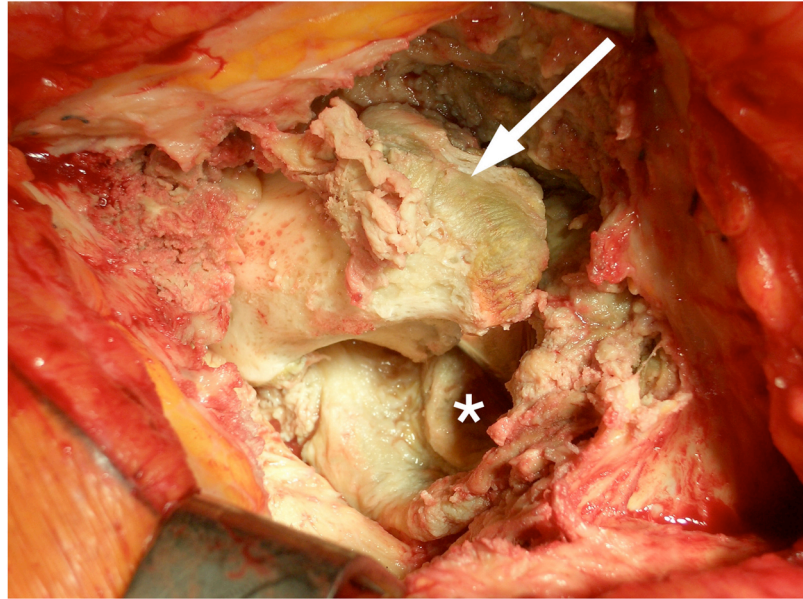
## **4.4 Clinical Studies**

### **4.41 Aseptic Lymphocytic Vasculitis-Associated Lesions (ALVAL)**

ALVAL was first described histologically in 2005 (158,159). A surgical and arthrographic description of necrosis associated with MOM THR from 1970s (160) suggest that the disease may have been seen much earlier but not identified as a discrete disease. It has, so far, only been described in patients with second generation MOM arthroplasties and is characterised by tissue necrosis, massive fibrin deposition, diffuse and dense perivascular lymphocytic infiltrate, occasional eosinophilic infiltrates and rarely microscopic metallic particles (161,162). It has been accompanied by an increasingly widespread appreciation that there is a large group of patients who suffer early failure with a number of MOM prostheses (161-165). The mechanism of this disease is not entirely understood but some consider the aetiology to be form of metal hypersensitivity (147,165) and others have suggested that a release of metal ions may follow galvanic corrosion (a

process that can occur when two different metals, in this case a titanium shell and a cobalt-chrome stem, are connected by an electrolyte solution) (166). The disease is different from metallosis, in which huge volumes of macroscopic metallic debris are shed in to the pseudocapsule of the THR by an abnormal MOM articulation. In metallosis metal debris forms a high attenuation deposit which can be identified on plain radiographs as a cloud-like opacity although these appearances can be confused with heterotopic ossification (HO). CT can correctly identify metallosis by localising the distribution of iron particles and distinguishing these opacities from the cortical and trabecula differentiation of HO (167). In comparison the metal shed into the soft tissues in ALVAL is present predominantly in ionic form and is only occasionally identified in the pseudocapsule of ALVAL by light microscopy (168). While not visible to the naked eye the cobalt-chromium ions are toxic and appear to cause extensive soft tissue necrosis.

In our practice ALVAL came to light in approximately 2004 when nearly 650 second generation metal-on-metal hips had been implanted in Norwich. A series of 12 patients, several of who were young, presented early in their post-operative course with pain, apparently normal plain radiographs and then at surgery extensive peri-prosthetic soft tissue necrosis (Figure 35). This triggered an investigation by the Medicines and Healthcare products Regulation Agency (MHRA)(169) and the following clinical studies.



**Figure 35. Intra-operative findings in ALVAL. The image demonstrates a posterior approach to the hip at revision with the prosthesis removed. The gluteal tendon (arrow) has avulsed from the greater trochanter and devitalised. The proximal femur (asterix) also failed to bleed when challenged with a curette indicating that it too was necrotic.**

#### **4.42 First description of MR appearances of ALVAL (161)**

This took the form of a retrospective case series and was the first publication in the scientific literature describing the MR appearances of ALVAL. The first 20 hips, in 19 patients, to present with pain early (within 3 years of operation) in their post-operative course were included in the study. All patients had undergone THR with a cobalt-chromium-on-cobalt-chromium alloy prosthesis (Ultima ® acetabular cup, 28mm Ultima ® head and an Ultima ® TPS stem: De Puy International Ltd., Leeds, UK). The study group comprised 11 women and 8 men aged from 37 to 74 years old. They all underwent MR examinations of the THR. The first seven patients underwent conventional MR studies and subsequently metal artefact reduction sequences (MARS), using the principles described above, were used for all further MR

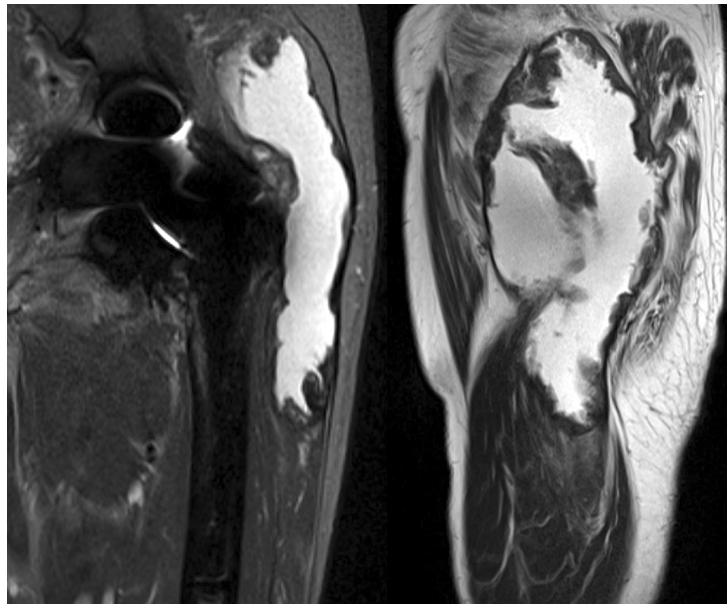
examinations. The surgical records, radiology and histology for all 19 patients were reviewed.

The plain radiographs for all patients were reviewed by a consultant radiologist and a consultant orthopaedic surgeon. The following measures were recorded on the first post-operative radiograph and the last radiograph prior to MRI: acetabular inclination, acetabular cup height, leg length, lateral offset and femoral stem angle as well as the presence of stress shielding, HO, osteolysis and cement mantle defects. All of these may be used to indicate failure of a prosthesis with time (104). The MR examinations were reviewed and scored for the presence of peri-prosthetic soft tissue abnormalities, tendon avulsions and bone marrow changes by three consultant radiologists in consensus. After MR 14 patients went on to revision surgery; the records for these patients were reviewed by a consultant orthopaedic surgeon and the histology was reviewed by a consultant histopathologist.

Descriptive statistics of the measures of plain radiographic appearances indicated that all the radiographs were within acceptable published normal ranges and that there was no significant difference between the measurements immediately after surgery and those prior to MR. In other words there was no plain radiographic evidence of failure of the THR.

In all 20 hips there was a peri-prosthetic soft tissue abnormality which consisted typically of a fluid filled (hyperintense on T2W, hypointense on T1W) cavity encircling or abutting the neck of the prosthesis (Figure 36). The cavity was demarcated by a “capsule” which was typically thick-walled and ragged and iso-intense to muscle on T1W and hypo-intense on T2W. The mean maximal diameter for the collection was 7cm and in most cases the collection extended in to the gluteal compartment. Other compartments around the hip were also involved but less frequently. Muscle atrophy was noted frequently in the gluteus medius and minimus as well as the short external rotators with gluteal tendon avulsion in 5 hips. Myositis, bone

marrow oedema and an occult fracture were diagnosed in a handful of cases.



**Figure 36. A coronal STIR<sup>10</sup> on the left and a sagittal T2W MR through a left MOM THR demonstrate a large fluid filled cavity, typical of the cystic mass found in ALVAL, surrounded by a ragged pseudocapsule extending lateral to the greater trochanter, cranial in to the gluteal compartment (where there is gluteal muscle atrophy) and caudal in to the lateral thigh.**

In all 14 revised hips a discrete peri-prosthetic soft tissue mass or fluid filled cavity lined by a thick avascular fibrinous membrane was noted.

Macroscopic necrosis was seen in over half of cases, with avascular soft tissue and/or bone in most. All the implants were firmly fixed but heavily corroded. Samples were sent for microscopy and culture from all cases but these yielded no growth. The predominant histological finding was one of necrosis and fibrin deposition. In just 5 patients was there the previously described characteristic peri-lymphocytic infiltrate of ALVAL and lesser numbers demonstrated features of eosinophilia, granulomas, inflammation

---

<sup>10</sup> Short tau inversion recovery

and one patient demonstrated metallic particles visible only by light microscopy.

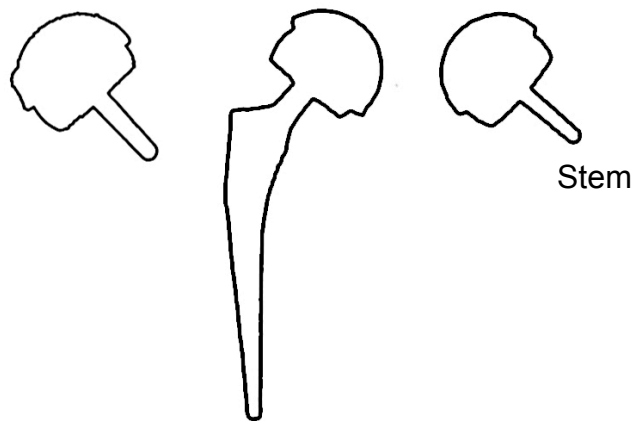
The conclusion of the study was that an as yet unknown number of patients with the MOM prosthesis described were failing early because of a disease process that was very similar to ALVAL. The MR appearances appear to be unique to this condition and, probably most importantly, this was the first instance in which a definitive role for MR in the imaging of THR had been described. Previous authors had described early experience with MR in complications of THR (145,147,150) but these had not led to subsequent widespread use of MR. This is because the conditions being diagnosed, such as gluteal avulsion, infection and polyethylene osteolysis were already being satisfactorily managed with a combination of clinical examination, plain radiographs and serological and haematological tests. With ALVAL it appeared that blood tests and x-rays were uniformly unhelpful and that MR might be the key to pre-operative diagnosis.

The limitations of the study left a number of questions unanswered. The radiological and histopathological data had not been collected systematically. There were no controls in the study which made it difficult to be certain whether all of the features described could be attributed to disease and what might be acceptable post-operative appearances. Up until then no reports of MRI in normal asymptomatic patients with THR had been published. The sample size was small, and likely to be biased to the most severe cases (presenting earliest) to be able to get a clear picture of what the complete spectrum of disease might be and there was no measure of reliability. A prospective study would be required to be able to correlate MR findings with clinical and histological measures of severity and the suspicion was that the high numbers of acellular histological samples might reflect inadequate biopsy technique; deeper samples were probably required to sample the cellular material outside the thick membranous pseudocapsule.



#### 4.43 Characteristics of ALVAL in other MOM prostheses (168,170)

The publication of these findings was followed shortly by a similar description of ALVAL in another MOM prosthesis; the Birmingham resurfacing THR (171,172)(Figure 37).

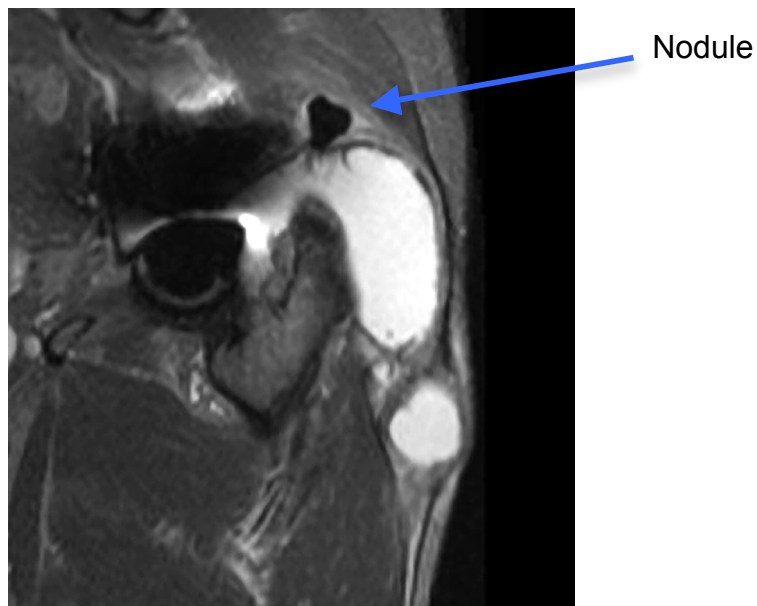


**Figure 37. Line drawing of the radiographic silhouettes of three second generation large bearing MOM THRs. On the left is a Birmingham, and on the right is an ASR<sup>11</sup>, both are resurfacing THRs which are difficult to distinguish on plain radiographs. These comprise a spherical shell that replaces the articular surface of the femoral head and which is anchored in the native femoral neck by a relatively short stem. The central THR is the same as that on the right; an ASR but with an uncemented conventional stem.**

Our experience of ALVAL in these prostheses was limited to a handful of cases but we were able to report the pre-operative diagnosis of lymphocytic spread of metal particles in a patient with ALVAL that presented only 6 months after undergoing bilateral resurfacing procedures (168). The patient presented bilateral hip discomfort and palpable swelling. On ultrasound examination there were large bilateral periprosthetic fluid filled cavities which extended through defects in the ilio-tibial tracts in to subcutaneous fat and

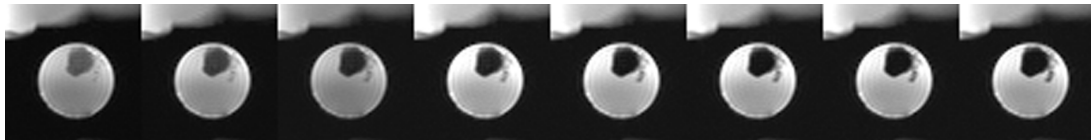
<sup>11</sup> Articular surface replacement hip – DePuy.

MR revealed the characteristic features of ALVAL. However a new feature was the presence of a small lobulated soft tissue mass (1.5cm diameter) which lay between gluteus medius and minimus some 6cm away from the prosthesis and, crucially, outside of the cystic ALVAL mass. This lesion demonstrated “blooming” in that the margins of the lesion were a little indistinct and very low signal on MR sequences with long echo times; a sign of metal deposition (Figure 38).



**Figure 38. Coronal STIR MR through a left Birmingham MOM THR demonstrating a markedly hypointense nodule in the left gluteal compartment just cranial and external to an aseptic periprosthetic fluid collection.**

The surgeon was directed to the lesion (this would not normally have been retrieved without the MR road map) and the specimen was re-examined with MR using a series of sequences with progressively long echo times which allowed us to generate a graph of signal intensity decay compared to echo time which demonstrated a complete loss of signal by 44ms (Figure 39).



**Figure 39. Series of MR images with increasingly long TE times demonstrating the loss of signal within the retrieved lymph node (within a specimen pot containing formalin) caused by the presence of metal debris.**

This is much shorter than would be expected from normal biological tissues and confirmed that the sample retrieved was the sample demonstrated at MR. Histology confirmed the presence of metal particles within macrophages as part of sinus histiocytosis (a condition where the sinuses within lymph nodes become distended and filled by histiocytes) and indicated for the first time that lymphoreticular spread of metal debris was possible with MOM prostheses.

The third and fourth MOM prostheses to present with ALVAL in our practice were the ASR acetabular components and matched cobalt-chrome large bearing heads implanted either as a resurfacing procedure or as an uncemented THR with a titanium (Corail – DePuy, Leeds, UK) stem. The MR appearances of ALVAL in these prostheses was not immediately recognised. In the first few patients the appearances were considered to be atypical of the appearance of ALVAL in the Ultima-TPS THR that we had previously described and it was not until histology from the revision procedures identified ALVAL that we considered this group separately in a prospective study. In all 75 ASR hip replacements (59 THR and 16 resurfacing) in 68 patients underwent MR examination (This was the whole cohort and not just symptomatic patients) (171). All of these MR studies were reported independently by two consultant radiologists and in 36% of patients there were features considered to be typical of ALVAL. The periprosthetic fluid collections, as described previously, were again a common feature however there were some notable differences. The cystic lesion often contained what appeared to be debris consisting of intermediate

MR signal whereas the Ultima-TPS associated lesions typically contained homogenous fluid signal. Whereas gluteal involvement was common, and gluteal avulsion not uncommon with the Ultima-TPS, these were uncommon findings in the ASR replacements.

After publication of our initial MR findings in the Ultima-TPS the discussion about the aetiology of this disease was often specific to that prosthesis and the implantation technique. For instance the exact proportions of the components of the alloy, the fixation cement and the engineering of the prosthesis were all considered as potential causes of ALVAL. However our subsequent experience clearly demonstrated that the only common link between the prostheses is the MOM bearing itself. Our experience included small and large bearing, stemmed and resurfacing, and cemented and uncemented prostheses and the same histological appearance was confirmed in all of them. However there were apparent differences in the MR appearances in the different groups. This might be because different alloys and different bearings produce different forms of ionic or metallic debris with different levels of toxicity but this needs to be considered with some caution. The Ultima-TPS and ASR groups had different demographics partly because of how we selected the cases. The Ultima-TPS study was retrospective and probably contained the most severely affected patients whereas the ASR study was prospective and cross-sectional. The different prostheses were implanted by different surgeons and their patient selection criteria for operating may not have been comparable. Despite this there do appear to be common MR appearances, typical of ALVAL across all the groups, and there is evidence that not only are MOM THR associated with elevated serum levels of cobalt and chromium ions but also lymphoreticular spread of metal debris.

#### **4.44 Clinical outcomes (165,171)**

The clinical outcomes for the Ultima-TPS and the ASR THRs in Norwich have been poor. The 5 year revision rate for the Ultima-TPS is 14% (165), and at

approximately 32 months it is 10% for the ASR prostheses (171). Of the 90 Ultima-TPS hips revised 45 had pre-operative MRI and in 41 (91%), on subsequent consensus review by two consultant radiologists, there were features considered to be typical of ALVAL. However as a retrospective study, that limited its analysis to patients who underwent revision surgery, there are significant limitations to what can be inferred about the diagnostic accuracy of MR. The decision to revise a patient is likely to have been significantly influenced by the results of the MR and therefore could create a self fulfilling prophecy of MR predicting outcome.

The ASR study, for radiologists, is more useful. In this study the data was collected systematically for the whole cohort of patients and MR studies were reported independently by two consultant radiologists prior to surgery. At the time of publication 8 patients (10%) had undergone revision. Of these eight patients six patients had MR that reported features of ALVAL and this was subsequently shown to be the case histologically. Two of the revision patients were reported as having normal post-operative appearances and there was no evidence of ALVAL on histology. At first this might seem encouraging; although the numbers are small the results seem to provide some evidence that MR might be an accurate test for ALVAL. But this is not the whole story. While at first there does appear to be an association between the presence of MR features of ALVAL and clinical status (Oxford Hip Score (OHS) mean 22.8 with normal MRI and mean 19.4 with features of ALVAL on MR; the higher the score the better) when subset analysis was performed comparing the grade of severity of the MR changes there appeared to be no correlation with symptoms. Patients with severe disease on MR imaging could be entirely asymptomatic.

#### **4.45 Asymptomatic patients and volunteers (70)**

One of the limitations of our initial, and subsequent, MR description of ALVAL on MRI was that we had no prior knowledge of an acceptable range of normal post-operative appearances after THR. We acknowledged that it

was likely that some of the abnormalities that we described were not related to MOM disease. In order to address this we carried out a pilot study of MRI in asymptomatic volunteers with either a POM or MOM THR. The study consisted of 10 patients in each group and demonstrated the following key features. The surgical approach to the THR influenced a number of the normal post-operative appearances. In particular a posterior approach to the hip, whereby the short external rotator (SER) muscles are divided and, in our institution, reattached to the greater trochanter after implantation of the prosthesis is invariably associated with atrophy of the SER muscles. This presents with wasting of, most notably on MRI, piriformis and obturator internus. While this appearance is normal, avulsion of the SER, as described in the Ultima-TPS but not the ASRs, is probably not (although some surgeons at other institutions do not reattach the SER). The lateral approach to the hip can cause wasting of the anterior half of the gluteus medius muscle and this may be mis-interpreted as a myopathy which can be seen in ALVAL. Interestingly gluteal muscle wasting was also identified in contra-lateral un-operated hips with OA and therefore may predate any intervention. Bone marrow oedema was described as part of the original and subsequent descriptions of ALVAL on MR but in asymptomatic volunteers the inter-rater reliability for reporting this feature was poor and therefore we would recommend caution when assigning significance to this. While thick walled fluid collections are a key feature of ALVAL, small thin walled fluid collections in the plane of the surgical approach, are normal and presumably represent benign post-operative seromas.

The results of this study confirmed our concerns about the significance of SER atrophy following a posterior approach to the hip and identified new areas which we now either consider to be normal or not robust MR features. The study was small and was deliberately designed as a pilot because there was no previous data on which to base a sample size estimate. The two groups (POM and MOM) were not directly comparable because by the time the study started the Ultima-TPS had stopped being used for several years and therefore the spread of times from operation was broader in the POM group and the age range was different with the MOM cohort being younger.

However the results suggest that a follow up study should start with pre-operative MR to assess OA related muscle loss and include a series of post-operative MR to assess stability or otherwise of supposed seromas around the hip.

#### **4.46 Grading of severity of ALVAL with MR (66)**

In April 2010 the MHRA issued a Medical Device Alert (169) covering all MOM THRs advising that all patients should be followed annually for a minimum of 5 years and that all symptomatic patients should be investigated with, among other tests, MRI. The Norwich cohort of Ultima-TPS has been under similar surveillance since 2008, and as described previously, the complete cohort of ASR hips had been imaged with MR. This surveillance will provide longitudinal data about ALVAL and in particular will allow us to correlate MR appearances with clinical outcomes. To do this will require an MR scoring system that categorises the severity of MR changes. Although we had some anecdotal opinions about what constituted severe changes the reliability of such a scoring system could not be evaluated until we had gained enough experience of MRI in asymptomatic and normal post-operative patients and patients with non-MOM complications.

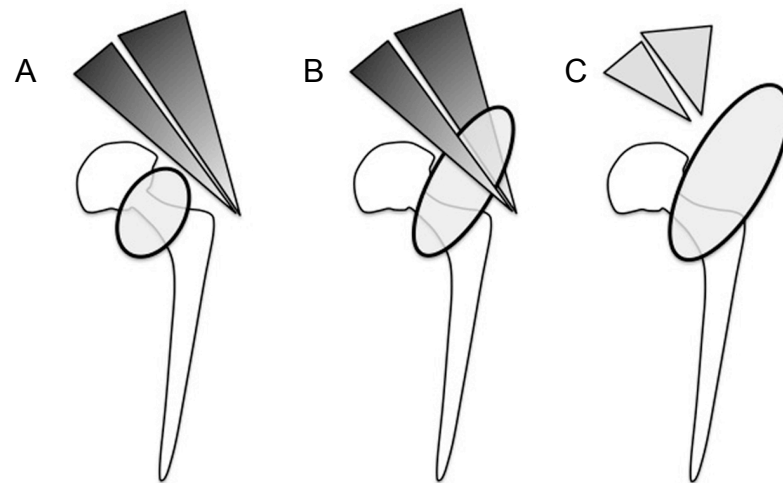
A scoring system to grade the severity of MR changes in MOM was devised as follows and for the following reasons (Table 10). Mild ALVAL was considered to comprise small peri-prosthetic soft tissues masses or collections less than 5 cm in maximal diameter without any other radiological abnormality. If the patient was asymptomatic or had only mild symptoms then the surgeon was typically prepared to wait and watch repeating the MR in 6 months. Moderate disease included peri-prosthetic masses or collection over 5 cm in maximal diameter. This measure was chosen because by this size the collection had typically extended in to one of the muscle compartments, typically the gluteal. Findings such as muscle oedema and atrophy, and bone marrow oedema was also considered to be moderately severe findings indicating that the disease had extended beyond the confines of the collection. In these patients routine elective revision of the prosthesis

was indicated. Severe disease included periprosthetic fractures, tendon avulsion and bone marrow replacement (Figure 40). Typically these patients would undergo urgent revision although initial reports indicate that by this stage the outcomes of revision are disappointing with poor functional outcomes (173).

Grade	Description	Criteria
A	Normal or acceptable	Normal post-op appearances Including seromas and small haematomas
B	Infection	Fluid filled cavity with high signal T2 wall Inflammatory changes in soft tissues +/-Bone marrow edema
C1	Mild MOM disease	Periprosthetic soft tissues mass with no hyperintense T2W fluid signal Or: Fluid filled peri-prosthetic cavity Either less than 5cm maximal diameter
C2	Moderate MOM disease	Periprosthetic soft tissue mass/fluid filled cavity greater than 5cm diameter Or C1 lesion with either of following Muscle atrophy or edema in any muscle other than short external rotators Bone marrow edema: hyperintense on STIR
C3	Severe MOM disease	Any one of the following Fluid filled cavity extending through deep fascia Tendon avulsion Intermediate T1W soft tissue cortical or marrow signal Fracture

**Table 10. Criteria for grading ALVAL using MR**





**Figure 40. Diagrammatic representation of the grading of ALVAL on MRI demonstrating the relationship of periprosthetic collection and the gluteus medius and minimus muscles and tendons. A. Mild disease comprises a small periprosthetic collection. B. The collection extends in to the gluteal compartment in moderate disease and causes gluteal tendon avulsion in severe disease (C).**

As discussed previously the inter-rater reliability for moderate and severe disease is good but it can be difficult to differentiate between mild ALVAL and other non-MOM complications such as infection. While the grading system, in part appears to be reliable this does not mean that it is valid. To test validity will require a correlation with a longitudinal outcome study. This is important for two reasons. The first is that our experience with the ASR hips suggests that measures of severity may not correlate well with symptoms. The second is that while it will be possible to correlate surgical outcomes in patients with moderate or severe disease, many of who will undergo surgery, correlation with mild disease will take time to confirm which patients progress to more severe disease and which remain stable.

## **4.5 Summary**

The story of the second generation MOM THR in Norwich has provided us with the opportunity to contribute a number of new pieces of knowledge and understanding to the world literature. We have been able to describe a new technique for metal artefact reduction in CT and define the optimal approach to metal artefact reduction on clinical MR systems. We have published the first clinical and radiological descriptions of a new disease, ALVAL, and followed this up with more in-depth understanding of the disease in four different second generation MOM prostheses with insights in to the pathophysiology of the disease and its relationship to symptomatology. We have further contributed the results of clinical studies of asymptomatic volunteers, and patients with MOM THR, which have allowed us to publish a scoring system for grading the severity of ALVAL with MRI.

## Conclusion

The digital evolution in radiology has made a significant impact in clinical medicine and opened new avenues of research. PACS ensures that a patient's radiological studies are never lost, accessible from multiple points and easily transportable. It also provides an archive for harvesting research material for population studies sometimes from seemingly redundant data. The DICOM image file format allows mathematicians to model human physiology *in vivo*, particularly with MR imaging, resulting in new insights in to tissue responses to disease and treatment. The computer workstations from which radiological studies are reported provide increasingly sophisticated tools for image analysis, some of which have now been demonstrated to be reliable, and superior to older analogue techniques. The contra-indications and indications for radiology continue to evolve and where once orthopaedic prostheses were an absolute contra-indication to MRI, in certain clinical circumstances, and with the correct technique, MR imaging is now an essential tool in the assessment of the total hip replacement.

## Appendix 1

### ***Publications***

This thesis is based on the following peer-review publications which are presented in the order that they are referred to in the text. My contribution to each of these papers is indicated in italics.

### **Chapter 1**

Crawford MB, Toms AP, Shepstone L. Defining normal vertebral angulation at the thoracolumbar junction. *AJR Am J Roentgenol*. 2009 Jul;193(1):W33-37.

*Senior and corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Fowkes LA, Petridou E, Zagorski C, Karuppiah A, Toms AP. Defining a reference range of acetabular inclination and center-edge angle of the hip in asymptomatic individuals. *Skeletal Radiol*. 2011 Nov;40(11):1427–34.

*Senior author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

### **Chapter 2**

Kingston A-R, Toms A, Ghosh-Ray S, Johnston-Downing S. Does running cause metatarsophalangeal joint effusions? A comparison of synovial fluid volumes on MRI in athletes before and after running. *Skeletal Radiology*. 2009 May 1;38(5):499-504.

*Senior and corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Patel SR, Toms AP, Rehman JM, Wimbhurst J. A reliability study of measurement tools available on standard picture archiving and communication system workstations for the evaluation of hip radiographs following arthroplasty. *J Bone Joint Surg Am*. 2011 Sep 21;93(18):1712–9.

*Senior author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Toms AP, Smith-Bateman C, Malcolm PN, Cahir J, Graves M. Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses. Clin Radiol. 2010 Jun;65(6):447-452.

*Corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Toms A, Cahir J, Swift L, Donell S. Imaging the femoral sulcus with ultrasound, CT, and MRI: reliability and generalizability in patients with patellar instability. Skeletal Radiology. 2009 Apr 1;38(4):329-338.

*Corresponding author, design of investigation, LREC approval, conduct of research, analysis of outcome, preparation for publication.*

Anderson H, Toms AP, Cahir JG, Goodwin RW, Wimhurst J, Nolan JF. Grading the severity of soft tissue changes associated with metal-on-metal hip replacements: reliability of an MR grading system. Skeletal Radiol. 2011 Mar;40(3):303-307.

*Senior and corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Toms AP, Kasmai B, Williams S, Wilson P. Building an anonymized catalogued radiology museum in PACS: a feasibility study. Br J Radiol. 2006 Aug 1;79(944):666-671.

*Corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

### Chapter 3

Toms AP, White LM, Kandel R, Bleakney RR, Noseworthy M, Lee S, et al. Limitations of single slice dynamic contrast enhanced MR in pharmacokinetic modeling of bone sarcomas. Acta Radiol. 2009 Jun;50(5):512-520.

*Corresponding author, IRB approval, conduct of research, analysis of outcome, preparation for publication.*

### Chapter 4

Toms AP, Botchu R, Nolan JF. (ii) Diagnostic plain film radiology of the failing hip replacement. Orthopaedics and Trauma. 2009 Apr;23(2):88-100.

*Corresponding author, preparation for publication.*

Cahir JG, Toms AP, Marshall TJ, Wimhurst J, Nolan J. CT and MRI of hip arthroplasty. Clin Radiol. 2007 Dec;62(12):1163-1171; discussion 1172-1173.

*Co-author, preparation for publication.*

Cahir JG, Toms AP. CT and MRI of hip replacements. Orthopaedics and Trauma. 2009 Apr;23(2):101-108.

*Co-author, preparation for publication.*

Donell ST, Darrah C, Nolan JF, Wimhurst J, Toms A, Barker THW, et al. Early failure of the Ultima metal-on-metal total hip replacement in the presence of normal plain radiographs. J Bone Joint Surg Br. 2010 Nov;92(11):1501-1508.

*Co-author, conduct of research, analysis of outcomes, preparation for publication.*

Khan RJK, Wimhurst J, Foroughi S, Toms A. The natural history of metallosis from catastrophic failure of a polyethylene liner in a total hip. The Journal of Arthroplasty. 2009;24(7):1144.

*Co-author, preparation for publication.*

Lewis M, Toms AP, Reid K, Bugg W. CT metal artefact reduction of total knee prostheses using angled gantry multiplanar reformation. Knee. 2010 Aug;17(4):279-282.

*Senior author, design of investigation, conduct of research, analysis of outcome, preparation for publication*

Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donell ST, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. Clin Radiol. 2008 Jan;63(1):49-58.

*Corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Toms AP, Nolan J, Barker T, Darrah C, Malcolm P. Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR. Br J Radiol. 2009 May;82(977):e87-91.

*Corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Toms AP, Smith-Bateman C, Malcolm PN, Cahir J, Graves M. Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses. Clin Radiol. 2010 Jun;65(6):447-452.

*Corresponding author, design of investigation, conduct of research, analysis of outcome, preparation for publication.*

Mistry A, Cahir J, Donell ST, Nolan J, Toms AP. MRI of asymptomatic patients with metal-on-metal and polyethylene-on-metal total hip arthroplasties. Clin Radiol. 2011 Jun;66(6):540-545.

*Senior author, design of investigation, LREC approval, conduct of research, analysis of outcome, preparation for publication.*

Wynn-Jones H, Macnair R, Wimbhurst J, Chirodian N, Derbyshire B, Toms A, et al. Silent soft tissue pathology is common with a modern metal-on-metal hip arthroplasty. Acta Orthop. 2011 Jun;82(3):301–7.

*Co-author, conduct of research, analysis of outcomes, preparation for publication.*

## Appendix 2

### ***Published Abstracts***

Mistry A, Toms AP, Cahir J, Donnell ST, Nolan J. Asymptomatic metal-on-metal (MOM) and polyethylene-on-metal (POM) total hip replacements (THR), spectrum and comparison of MRI findings. Proceedings of the UK Radiological Congress 2010; p36

Anderson HL, Toms AP, Cahir JG, Goodwin RG, Nolan JF, Donnell ST. Validation of an MR grading system for scoring soft tissue disease in metal-on-metal hip replacements. Proceedings of the UK Radiological Congress 2010; p36

Patel SR, Toms AP, Rehman J, Wimhurst J. A reliability study of measurement tools available on standard PACS workstations for the evaluation of the post-arthroplasty radiograph. In: Radiological Society of North America scientific assembly and annual meeting program. Oak Brook, Ill: Radiological Society of North America, 2009; 596.

Toms AP, Smith C, Malcolm PNM. Optimisation of metal artefact reduction sequences (MARS) for MR imaging of total hip prostheses. Eur Radiol 2009; 19:S201

Ghosh-Ray S, Rosa S, Toms AP, Clark A. Automated segmentation of volumes on MR imaging: Validation, accuracy and relationship of error to imaging parameters. Eur Radiol 2009; 19:S158

Toms AP, White LM, Kandel R, Bleakney RR, Noseworthy MD, Shepstone L, Blackstein ME, Wunder J. Dynamic contrast enhanced MRI of spindle cell sarcomas: Counterintuitive relationship of perfusion to tumour necrosis. Eur Radiol 2008; 18(S1): 306

Toms AP, Chojnowski A, Cahir JG. Midcarpal instability: A diagnostic role for dynamic carpal ultrasound? Eur Radiol 2008; 18(S1): 199

Crawford M, Toms AP, Shenoy V, Shepstone L. Defining the normal range of anatomical anterior vertebral “wedging” at the thoraco-lumbar junction. Radiology 2006.

Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donnell ST, Barker T, Tucker JK. MR Imaging of early failing metal-on-metal total hip arthroplasty with surgical and histological correlation. Eur Radiol 2006; 16 (S1): 164



## ***Scientific presentations***

### **Oral**

Wynn Jones H, Wimhurst J, Macnair R, Chirodian N, Derbyshire B, Toms A, Cahir J. High early failure rate with a modern large head metal on metal hip replacement: the importance of metal artefact reduction MRI imaging. Combined Meeting of the Orthopaedic Associations, Glasgow. 13th September 2010.

Mistry A, Toms AP, Cahir J, Donnell ST, Nolan J. Asymptomatic metal-on-metal (MOM) and polyethylene-on-metal (POM) total hip replacements (THR), spectrum and comparison of MRI findings. UK Radiological Congress, Birmingham. 9<sup>th</sup> June 2010

Anderson HL, Toms AP, Cahir JG, Goodwin RG, Nolan JF, Donnell ST. Validation of an MR grading system for scoring soft tissue disease in metal-on-metal hip replacements. UK Radiological Congress, Birmingham. 9<sup>th</sup> June 2010

Patel S, Toms AP, Rehman J, Wimhurst J. A Reliability Study of Measurement Tools Available on Standard PACS (Picture Archiving and Communications System) Workstations for the Evaluation of the Post-arthroplasty Radiograph. Radiological Society of North America, Chicago 8<sup>th</sup> December 2009.

Toms AP, Smith C, Malcolm PN. Optimisation of metal artefact reduction sequences (MARS) for MR imaging of total hip prostheses. European Congress of Radiology, Vienna. 7<sup>th</sup> March 2009

Ghosh-Ray S, Rosa S, Toms AP, Clark A. Automated segmentation of volumes on MR imaging: Validation, accuracy and relationship of error to imaging parameters. European Congress of Radiology, Vienna. 6<sup>th</sup> March 2009

Fowkes L, Karuppiyah A, Toms A. Acetabular Inclination in a "Normal" Population. British Association of Clinical Anatomists Winter Scientific Meeting, Norwich. 18<sup>th</sup> December 2008

Toms AP, White LM, Kandel R, Bleakney RR, Noseworthy MD, Shepstone L, Blackstein ME, Wunder J. Dynamic contrast enhanced MRI of spindle cell sarcomas: Counterintuitive relationship of perfusion to tumour necrosis. European Congress of Radiology, Vienna. 11<sup>th</sup> March 2008.

Crawford M, Toms AP, Shenoy V, Shepstone L. Defining the normal range of anatomical anterior vertebral "wedging" at the thoraco-lumbar junction. Radiological Society of North America. Chicago 1<sup>st</sup> December 2006

Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donnell ST, Barker T, Tucker JK. MR imaging of early failing metal-on-metal total hip arthroplasty with surgical and histological correlation. European Congress of Radiology. Vienna. 3<sup>rd</sup> March 2006.

## Posters

Yanny S, Cahir J, Barker T, Wimhurst J, Nolan JF, Toms AP. MRI of aseptic lymphocytic vasculitis-associated lesions (ALVAL) in metal-on-metal hip arthroplasties: a pictorial review. International Skeletal Society, Athens 2010. Winner of the prize for Best Scientific Poster.

Anderson H, Toms AP, Cahir JG, Goodwin RG, Nolan JF, Donnell ST. Validation of an MR grading system for scoring soft tissue disease in metal-on-metal hip replacements. European Congress of Radiology. Vienna March 2010.

Mistry A, Toms AP, Cahir JG, Donnell ST, Nolan JF. Asymptomatic total hip replacements; a cross-sectional study defining normal MR appearances. European Congress of Radiology. Vienna March 2010.

Patel S, Toms AP, Rehman J, Wimhurst J. A Reliability Study of Measurement Tools Available on Standard PACS (Picture Archiving and Communications System) Workstations for the Evaluation of the Post-arthroplasty Radiograph. Radiological Society of North America, Chicago 2009.

## Definitions

$\alpha$	Acetabular inclination angle
A	Mean arterial slope
AI	Acetabular inclination
ALVAL	Aseptic lymphocytic vasculitis-associated lesions
AP	Antero-posterior
$B_0$	$B_0$ magnetic field
BW	Receiver bandwidth
CEA	Centre-edge-angle
CI	Confidence intervals
CNC	Computer numeric control
CT	Computed tomography
$\delta$	Pelvic tilt angle
DCE-MRI	Dynamic contrast-enhanced magnetic resonance imaging
DDH	Developmental dysplasia of the hip
DEXA	Dual energy x-ray absorptiometry
DICOM	Digital imaging and communications in medicine
$\Delta\chi$	Voxel size
EES	Extravascular extracellular space
FAI	Femoro-acetabular impingement
FOI	Foramen obturator index
FOV	Field of view
G	Generalizability coefficient
$\gamma$	Gyromagnetic ratio
$G_{FE}$	Frequency encoding gradient
GT	Generalizability theory
HO	Heterotopic ossification
HU	Hounsfield units
Hz	Hertz
ICC	Intraclass correlation coefficient
$\kappa$	Kappa coefficient

$K_f$	Knee facet
$K^{\text{trans}}$	Endothelial permeability coefficient
kVp	Peak kilovoltage
MARS	Metal artefact reduction sequences
mAs	Milliampere second
MDCT	Multidetector CT
MHRA	Medicines and Healthcare products Regulation Agency
MOM	Metal-on-metal
MPR	Multi-planar reformat
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MTPJ	Metatarsophalangeal joint
NMR	Nuclear magnetic resonance
NNUH	Norfolk & Norwich University Hospital
OA	Osteoarthritis
OHS	Oxford hip score
PACS	Picture archiving and communication systems
PBPK	Physiologically-based pharmacokinetic modelling
PD	Proton density
$P_f$	Patient facet
POM	Polyethylene-on-metal
R	Reliability
ROI	Region of interest
SER	Short external rotator
SCP	Sacro-coccygeal-pubis
SD	Standard deviation
$\Sigma f$	Sum of all facets
SNR	Signal to noise ratio
STARD	Standards for the reporting of diagnostic accuracy studies
STIR	Short tau inversion recovery
SUFE	Slipped upper femoral epiphysis
T	Tesla
T1W	T1 weighted
T2W	T2 weighted

TE	Echo time
THR	Total hip replacement
TKR	Total knee replacement
TR	Relaxation time
US	Ultrasound
$v_e$	Extravascular extracellular space volume

## Glossary

### **Absolute agreement**

Subtype of interclass and intraclass correlation coefficient that provides a measure of the reliability of a test for any observer.

### **Acetabular inclination**

The measure of infero-medial to supero-lateral position of the plane of the acetabulum on an antero-posterior radiograph of the pelvis relative to the inter-tear drop base line. The plane of the acetabulum is defined as a line drawn from the supero-lateral margin to the inferior aspect of the adjacent tear-drop.

### **Angiogenesis**

The proliferation of new blood vessels.

### **Aseptic lymphocytic vasculitis-associated lesions (ALVAL)**

Disease associated with metal-on-metal hip replacements that is characterised by peri-prosthetic soft tissue necrosis and a specific histological picture of perivascular lymphocytic infiltrates.

### **B<sub>0</sub> magnetic field**

The magnetic field of an MR machine.

### **Beam hardening artefact**

Artefact consisting of high and low attenuation streaks radiating from metalwork in a starburst pattern on CT.

### **Bland-Altman plot**

A data plot used to analyse differences between two different datasets. Also known as a *difference plot* and a *Tukey mean-difference plot*.

**Bone marrow oedema**

An inexact description of fluid signal changes on MR characterised by hyperintense signal on T2W fat saturated sequences and hypointense signal on T1W sequences.

**Centre-edge-angle**

A measure of lateral cover of the femoral head by the supero-lateral acetabular roof on an antero-posterior radiograph. The measure is obtained by subtending an angle between two lines drawn from the centre of the femoral head; one to the superolateral acetabulum and one vertically parallel to the central axis of the body.

**Centre of rotation**

The position of the left femoral head, when considered as a bearing, in relation to the pelvis. The centre of each femoral head is considered in relation to the inferior aspect of the adjacent tear-drop.

**Classical test theory**

Psychometric theory aimed at understanding and improving the reliability of psychological tests.

**Clearance model**

Pharmacokinetic model used in DCE-MRI to describe the movement of gadolinium from the extracellular extravascular space back in to blood plasma as the chelate is excreted by the kidneys.

**Cobalt-chromium**

Short hand description of the most common steel alloy used in the manufacture of orthopaedic implants. The other major constituents being iron and carbon.

**Cohen's kappa coefficient**

A statistical measure of inter-rater agreement

**Correlation**

Statistical relationship between two or more sets of data.

**CT radiograph**

A planar digital radiograph acquired using a CT machine, at the beginning of each CT examination, which is used to plan the position and extent of subsequent CT images.

**Consistency**

Subtype of intraclass and interclass correlation coefficient that indicates reliability of a particular set of observations.

**Digital imaging and communications in medicine (DICOM)**

A standard which defines the protocols that ensure the connectivity of medical imaging instruments with networks, workstations and image archives. The acronym DICOM is also applied to a specific image file format to which all manufacturers adhere when exporting digital medical images

**Dynamic contrast enhanced MR imaging (DCE-MRI)**

Technique for describing the perfusion characteristic of tissues using MR imaging.

**Endothelial permeability coefficient**

Volume transfer constant for gadolinium crossing from blood plasma to the extravascular extracellular space that is a reflection of endothelial permeability and known as  $K^{\text{trans}}$ .

**Endplate angle**

The angle subtended by two lines representing the superior and inferior vertebral endplates.



**Extravascular extracellular space**

One of the two compartments considered during modified two compartment modelling of DCE-MRI data (EES). The space outside of the blood vessels and between the cells. The volume of which is denoted by the symbol  $v_e$ .

**Facet**

Name given to discrete entities that might independently influence reliability within Generalizability theory.

**Foramen obturator index**

The ratio of the maximal medial to lateral dimensions of the obturator foramina on an antero-posterior radiograph of the pelvis which provides a measure of axial rotation of the pelvis.

**Frequency encoding gradient**

Radiofrequency gradient applied to the  $B_0$  magnetic field that causes a gradual change in the frequency of precession (spin) of protons along its length. These differences in precession can be recognised and use to spatially locate the signal.

**Gadolinium**

Rare earth metal that is superparamagnetic. When chelated can be used as an intravascular contrast agent in MR imaging.

**Gantry**

The part of a CT machine that house the x-ray tube as it rotates around the patient.

**Generalizability theory**

A statistical method for analysing the reliability of test by considering all sources of variance that might influence outcomes.

**Gyromagnetic ratio**

Property of a nucleus that determines its frequency of precession within a given magnetic field.

**Hounsfield unit**

Measure of attenuation on CT named after Sir Godfrey Hounsfield the inventor of CT.

**Interclass correlation**

Correlation between two or more different measures.

**Inter-rater reliability**

The consistency of observations between two or more observers.

**Intraclass correlation**

Correlation between two or more sets of observations of the same measure.

**Intra-rater reliability**

The consistency of repeated observations by one observer.

**Lateral offset**

The relative difference in lateral position of the proximal femora in relation to the pelvis measured as the difference in length between the two lesser trochanters and a line drawn from the tear-drop perpendicular to the inter-tear drop line.

**Leg length**

The relative difference in cranio-caudal position of the proximal femora in relation to the pelvis measured as the difference in length between the two lesser trochanters and the inter-tear drop line.

**Limits of agreement**

95% confidence limits in the difference between two observers used to determine whether differences are consistent or random (See Bland-Altman plot).

**Lymphoreticular**

Of cells and tissues of the lymphoreticular system which includes lymphoid and reticuloendothelial systems

**Matrix**

Grid of pixels or voxels that comprises a digital image.

**Metal artefact**

See *Susceptibility artefact*

**Metal artefact reduction sequences (MARS)**

MR imaging sequences optimised to reduce the effects of susceptibility to metal.

**Metallosis**

Condition where failure of an arthroplasty leads to large amounts of metal debris being shed in to the pseudocapsule of the joint.

**Metal-on-metal**

Term used to describe a hip replacement where the bearing comprises a metal acetabular cup and a metal femoral head component.

**Multiplanar reformation**

The reconstruction of an image from a three dimensional dataset in a plane other than the plane the data was acquired in.

**Normal range**

Statistically defined as two standard deviations above and below the mean, that includes 95% of a normally distributed dataset.

**Osteolysis**

Resorption of bone manifesting as lucencies on conventional radiographs.

**Parallax**

In radiography refers to the distortion in a radiograph caused by the projection of a cone beam of x-rays through the subject onto a flat radiographic plate.

**Parametric map**

Coloured code map of an endpoint measure that is typically, in radiological research, superimposed upon a corresponding radiological anatomical image. Each pixel is spectral colour coded to represent a magnitude of the particular measure at that point in space.

**Pearson product-moment correlation coefficient**

A statistical measure of linear correlation between two continuous datasets.

**Phantom**

A model that is design to represent a simplified element of the human body which can then be imaged without complications of radiation exposure, movement artefact, magnetic field inhomogeneity or ethical limitations.

**Pharmacokinetic modelling**

The mathematical description of the concentration, distribution and movement of compounds between compartments within a body.

**Photopaenia**

A focal reduction in the number of x-rays reaching a detector or film.

**Picture archiving and communication systems (PACS)**

A network of radiological image acquisition machines linked to an archive which can be accessed by a network of workstations for image review.

**Polyethylene-on-metal**

A generic description of a total hip replacement where the acetabular cup is made from polyethylene and the head of the femoral prosthesis is metallic.

**Receiver bandwidth**

The range between the highest and lowest frequencies of signal detected by the receiver coils contributing to the final MR image.

**Reliability**

The consistency of a test which in radiology can take a number of forms including intra-rater, inter-rater and test-retest reliability.

**Reference range**

Sometimes used instead of the term *normal range* to distinguish the statistical definition of normality from clinical normality.

**Region of interest**

A part of an image which is the focus of analysis and to which segmentation techniques are applied.

**Resurfacing**

Type of hip replacement where the femoral neck is preserved and the articular surface is replaced by a large metal-on-metal bearing.

**Sacro-coccygeal-pubis distance**

Indirect measurement of pelvic tilt measured from an antero-posterior radiograph of the pelvis.

**Segmentation**

In image analysis segmentation refers to the partitioning and simplification of sections of an image according to pre-defined criteria in order to analyse the properties of the resulting segment.

**Short external rotators**

Group of muscles that arise from the pelvis and insert in to the proximal femur and act to externally rotate the leg. They comprise piriformis, the gemelli, obturator internus and quadratus femoris.

**Signal-to-noise ratio**

The ratio of information within an image that is derived from the subject of that image (signal) to the random data that is generated by the instruments and processes required to produce the image.

**Sinus histiocytosis**

Condition where sinuses within lymph nodes become filled with histiocytes.

**Small particle granulomas**

Foreign body granulomas stimulated by polyethylene debris from acetabular cups. Peri-prosthetic enlargement of the granulomas causes osteolysis and therefore loosening of the prosthesis.

**Standards for the reporting of diagnostic accuracy studies (STARD)**

Initiative designed to standardise the reporting of radiological research studies dealing with accuracy in order to improve external evaluation of the findings.

**Stem angle**

The angle subtended by the central axis of the femoral component of a THR to the central axis of the host medulla cavity.

**Susceptibility artefact**

Artefact resulting in signal loss and geometric distortion resulting from magnetisation of ferrous material within a clinical MR machine.

**Test-retest reliability.**

The consistency of a repeated application of a test to the same subject.

**Trochlear dysplasia**

Abnormal development of the trochlea (femoral sulcus) typically resulting in a shallow or malformed notch.

**Voxel**

Word derived from *volume element* analogous to pixel (*picture element*) but in three dimensions.

**Weighted kappa coefficient**

Agreement coefficient that applies a weighting to the degree of disagreement when comparing ordinal data.

## References

1. Jones WA. The Role of Anatomy in the Radiological Study of the Spine. *Can Med Assoc J.* 1936;34(3):265-269.
2. Schweitzer ME, Hodler J, Cervilla V, Resnick D. Craniovertebral junction: normal anatomy with MR correlation. *AJR Am J Roentgenol.* 1992;158(5):1087-1090.
3. Black DM, Cummings SR, Stone K, Hudes E, Palermo L, Steiger P. A new approach to defining normal vertebral dimensions. *Journal of Bone and Mineral Research.* 1991;6(8):883-892.
4. Choplin RH, Boehme JM, Maynard CD. Picture archiving and communication systems: an overview. *Radiographics.* 1992;12(1):127.
5. PACS Stakeholder brochure: celebrating success - and looking to the future [Internet]. 2008 [cited 2011 Feb 3];Available from: [www.connectingforhealth.nhs.uk/.../pacs/materials/stakebrochure.pdf](http://www.connectingforhealth.nhs.uk/.../pacs/materials/stakebrochure.pdf)
6. Crawford MB, Toms AP, Shepstone L. Defining normal vertebral angulation at the thoracolumbar junction. *AJR Am J Roentgenol.* 2009;193(1):W33-37.
7. Gallagher JC, Hedlund LR, Stoner S, Meeger C. Vertebral morphometry: normative data. *Bone and mineral.* 1988;4(2):189.
8. Holmes JF, Miller PQ, Panacek EA, Lin S, Horne NS, Mower WR. Epidemiology of thoracolumbar spine injury in blunt trauma. *Academic Emergency Medicine.* 2001;8(9):866-872.
9. Lenchik L, Rogers LF, Delmas PD, Genant HK. Diagnosis of Osteoporotic Vertebral Fractures: Importance of Recognition and Description by Radiologists. *Am. J. Roentgenol.* 2004;183(4):949-958.
10. Davies KM, Recker RR, Heaney RP. Normal vertebral dimensions and normal variation in serial measurements of vertebrae. *Journal of bone and mineral research.* 1989;4(3):341-349.
11. McCloskey EV, Spector TD, Eyres KS, Fern ED, O'rourke N, Vasikaran S, et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporosis International.* 1993;3(3):138-147.
12. Dai L. Radiographic measurements of the heights of vertebral bodies in thoracic and lumbar spine. *Chinese Med Sci J.* 1996;11(2):117.



13. Hurxthal LM. Measurement of anterior vertebral compressions and biconcave vertebrae. *American Journal of Roentgenology*. 1968;103(3):635.
14. Vedantam R, Lenke LG, Keeney JA, Bridwell KH. Comparison of standing sagittal spinal alignment in asymptomatic adolescents and adults. *Spine*. 1998;23(2):211.
15. Gangnet N, Dumas R, Pomoero V, Mitulescu A, Skalli W, Vital JM. Three-dimensional spinal and pelvic alignment in an asymptomatic population. *Spine*. 2006;31(15):E507.
16. Roussouly P, Gollogly S, Berthonnaud E, Dimnet J. Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position. *Spine*. 2005;30(3):346.
17. Vialle R, Levassor N, Rillardon L, Templier A, Skalli W, Guigui P. Radiographic analysis of the sagittal alignment and balance of the spine in asymptomatic subjects. *J Bone Joint Surg Am*. 2005;87(2):260–7.
18. Eng J. Sample size estimation: how many individuals should be studied? *Radiology*. 2003;227(2):309–13.
19. Fowkes LA, Petridou E, Zagorski C, Karuppiah A, Toms AP. Defining a reference range of acetabular inclination and center-edge angle of the hip in asymptomatic individuals. *Skeletal Radiol*. 2011;40(11):1427–34.
20. Murray DW. The definition and measurement of acetabular orientation. *J Bone Joint Surg Br*. 1993;75(2):228-232.
21. Giori NJ, Trousdale RT. Acetabular retroversion is associated with osteoarthritis of the hip. *Clinical Orthop Rel Res*. 2003;417:263.
22. Wagner S, Hofstetter W, Chiquet M, Mainil-Varlet P, Stauffer E, Ganz R, et al. Early osteoarthritic changes of human femoral head cartilage subsequent to femoro-acetabular impingement. *Osteoarthr. Cartil*. 2003;11(7):508-518.
23. Tönnis D, Heinecke A. Acetabular and femoral anteversion: relationship with osteoarthritis of the hip. *J Bone Joint Surg Am*. 1999;81(12):1747-1770.
24. Ganz R, Leunig M, Leunig-Ganz K, Harris WH. The etiology of osteoarthritis of the hip. *Clinical Orthopaedics and Related Research*. 2008;466(2):264-272.
25. Steppacher SD, Tannast M, Werlen S, Siebenrock KA. Femoral morphology differs between deficient and excessive acetabular

- coverage. *Clinical Orthopaedics and Related Research*. 2008;466(4):782-790.
26. Vossinakis IC, Georgiades G, Kafidas D, Hartofilakidis G. Unilateral hip osteoarthritis: can we predict the outcome of the other hip? *Skeletal radiology*. 2008;37(10):911-916.
  27. Siebenrock KA, Kalbermatten DF, Ganz R. Effect of pelvic tilt on acetabular retroversion: a study of pelves from cadavers. *Clin Orthop Relat Res*. 2003 Feb;(407):241-248.
  28. Nagao Y, Aoki H, Ishii S-J, Masuda T, Beppu M. Radiographic method to measure the inclination angle of the acetabulum. *J Orthop Sci*. 2008;13(1):62-71.
  29. Kalteis TA, Handel M, Herbst B, Grifka J, Renkawitz T. In vitro investigation of the influence of pelvic tilt on acetabular cup alignment. *J Arthroplasty*. 2009;24(1):152-157.
  30. Haenle M, Heitner A, Mittelmeier W, Barbano R, Scholz R, Steinhäuser E, et al. Assessment of cup position from plain radiographs: impact of pelvic tilting. *Surgical and Radiologic Anatomy*. 2007;29(1):29-35.
  31. Armbuster TG, Guerra J Jr, Resnick D, Goergen TG, Feingold ML, Niwayama G, et al. The adult hip: an anatomic study. Part I: the bony landmarks. *Radiology*. 1978;128(1):1-10.
  32. Stulberg SD, Harris WH. Acetabular dysplasia and development of osteoarthritis of the hip. In: *The Hip. Proceedings of the Second Open Scientific Meeting of the Hip Society*. St Louis, MO: Mosby. 1974. p.82-93.
  33. Sharp IK. Acetabular dysplasia: the acetabular angle. *Journal of Bone and Joint Surgery-British Volume*. 1961;43(2):268.
  34. Nishihara S, Sugano N, Nishii T, Ohzono K, Yoshikawa H. Measurements of pelvic flexion angle using three-dimensional computed tomography. *Clin. Orthop. Relat. Res*. 2003 Jun;(411):140-151.
  35. Perreira AC, Hunter JC, Laird T, Jamali AA. Multilevel measurement of acetabular version using 3-D CT-generated models: implications for hip preservation surgery. *Clin Orthop Relat Res*. 2011;469(2):552-561.
  36. Zheng G, Zhang X. Computer assisted determination of acetabular cup orientation using 2D-3D image registration. *Int J Comput Assist Radiol Surg*. 2010;5(5):437-447.

37. Richards PJ, Pattison JM, Belcher J, Decann RW, Anderson S, Wynn-Jones C. A new tilt on pelvic radiographs: a pilot study. *Skeletal Radiol.* 2009;38(2):113-122.
38. Tannast M, Zheng G, Anderegg C, Burckhardt K, Langlotz F, Ganz R, et al. Tilt and rotation correction of acetabular version on pelvic radiographs. *Clin. Orthop. Relat. Res.* 2005 Sep;438:182-190.
39. Streiner D, Norman G. *Health Measurement Scales: A practical guide to their development and use.* 3rd ed. OUP Oxford; 2003.
40. Mackenzie R, Dixon AK. Measuring the effects of imaging: an evaluative framework. *Clin Radiol.* 1995;50(8):513–8.
41. Sardanelli F, Hunink MG, Gilbert FJ, Di Leo G, Krestin GP. Evidence-based radiology: why and how? *Eur Radiol.* 2010;20(1):1-15.
42. Bossuyt PMM. STARD statement: still room for improvement in the reporting of diagnostic accuracy studies. *Radiology.* 2008;248(3):713-714.
43. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428.
44. Fisher RA. *Statistical methods for research workers.* 9th ed. London: Oliver and Boyd Ltd; 1944.
45. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological methods.* 1996;1(1):30-46.
46. Cohen J. A coefficient of agreement for nominal scales. *Educational and psychological measurement.* 1960;20(1):37-46.
47. Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin.* 1968;70(4):213-220.
48. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics.* 1977;33(2):363-374.
49. Altman DG. *Practical Statistics for Medical Research.* 1st ed. Chapman & Hall; 1990.
50. Shrout PE. Measurement reliability and agreement in psychiatry. *Statistical methods in medical research.* 1998;7(3):301.
51. Nunnally JC, Bernstein RH. *The Assessment of reliability. Psychometric Theory.* New York: McGraw-Hill, Inc. 1994:248-292.

52. Petrie A, Sabin C. Medical Statistics at a Glance. 3rd ed. Wiley-Blackwell; 2009.
53. Cronbach LJ, Rajaratnam N, Gleser GC. Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*. 1963;2:137-63.
54. Alemparte J, Ekdahl M, Burnier L, Hernández R, Cardemil A, Cielo R, et al. Patellofemoral evaluation with radiographs and computed tomography scans in 60 knees of asymptomatic subjects. *Arthroscopy*. 2007;23(2):170-177.
55. Beals RK, Buehler K. Treatment of patellofemoral instability in childhood with creation of a femoral sulcus. *J Pediatr Orthop*. 1997;17(4):516-519.
56. Bonnard C, Nocquet P, Sollogoub I, Glorion B. [Patellar instability in children. Result of transposition of the medial third of patellar tendon]. *Rev Chir Orthop Reparatrice Appar Mot*. 1990;76(7):473-479.
57. Inoue M, Shino K, Hirose H, Horibe S, Ono K. Subluxation of the patella. Computed tomography analysis of patellofemoral congruence. *J Bone Joint Surg Am*. 1988;70(9):1331-1337.
58. Malghem J, Maldague B. Subluxation of the patella. Computed tomography analysis of patellofemoral congruence. *J Bone Joint Surg Am*. 1989;71(10):1575-1576.
59. Biedert RM, Gruhl C. Axial computed tomography of the patellofemoral joint with and without quadriceps contraction. *Arch Orthop Trauma Surg*. 1997;116(1-2):77-82.
60. Tardieu C, Glard Y, Garron E, Boulay C, Jouve J-L, Dutour O, et al. Relationship between formation of the femoral bicondylar angle and trochlear shape: independence of diaphyseal and epiphyseal growth. *Am. J. Phys. Anthropol*. 2006;130(4):491-500.
61. Nietosvaara Y. The femoral sulcus in children. An ultrasonographic study. *J Bone Joint Surg Br*. 1994;76(5):807-809.
62. Nietosvaara Y, Aalto K. The cartilaginous femoral sulcus in children with patellar dislocation: an ultrasonographic study. *J Pediatr Orthop*. 1997;17(1):50-53.
63. Martino F, De Serio A, Macarini L, Rizzo A, Laforgia R, Rotondo A, et al. Ultrasonography versus computed tomography in evaluation of the femoral trochlear groove morphology: a pilot study on healthy, young volunteers. *Eur Radiol*. 1998;8(2):244-247.

64. Modest VE, Murphy MC, Mann RW. Optical verification of a technique for in situ ultrasonic measurement of articular cartilage thickness. *J Biomech.* 1989;22(2):171-176.
65. Shavelson RJ, Webb NM, Rowley GL. Generalizability theory. *American Psychologist.* 1989;44(6):922-932.
66. Anderson H, Toms AP, Cahir JG, Goodwin RW, Wimhurst J, Nolan JF. Grading the severity of soft tissue changes associated with metal-on-metal hip replacements: reliability of an MR grading system. *Skeletal Radiol.* 2011;40(3):303–7.
67. Padhani AR, Ollivier L. The RECIST criteria: implications for diagnostic radiologists. *British Journal of Radiology.* 2001;74(887):983.
68. Toms AP, Kasmai B, Williams S, Wilson P. Building an anonymized catalogued radiology museum in PACS: a feasibility study. *Br J Radiol.* 2006;79:666-671.
69. RANDOM.ORG - True Random Number Service [Internet]. [cited 2011 Feb 1]; Available from: <http://www.random.org/>
70. Mistry A, Cahir J, Donell ST, Nolan J, Toms AP. MRI of asymptomatic patients with metal-on-metal and polyethylene-on-metal total hip arthroplasties. *Clin Radiol.* 2011;66(6):540-545.
71. Toms AP, Smith-Bateman C, Malcolm PN, Cahir J, Graves M. Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses. *Clin Radiol.* 2010;65(6):447-452.
72. Kingston A-R, Toms A, Ghosh-Ray S, Johnston-Downing S. Does running cause metatarsophalangeal joint effusions? A comparison of synovial fluid volumes on MRI in athletes before and after running. *Skeletal Radiology.* 2009;38(5):499-504.
73. Patel SR, Toms AP, Rehman JM, Wimhurst J. A reliability study of measurement tools available on standard picture archiving and communication system workstations for the evaluation of hip radiographs following arthroplasty. *J Bone Joint Surg Am.* 2011;93(18):1712–9.
74. Clohisy JC, Carlisle JC, Trousdale R, Kim Y-J, Beaule PE, Morgan P, et al. Radiographic evaluation of the hip has limited reliability. *Clin. Orthop. Relat. Res.* 2009;467(3):666-675.
75. Li WW. Tumor angiogenesis: molecular pathology, therapeutic targeting, and imaging. *Acad Radiol.* 2000;7(10):800-811.
76. Tofts PS. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging.* 1997;7(1):91-101.

77. Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging*. 1999;10(3):223-232.
78. van der Woude HJ, Verstraete KL, Hogendoorn PC, Taminiau AH, Hermans J, Bloem JL. Musculoskeletal tumors: does fast dynamic contrast-enhanced subtraction MR imaging contribute to the characterization? *Radiology*. 1998;208(3):821-828.
79. Verstraete KL, De Deene Y, Roels H, Dierick A, Uyttendaele D, Kunnen M. Benign and malignant musculoskeletal lesions: dynamic contrast-enhanced MR imaging--parametric "first-pass" images depict tissue vascularization and perfusion. *Radiology*. 1994;192(3):835.
80. de Baere T, Vanel D, Shapeero LG, Charpentier A, Terrier P, di Paola M. Osteosarcoma after chemotherapy: evaluation with contrast material-enhanced subtraction MR imaging. *Radiology*. 1992;185(2):587-592.
81. Dyke JP, Panicek DM, Healey JH, Meyers PA, Huvos AG, Schwartz LH, et al. Osteogenic and Ewing sarcomas: estimation of necrotic fraction during induction chemotherapy with dynamic contrast-enhanced MR imaging. *Radiology*. 2003;228(1):271-278.
82. Egmont-Petersen M, Hogendoorn PC, van der Geest RJ, Vrooman HA, van der Woude H, Janssen JP, et al. Detection of areas with viable remnant tumor in postchemotherapy patients with Ewing's sarcoma by dynamic contrast-enhanced MRI using pharmacokinetic modeling. *Magn Reson Imaging*. 2000;18(5):525-535.
83. Hanna SL, Parham DM, Fairclough DL, Meyer WH, Le AH, Fletcher BD. Assessment of osteosarcoma response to preoperative chemotherapy using dynamic FLASH gadolinium-DTPA-enhanced magnetic resonance mapping. *Invest Radiol*. 1992;27(5):367-373.
84. Reddick WE, Bhargava R, Taylor JS, Meyer WH, Fletcher BD. Dynamic contrast-enhanced MR imaging evaluation of osteosarcoma response to neoadjuvant chemotherapy. *J Magn Reson Imaging*. 1995;5(6):689-694.
85. van der Woude HJ, Bloem JL, Verstraete KL, Taminiau AH, Nooy MA, Hogendoorn PC. Osteosarcoma and Ewing's sarcoma after neoadjuvant chemotherapy: value of dynamic MR imaging in detecting viable tumor before surgery. *AJR Am J Roentgenol*. 1995;165(3):593-598.
86. Toms AP, White LM, Kandel R, Bleakney RR, Noseworthy M, Lee S, et al. Limitations of single slice dynamic contrast enhanced MR in

- pharmacokinetic modeling of bone sarcomas. *Acta Radiol.* 2009;50(5):512-520.
87. Rosen G, Caparros B, Huvos AG, Kosloff C, Nirenberg A, Cacavio A, et al. Preoperative chemotherapy for osteogenic sarcoma: selection of postoperative adjuvant chemotherapy based on the response of the primary tumor to preoperative chemotherapy. *Cancer.* 1982;49(6):1221–30.
  88. Erlemann R, Sciuk J, Bosse A, Ritter J, Kusnierz-Glaz CR, Peters PE, et al. Response of osteosarcoma and Ewing sarcoma to preoperative chemotherapy: assessment with dynamic and static MR imaging and skeletal scintigraphy. *Radiology.* 1990;175(3):791–6.
  89. Fletcher BD, Hanna SL, Fairclough DL, Gronemeyer SA. Pediatric musculoskeletal tumors: use of dynamic, contrast-enhanced MR imaging to monitor response to chemotherapy. *Radiology.* 1992;184(1):243–8.
  90. Kawai A, Sugihara S, Kunisada T, Uchida Y, Inoue H. Imaging assessment of the response of bone tumors to preoperative chemotherapy. *Clin Orthop Relat Res.* 1997;(337):216–25.
  91. Uhl M, Saueressig U, van Buijen M, Kontny U, Niemeyer C, Köhler G, et al. Osteosarcoma: preliminary results of in vivo assessment of tumor necrosis after chemotherapy with diffusion- and perfusion-weighted magnetic resonance imaging. *Invest Radiol.* 2006;41(8):618–23.
  92. Wunder JS, Paulian G, Huvos AG, Heller G, Meyers PA, Healey JH. The histological response to chemotherapy as a predictor of the oncological outcome of operative treatment of Ewing sarcoma. *J Bone Joint Surg Am.* 1998;80(7):1020-1033.
  93. Meyers PA, Gorlick R, Heller G, Casper E, Lane J, Huvos AG, et al. Intensification of preoperative chemotherapy for osteogenic sarcoma: results of the Memorial Sloan-Kettering (T12) protocol. *J. Clin. Oncol.* 1998;16(7):2452-2458.
  94. Bonnerot V, Charpentier A, Frouin F, Kalifa C, Vanel D, Di Paola R. Factor analysis of dynamic magnetic resonance imaging in predicting the response of osteosarcoma to chemotherapy. *Invest Radiol.* 1992 Oct;27(10):847-855.
  95. Ongolo-Zogo P, Thiesse P, Sau J, Desuzinges C, Blay JY, Bonmartin A, et al. Assessment of osteosarcoma response to neoadjuvant chemotherapy: comparative usefulness of dynamic gadolinium-enhanced spin-echo magnetic resonance imaging and technetium-99m skeletal angioscintigraphy. *Eur Radiol.* 1999;9(5):907-914.

96. Reynolds LA, Tansey EM. Early development of total hip replacement. London: The Wellcome Trust Centre for the History of Medicine; 2007.
97. Shellock FG, Spinazzi A. MRI safety update 2008: part 1, MRI contrast agents and nephrogenic systemic fibrosis. *AJR Am J Roentgenol*. 2008;191(4):1129–39.
98. August AC, Aldam CH, Pynsent PB. The McKee-Farrar hip arthroplasty. A long-term study. *Journal of Bone & Joint Surgery, British Volume*. 1986;68(4):520.
99. McKee GK, Watson-Farrar J. Replacement of arthritic hips by the McKee-Farrar prosthesis. *Journal of Bone & Joint Surgery, British Volume*. 1966;48(2):245.
100. Charnley J. Arthroplasty of the hip. A new operation. *Lancet*. 1961;1(7187):1129-1132.
101. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br*. 1972;54(1):61-76.
102. Cahir JG, Toms AP. (iii) CT and MRI of hip replacements. *Orthopaedics and Trauma*. 2009;23(2):101-108.
103. Toms AP, Botchu R, Nolan JF. (ii) Diagnostic plain film radiology of the failing hip replacement. *Orthopaedics and Trauma*. 2009;23(2):88-100.
104. Charnley J. Low friction arthroplasty of the hip: theory and practice. Springer; 1979.
105. Amstutz HC, Grigoris P. Metal on metal bearings in hip arthroplasty. *Clin. Orthop. Relat. Res*. 1996;(329 Suppl):S11-34.
106. Kothari M, Bartel DL, Booker JF. Surface geometry of retrieved McKee-Farrar total hip replacements. *Clin Orthop Relat Res*. 1996;(329 Suppl):S141-147.
107. Sauvé P, Mountney J, Khan T, De Beer J, Higgins B, Grover M. Metal ion levels after metal-on-metal Ring total hip replacement: a 30-year follow-up study. *J Bone Joint Surg Br*. 2007;89(5):586-590.
108. Brodner W, Bitzan P, Meisinger V, Kaider A, Gottsauner-Wolf F, Kotz R. Serum cobalt levels after metal-on-metal total hip arthroplasty. *J Bone Joint Surg Am*. 2003;85-A(11):2168-2173.
109. Ladon D, Doherty A, Newson R, Turner J, Bhamra M, Case CP. Changes in metal levels and chromosome aberrations in the peripheral blood of patients after metal-on-metal hip arthroplasty. *J Arthroplasty*. 2004;19(8 Suppl 3):78-83.



110. Valentí JR, Del Río J, Amillo S. Catastrophic wear in a metal-on-ceramic total hip arthroplasty. *J Arthroplasty*. 2007;22(6):920-922.
111. Gruber FW, Böck A, Trattinig S, Lintner F, Ritschl P. Cystic lesion of the groin due to metallosis: a rare long-term complication of metal-on-metal total hip arthroplasty. *J Arthroplasty*. 2007;22(6):923-927.
112. Bauer TW, Schils J. The pathology of total joint arthroplasty.II. Mechanisms of implant failure. *Skeletal Radiol*. 1999;28(9):483-497.
113. Saito S, Ryu J, Watanabe M, Ishii T, Saigo K. Midterm results of Metasul metal-on-metal total hip arthroplasty. *J Arthroplasty*. 2006;21(8):1105-1110.
114. Müller ME. The benefits of metal-on-metal total hip replacements. *Clin. Orthop. Relat. Res*. 1995;(311):54-59.
115. Daniel J, Pynsent PB, McMinn DJW. Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *J Bone Joint Surg Br*. 2004;86(2):177-184.
116. Massin P, Schmidt L, Engh CA. Evaluation of cementless acetabular component migration. An experimental study. *J Arthroplasty*. 1989;4(3):245-251.
117. Ilchmann T. Radiographic assessment of cup migration and wear after hip replacement. *Acta Orthop Scand Suppl*. 1997;276:1-26.
118. Barrack RL, Mulroy RD, Harris WH. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. *J Bone Joint Surg Br*. 1992;74(3):385-389.
119. Wan Z, Malik A, Jaramaz B, Chao L, Dorr LD. Imaging and navigation measurement of acetabular component position in THA. *Clin. Orthop. Relat. Res*. 2009;467(1):32-42.
120. Yoon Y-S, Hodgson AJ, Tonetti J, Masri BA, Duncan CP. Resolving inconsistencies in defining the target orientation for the acetabular cup angles in total hip arthroplasty. *Clin Biomech (Bristol, Avon)*. 2008;23(3):253-259.
121. Murray DW. The definition and measurement of acetabular orientation. *J Bone Joint Surg Br*. 1993;75(2):228-232.
122. Little NJ, Busch CA, Gallagher JA, Rorabeck CH, Bourne RB. Acetabular polyethylene wear and acetabular inclination and femoral offset. *Clin Orthop Relat Res*. 2009;467(11):2895-2900.

123. Su EP, Callander PW, Salvati EA. The bubble sign: a new radiographic sign in total hip arthroplasty. *J Arthroplasty*. 2003;18(1):110-112.
124. Aspenberg P, van der Vis H. Fluid pressure may cause periprosthetic osteolysis. Particles are not the only thing. *Acta Orthop Scand*. 1998;69(1):1-4.
125. Claus AM, Sychterz CJ, Hopper RH, Engh CA. Pattern of osteolysis around two different cementless metal-backed cups: retrospective, radiographic analysis at minimum 10-year follow-up. *J Arthroplasty*. 2001;16(8 Suppl 1):177-182.
126. Seelen JL, Bruijn JD, Hansen BE, Kingma LM, Bloem JL. Reproducible radiographs of acetabular prostheses. A method assessed in 35 patients. *Acta Orthop Scand*. 1994;65(3):258-262.
127. O'Neill DA, Harris WH. Failed total hip replacement: assessment by plain radiographs, arthrograms, and aspiration of the hip joint. *J Bone Joint Surg Am*. 1984;66(4):540-546.
128. Mjöberg B, Brismar J, Hansson LI, Pettersson H, Selvik G, Onnerfält R. Definition of endoprosthetic loosening. Comparison of arthrography, scintigraphy and roentgen stereophotogrammetry in prosthetic hips. *Acta Orthop Scand*. 1985;56(6):469-473.
129. Claus AM, Engh CA, Sychterz CJ, Xenos JS, Orishimo KF, Engh CA. Radiographic definition of pelvic osteolysis following total hip arthroplasty. *J Bone Joint Surg Am*. 2003;85-A(8):1519-1526.
130. Chiang PP, Burke DW, Freiberg AA, Rubash HE. Osteolysis of the pelvis: evaluation and treatment. *Clin Orthop Relat Res*. 2003;(417):164-174.
131. Engh CA, Sychterz CJ, Young AM, Pollock DC, Toomey SD, Engh CA. Interobserver and intraobserver variability in radiographic assessment of osteolysis. *J Arthroplasty*. 2002;17(6):752-759.
132. Tannast M, Langlotz U, Siebenrock K-A, Wiese M, Bernsmann K, Langlotz F. Anatomic referencing of cup orientation in total hip arthroplasty. *Clin. Orthop. Relat. Res*. 2005;(436):144-150.
133. Beckmann E. CT scanning the early days. *Br J Radiol*. 2006;79(937):5-8.
134. Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. *Radiographics*. 2004;24(6):1679-1691.
135. Lee M-J, Kim S, Lee S-A, Song H-T, Huh Y-M, Kim D-H, et al. Overcoming artifacts from metallic orthopedic implants at high-field-

- strength MR imaging and multi-detector CT. *Radiographics*. 2007;27(3):791-803.
136. Haramati N, Staron RB, Mazel-Sperling K, Freeman K, Nickoloff EL, Barax C, et al. CT scans through metal scanning technique versus hardware composition. *Comput Med Imaging Graph*. 1994;18(6):429-434.
  137. Moon SG, Hong SH, Choi J-Y, Jun WS, Kang H-G, Kim H-S, et al. Metal artifact reduction by the alteration of technical factors in multidetector computed tomography: a 3-dimensional quantitative assessment. *J Comput Assist Tomogr*. 2008;32(4):630-633.
  138. Schindera ST, Nelson RC, Yoshizumi T, Toncheva G, Nguyen G, DeLong DM, et al. Effect of automatic tube current modulation on radiation dose and image quality for low tube voltage multidetector row CT angiography: phantom study. *Acad Radiol*. 2009;16(8):997-1002.
  139. Kalender WA, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. *Radiology*. 1987;164(2):576-577.
  140. Wang G, Frei T, Vannier MW. Fast iterative algorithm for metal artifact reduction in X-ray CT. *Acad Radiol*. 2000;7(8):607-614.
  141. Link TM, Berning W, Scherf S, Joosten U, Joist A, Engelke K, et al. CT of metal implants: reduction of artifacts using an extended CT scale technique. *J Comput Assist Tomogr*. 2000;24(1):165-172.
  142. Lauterbur PC. Magnetic resonance zeugmatography. *Pure Appl Chem*. 1974;40(1-2):149-157.
  143. Mansfield P, Maudsley AA. Medical imaging by NMR. *Br J Radiol*. 1977;50(591):188-194.
  144. White LM, Kim JK, Mehta M, Merchant N, Schweitzer ME, Morrison WB, et al. Complications of total hip arthroplasty: MR imaging-initial experience. *Radiology*. 2000;215(1):254-262.
  145. Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donell ST, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol*. 2008;63(1):49-58.
  146. Pfirrmann CWA, Notzli HP, Dora C, Hodler J, Zanetti M. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology*. 2005;235(3):969-976.

147. Olsen RV, Munk PL, Lee MJ, Janzen DL, MacKay AL, Xiang QS, et al. Metal artifact reduction sequence: early clinical applications. *Radiographics*. 2000;20(3):699-712.
148. Cahir JG, Toms AP, Marshall TJ, Wimbhurst J, Nolan J. CT and MRI of hip arthroplasty. *Clin Radiol*. 2007;62(12):1163-1171; discussion 1172-1173.
149. Chang SD, Lee MJ, Munk PL, Janzen DL, MacKay A, Xiang QS. MRI of spinal hardware: comparison of conventional T1-weighted sequence with a new metal artifact reduction sequence. *Skeletal Radiol*. 2001;30(4):213-218.
150. Potter HG, Nestor BJ, Sofka CM, Ho ST, Peters LE, Salvati EA. Magnetic resonance imaging after total hip arthroplasty: evaluation of periprosthetic soft tissue. *J Bone Joint Surg Am*. 2004;86-A(9):1947-1954.
151. Eustace S, Goldberg R, Williamson D, Melhem ER, Oladipo O, Yucel EK, et al. MR imaging of soft tissues adjacent to orthopaedic hardware: techniques to minimize susceptibility artefact. *Clin Radiol*. 1997;52(8):589-594.
152. Eustace S, Jara H, Goldberg R, Fenlon H, Mason M, Melhem ER, et al. A comparison of conventional spin-echo and turbo spin-echo imaging of soft tissues adjacent to orthopedic hardware. *AJR Am J Roentgenol*. 1998;170(2):455-458.
153. Suh JS, Jeong EK, Shin KH, Cho JH, Na JB, Kim DH, et al. Minimizing artifacts caused by metallic implants at MR imaging: experimental and clinical studies. *AJR Am J Roentgenol*. 1998;171(5):1207-1213.
154. Toms AP, Smith-Bateman C, Malcolm PN, Cahir J, Graves M. Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses. *Clin Radiol*. 2010;65(6):447-452.
155. Lee MJ, Janzen DL, Munk PL, MacKay A, Xiang QS, McGowen A. Quantitative assessment of an MR technique for reducing metal artifact: application to spin-echo imaging in a phantom. *Skeletal Radiol*. 2001;30(7):398-401.
156. Toms AP, Smith-Bateman C, Malcolm PN, Cahir J, Graves M. Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses. *Clin Radiol*. 2010;65(6):447-452.
157. Lewis M, Toms AP, Reid K, Bugg W. CT metal artefact reduction of total knee prostheses using angled gantry multiplanar reformation. *Knee*. 2010;17(4):279-82.

158. Willert H-G, Buchhorn GH, Fayyazi A, Flury R, Windler M, Köster G, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am.* 2005;87(1):28-36.
159. Davies AP, Willert HG, Campbell PA, Learmonth ID, Case CP. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am.* 2005;87(1):18-27.
160. Jones DA, Lucas HK, O'Driscoll M, Price CH, Wibberley B. Cobalt toxicity after McKee hip arthroplasty. *J Bone Joint Surg Br.* 1975;57(3):289-96.
161. Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donell ST, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol.* 2008;63(1):49-58.
162. Boardman DR, Middleton FR, Kavanagh TG. A benign psoas mass following metal-on-metal resurfacing of the hip. *J Bone Joint Surg Br.* 2006;88(3):402-404.
163. Hart AJ, Sabah S, Henckel J, Lewis A, Cobb J, Sampson B, et al. The painful metal-on-metal hip resurfacing. *J Bone Joint Surg Br.* 2009;91(6):738-744.
164. Park Y-S, Moon Y-W, Lim S-J, Yang J-M, Ahn G, Choi Y-L. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am.* 2005;87(7):1515-1521.
165. Donell ST, Darrah C, Nolan JF, Wimhurst J, Toms A, Barker THW, et al. Early failure of the Ultima metal-on-metal total hip replacement in the presence of normal plain radiographs. *J Bone Joint Surg Br.* 2010;92(11):1501-8.
166. Böhler M, Kanz F, Schwarz B, Steffan I, Walter A, Plenk H, et al. Adverse tissue reactions to wear particles from Co-alloy articulations, increased by alumina-blasting particle contamination from cementless Ti-based total hip implants. A report of seven revisions with early failure. *J Bone Joint Surg Br.* 2002;84(1):128-136.
167. Khan RJK, Wimhurst J, Foroughi S, Toms A. The natural history of metallosis from catastrophic failure of a polyethylene liner in a total hip. *The Journal of Arthroplasty.* 2009;24(7):1144.
168. Toms AP, Nolan J, Barker T, Darrah C, Malcolm P. Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR. *Br J Radiol.* 2009;82(977):e87-91.

169. Medicines and Healthcare products Regulatory Agency (MHRA) [www.mhra.gov.uk](http://www.mhra.gov.uk/Publications/Safetywarnings/MedicalDeviceAlerts/CON079157). Medical Device Alert: All metal-on-metal (MoM) hip replacements (MDA/2010/033) [Internet]. [cited 2010 Oct 5]; Available from: <http://www.mhra.gov.uk/Publications/Safetywarnings/MedicalDeviceAlerts/CON079157>
170. Wynn-Jones H, Macnair R, Wimhurst J, Chirodian N, Derbyshire B, Toms A, et al. Silent soft tissue pathology is common with a modern metal-on-metal hip arthroplasty. *Acta Orthop*. 2011;82(3):301–7.
171. Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons CLM, et al. Pseudotumours associated with metal-on-metal hip resurfacings. *J Bone Joint Surg Br*. 2008;90(7):847-851.
172. Fang CSJ, Harvie P, Gibbons CLMH, Whitwell D, Athanasou NA, Ostlere S. The imaging spectrum of peri-articular inflammatory masses following metal-on-metal hip resurfacing. *Skeletal Radiol*. 2008;37(8):715-722.
173. Grammatopolous G, Pandit H, Kwon Y-M, Gundle R, McLardy-Smith P, Beard DJ, et al. Hip resurfacings revised for inflammatory pseudotumour have a poor outcome. *J Bone Joint Surg Br*. 2009;91(8):1019-1024.

## Index

### A

Absolute agreement.....	37
Acetabular anatomy .....	26
Acetabular cup height .....	97
Acetabular inclination .....	27, 29, 31, 32, 58, 60, 61, 64, 97
ALVAL .....	46, 47, 94, 96, 98, 100, 103, 104, 106
Angiogenesis.....	66, 75
Artefact.....	88
beam hardening .....	84
metal.....	53, 86, 90, 92
metal artefact reduction .....	91, 96
susceptibility .....	86
Articular cartilage .....	44
Asymptomatic.....	99, 104
Atrophy.....	97, 106

### B

Birmingham resurfacing .....	100
Bland-Altman plot.....	39, 63, 64
Bone marrow oedema.....	98, 105, 106

### C

Centre of rotation .....	58, 59, 61
Centre-edge-angle .....	27, 29, 31, 64
Charnley.....	79, 80
Chemotherapy.....	69, 71, 75
Classical Test Theory.....	36
Clearance Model .....	68
Cobalt-chromium.....	80, 95, 96

Coefficient	
Cohen's Kappa .....	38
Cohen's Kappa .....	51, 62
correlation .....	54
generalizability .....	39, 42
intra-class .....	42, 54, 55, 60, 61, 64
Pearson product-moment correlation .....	63
repeatability .....	39
Weighted Kappa .....	38, 50, 51
Computed tomography .....	83, 88
Consistency .....	37
Contingency tables .....	50
Correlation	
interclass .....	37
intraclass .....	37
linear .....	30, 60
Pearson product-moment .....	37
CT radiograph .....	28
D	
DCE-MRI .....	67, 69, 71, 77
Developmental dysplasia of the hip .....	26
DICOM .....	17, 34, 47, 110
E	
Endplate angle .....	20, 21, 23, 25, 64
Error .....	36
standard error .....	24
Ewing's sarcoma .....	74
F	
Facets .....	42, 43
Femoral sulcus .....	39
Femoroacetabular impingement .....	26



Fracture	
osteoporotic .....	19, 25
Frequency encoding gradient .....	87, 94
 G	
Gadolinium .....	66
Gantry .....	89
Generalizability theory.....	39
Gluteus medius .....	97, 101, 105, 106
Gluteus minimus .....	97, 101
Grade of severity.....	104, 106
Grading systems .....	46
Gyromagnetic ratio.....	88
 I	
Inter-tear drop line.....	58
 K	
$K^{\text{trans}}$ .....	67, 68
 L	
Lateral offset .....	58, 59, 61, 97
Leg length .....	58, 61, 97
Limits of agreement.....	39, 63, 64
Lymphoreticular.....	100, 102, 103
 M	
Magnetic resonance imaging .....	86, 91
Matrix .....	91, 94
McKee .....	79
McKee-Farrar.....	80, 81
Metallosis .....	80, 82
metal-on-metal .....	46, 94, 99, 100, 103
Metatarsophalangeal joints .....	55

MHRA.....	106
Multiplanar reconstruction .....	85, 88, 91
Muscle oedema.....	106
Myositis .....	97
N	
Necrosis .....	70, 71, 75, 94, 95, 98
O	
Osteoarthrosis.....	2, 26, 81, 105
Osteolysis.....	80, 82
Osteosarcoma.....	70, 74
P	
Patella instability .....	39
Pelvic tilt.....	30
Perfusion .....	70, 74
Phantom study .....	53, 88
Pharmacokinetic modelling .....	66
Photopaenia .....	85
Polyethylene	
debris.....	80, 82
R	
Range	
normal.....	25, 27
reference .....	21, 26, 28, 31, 33, 34
Receiver bandwidth.....	53, 88, 91, 94
Reliability.....	35, 36, 52, 57, 58, 62, 99
inter-rater .....	37, 42, 83, 108
intra-rater .....	37, 42
Reproducibility.....	35
ROI.....	51, 53, 54, 55, 70, 71, 74, 89

## S

Sample size.....	74
Segmentation .....	51, 52, 53, 55
Selection bias.....	25, 74
Short external rotators.....	97, 105
Signal to noise ratio.....	88, 92
Signal void.....	91
Sinus histiocytosis.....	102
Slipped upper femoral epiphysis .....	26
Stem angle .....	58, 59, 61, 97
Sulcus angle.....	39, 42, 45
Synovial fluid .....	55

## T

Tendon avulsion.....	97, 103
Total hip replacement.....	53, 54, 79, 94, 99
Metal-on-metal.....	79, 80
polyethylene-on-metal .....	79
Total knee replacement .....	89
Trochlear dysplasia .....	39, 40

## U

Upper limit reference.....	25
----------------------------	----

## V

## Vertebral

angulation .....	20
height.....	19
wedging .....	19
Voxel .....	53, 88, 91

## U

$u_e$ .....	67, 68
-------------	--------

# Defining Normal Vertebral Angulation at the Thoracolumbar Junction

Michael B. Crawford<sup>1</sup>  
Andoni P. Toms<sup>1</sup>  
Lee Shepstone<sup>2</sup>

**OBJECTIVE.** The purpose of this cross-sectional study is to define the normal range of endplate angulation at T12 and L1 and, by doing so, to validate the angle measurement tools that are readily available on nearly all PACS.

**MATERIALS AND METHODS.** Two hundred consecutive lateral scout CT scans were examined in patients who were either 25 ( $n = 100$ ) or 35 ( $n = 100$ ) years old. The endplate angles for T12 and L1 were measured using a “Cobb angle” tool on a standard PACS workstation. Twenty-two cadaveric vertebrae were also imaged, and measurements obtained from the lateral scout CT image using electronic calipers were compared with measurements obtained with a goniometer.

**RESULTS.** The mean endplate angle at T12 measures  $4.34^\circ$  (2 SD,  $4.5^\circ$ ) and at L1,  $4.48^\circ$  ( $4.26^\circ$ ). The normal range of endplate angulation is therefore  $-0.16^\circ$  to  $8.84^\circ$  at T12 and  $0.22^\circ$ – $8.74^\circ$  at L1. No statistically significant difference was seen in the endplate angulation when men were compared with women or 25- and 35-year-old age groups were compared. A strong correlation exists between direct and CT-derived endplate angle measurements.

**CONCLUSION.** Vertebral endplate angulation can be reliably measured using widely available PACS workstation tools. The mean endplate angle for T12 and L1 is approximately  $4.5^\circ$ , with an approximate range extending from  $0^\circ$  to  $9^\circ$ . For practical purposes, an endplate angle of  $10^\circ$  or more can be considered outside the normal range.

**Keywords:** anatomy, CT, thoracolumbar junction, wedging

DOI:10.2214/AJR.08.2026

Received October 29, 2008; accepted after revision January 2, 2009.

<sup>1</sup>Department of Radiology, Norfolk and Norwich University Hospital, NHS Foundation Trust, Colney Ln., Norwich, Norfolk NR4 7UY, United Kingdom. Address correspondence to M. B. Crawford (Michael.crawford@nnuh.nhs.uk).

<sup>2</sup>School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, Norfolk, United Kingdom.

## WEB

This is a Web exclusive article.

AJR 2009; 193:W33–W37

0361–803X/09/1931–W33

© American Roentgen Ray Society

**A**nterior angulation or “wedging” of the vertebral bodies at the thoracolumbar junction is a recognized normal anatomic feature [1]. The thoracolumbar junction is also the most common location for osteoporotic and traumatic vertebral fractures. More than 50% of traumatic fractures occur between T12 and L2 [1, 2]. The lower thoracic spine is also the most common site of Scheuermann’s disease and is a further cause of vertebral wedging [3, 4]. Therefore, it can sometimes be difficult to differentiate between grade 1 osteoporotic fractures, mild traumatic fractures, and normal anatomic wedging. Several techniques have been described to quantify vertebral deformity. These usually involve measuring and comparing the relative anterior and posterior heights of vertebral bodies [1, 5–7]. Although this quantification of vertebral shape provides objective measures that can assist in the interpretation of thoracolumbar radiographs, the technique is time-consuming and infrequently used. The purpose of this study was to define the normal range of angulation of the

endplates of the T12 and L1 vertebral bodies using electronic tools that are widely available on all diagnostic DICOM workstations. The validity of these tools to obtain measurements of vertebral angulation was also evaluated.

## Material and Methods

An analysis of lateral scout CT scans from 200 consecutive thoracoabdominal CT examinations (LightSpeed Plus, GE Healthcare) was performed. Two hundred consecutive patients, 100 each at 25 and 35 years old, were included in the study beginning January 1, 2002, and ending December 31, 2004. All CT examinations were harvested from the hospital PACS. Only the first CT scan of a given patient was included. A lateral tomograph showing the entire lumbosacral spine was required for inclusion in study. Exclusion criteria were as follows: major trauma (such as an motor vehicle collision); any disease of the vertebral column, spinal canal, paravertebral soft tissues, or retroperitoneum; traumatic vertebral fractures; and known vertebral collapse. Patients were also excluded if there were fewer or more than five lumbar vertebrae or segmentation anomalies at the

lumbosacral junction. In total, 222 CT studies were examined, with 22 being excluded, leaving a data set of 200. Eight were excluded because the lateral tomographs were too pixilated to clearly identify the outline of the vertebrae. Thirteen were excluded because of lumbosacral ambiguity. One study was excluded because of retroperitoneal disease.

### Measurements

Endplate angles for T12 and L1 were measured using electronic calipers from the lateral CT tomographs. The superior endplate was defined as a line drawn between the most anterosuperior and the most posterosuperior corners of the vertebral body. The inferior endplate angle was similarly defined for the anteroinferior and posteroinferior corners. The shape of the intervening endplate was ignored (Fig. 1). The angle between the two endplates was measured using the "Cobb angle" tool on the PACS workstation (Fig. 2).

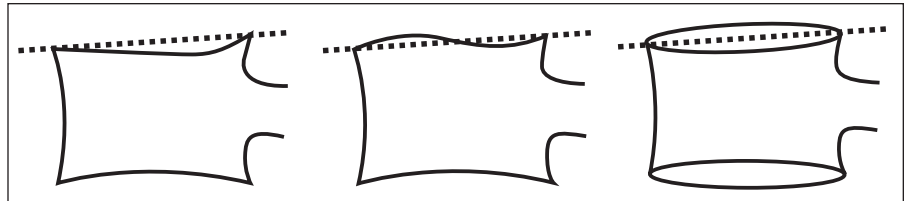
CT examinations for a total of 200 patients (114 men and 86 women) were included in the study and were independently examined by two observers. All CT examinations were obtained with the patient in the supine position. The data sets were then compared, and when there was a difference of more than  $3^\circ$ , the case was reviewed by the two reporting observers and remeasured independently in order to minimize discrepancies between the two observers that might result from inadvertently measuring the incorrect vertebral level. The second measurement was then accepted regardless of the difference between observers. A total of 15 examinations were reviewed for which the second readings were used in the final data set. Fifty examinations were measured twice by the same observer (blinded), with an interval of 8 weeks between measurements.

The rationale for using CT scout tomograms for assessing vertebral angulation was tested using 22 cadaveric vertebrae. A lateral scout tomogram was performed on the dry specimens using the standard CT protocol for chest and abdominal CT (Fig. 3). Thereafter, the vertebral endplate angles were measured independently by two observers. These measurements were obtained from the dry bones using a goniometer and from the CT images using the electronic Cobb angle tool.

### Statistical Analysis

For the cadaveric vertebral study, interobserver reliability was tested using intraclass correlation (ICC); a comparison of the dry bone and CT measurements was performed using the Bland-Altman plot and the Pearson's correlation coefficient test.

For the study of the healthy CT population, the mean endplate angle was calculated for T12 and



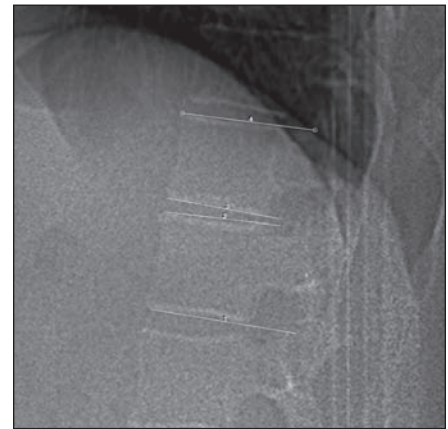
**Fig. 1**—Diagrammatic representation of measurements of variable endplate morphology on lateral projection. Lines intersecting anterior and posterior superior corners and anterior and posterior inferior corners were drawn without regard for shape or orientation of intervening endplate.

L1 with a normal range defined as 2 SD from the mean. Intra- and interobserver reliability was tested using ICCs. Comparisons between T12 and L1 were performed using the paired Student's *t* test, and between men and women and between patients 25 and 35 years old were performed using the unpaired Student's *t* test.

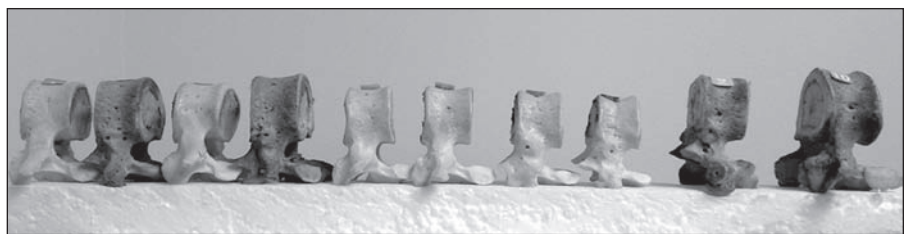
To assess whether the results are applicable to conventional radiography (in which the thoracolumbar junction is typically projected over the superior region of the radiographic plate and subjected to magnification and parallax distortion), a trigonometric model was used to estimate the maximal possible alteration in endplate angulation. A hypothetical vertebral body measuring 4.5 cm in the anteroposterior dimension, having a 3.5 cm posterior body height, and with an endplate angle of  $9^\circ$ —the upper limit of the normal range—was used for the calculations. The hypothetical vertebral body was considered to be positioned 20 cm from a  $42 \times 35$  radiographic film with a film-focus distance of 100 cm and projected over the superior edge of the film, which is the point of maximal parallax error (Fig. 4).

### Results

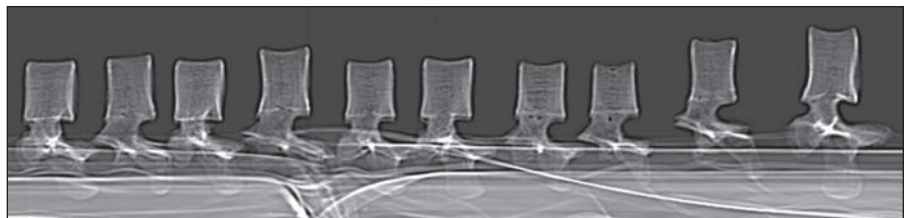
The mean endplate angle for T12 was  $4.34^\circ$  (Table 1). Therefore, the calculated normal statistical range is from  $-0.16^\circ$  to  $8.84^\circ$ . The mean endplate angle for L1 was  $4.48^\circ$ , with a normal statistical range of  $0.22$ – $8.74^\circ$  (Fig.



**Fig. 2**—Lateral scout radiograph shows use of "Cobb angle" tool on PACS workstation.



**A**

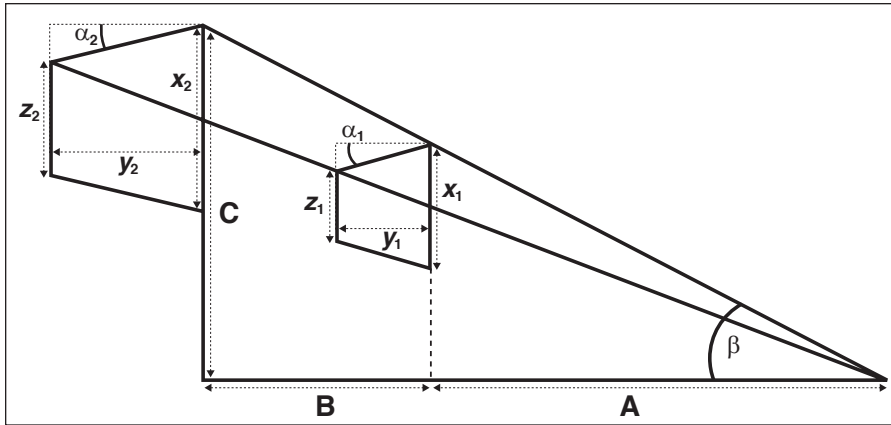


**B**

**Fig. 3**—Reliability study of CT scout tomography.

**A and B**, Lateral photograph (**A**) and scout CT scan (**B**) of part of set of cadaveric lumbar vertebrae used to assess reliability of CT scout tomograms for reproducing vertebral geometry.

## Vertebral Angulation at the Thoracolumbar Junction



**Fig. 4**—Diagrammatic representation of magnification and parallax distortion of thoracolumbar vertebra projected onto superior edge of 42 × 35 cm radiographic plate. With film–focus distance of 100 cm (A + B), beam at edge of film subtends angle of 11.9° to horizontal [ $\tan^{-1} (21 / 100)$ ]. A vertebral body with endplate angle of 9° ( $2\alpha_1$ ), measuring 3.5 cm in posterior body height ( $x_1$ ) and 4.5 cm in anteroposterior width ( $y_1$ ), positioned 20 cm (A) from film will be magnified and distorted because of parallax errors. From this the following can be estimated using simple trigonometry. Position of vertebra from central beam (C) equals 16.8 cm, anterior body height ( $z_1$ ) equals 3.15 cm, and superior and inferior endplate angles ( $\alpha_1$ ) measure 4.5°. Posterior vertebral body height of projected image increases to 4.4 cm, anterior body height increases to 3.5 cm, superior and inferior endplate angles ( $\alpha_2$ ) decrease to 4.35° and 4.39°, respectively. Dimensions  $x_2$ ,  $y_2$ , and  $z_2$  refer to posterior vertebral body.  $\beta$  = angle from center to edge of x-ray beam.

**TABLE 1: Measurements of Endplate Angulation at T12 and L1, Including Subset Measurements for Age and Sex**

Group	T12 (°)	L1 (°)
All (n = 200)	4.34 (4.5)	4.48 (4.26)
Women (n = 86)	4.12 (4.34)	4.49 (4.24)
Men (n = 114)	4.50 (4.6)	4.70 (4.56)
Age 25 (n = 100)	4.60 (4.8)	4.44 (4.1)
Age 35 (n = 100)	4.12 (4.12)	4.50 (4.4)

Note—Numbers in parentheses are 2 SD.

**TABLE 2: Differences in Endplate Angulation at T12 and L1 for Age and Sex**

Comparison	Difference <sup>a</sup> (°)	$p^b$
T12 vs L1	0.137 (−0.293 to 0.568)	0.53
Sex difference at T12	0.387 (1.02 to 0.247)	0.23
Sex difference at L1	0.024 (−0.575 to 0.625)	0.94
Age difference at T12	0.445 (−0.182 to 1.072)	0.16
Age difference at L1	0.090 (−0.684 to 0.504)	0.77

<sup>a</sup>Mean (95% CIs).

<sup>b</sup>Independent samples Student's *t* test.

**TABLE 3: Trigonometric Estimate of Effect of Parallax Error During Conventional Radiography on Endplate Angulation of a Vertebra Projected Onto Edge of 42 x 35 cm Radiographic Plate**

Measurement	Original Vertebral Dimensions	Magnified Dimensions 21 cm From Center of Film (12° arc)
Anterior height (mm)	31.5	35
Posterior height (mm)	35	43.75
Anteroposterior dimension (mm)	45	≈57
Endplate angle ( $2\alpha^\circ$ )	9	8.7

Note— $2\alpha^\circ$  = combined superior and inferior endplate angles.

5). No statistical difference was seen between the two levels ( $p = 0.40$ ) (Table 2).

In women, the mean angle at T12 was 4.12° and at L1, 4.49°. In men, the mean endplate angle at T12 was 4.5° and at L1, 4.7° degrees (Figs. 6 and 7). No significant difference was seen between men and women at either T12 ( $p = 0.23$ ) or L1 ( $p = 0.94$ ) (Table 2).

In the 25-year-old patients, the mean vertebral endplate angle at T12 was 4.6° and at L1, 4.44° (Table 1). In the 35-year-old patients, the mean endplate angulation at T12 measured 4.12° and at L1, 4.5° (Figs. 8 and 9). Again, no significant difference was seen in the two age groups at either T12 ( $p = 0.16$ ) or L1 ( $p = 0.77$ ) (Table 2).

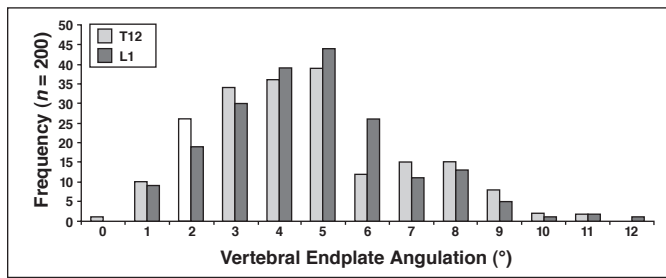
The ICC for intraobserver reliability, which was based on the first 50 CT examinations, was 0.800 (95% CI, 0.674–0.962) for T12 and 0.794 (0.665–0.923) for L1. The interobserver reliability for all 200 measurements was 0.858 (0.794–0.923) for T12 and 0.808 (0.723–0.893) for L1. The mean difference for the two observers was 1.2° at both T12 and L1 (2 SD, 1.8°).

The interobserver reliability for the direct (ICC, 0.972; 95% CI, 0.932–0.988) and CT-derived (0.963; 0.91–0.984) measures of the dry cadaveric vertebra were very good. The correlation between the mean observed measures showed a strong correlation between direct and CT-derived endplate angle measurements ( $r = 0.961$ ,  $p < 0.01$ ). The Bland-Altman plot shows that the mean CT measurement is 0.8° smaller than the dry bone measurement (95% CI, −3.2° to 1.6°).

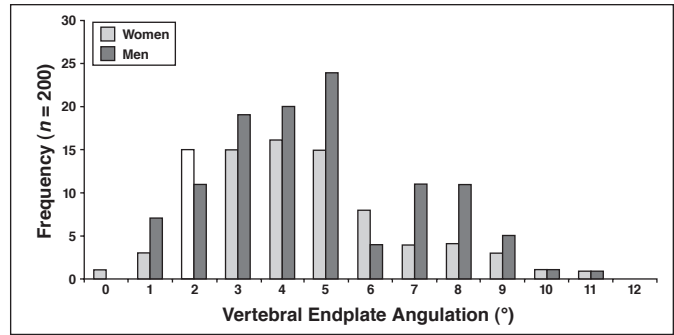
Trigonometric modeling of the alteration in endplate angulation on conventional radiography yielded the following results. The angle of the superior endplate is reduced by a maximum of 0.15°, and the inferior endplate angle is reduced by a maximum of 0.11° (Table 3).

## Discussion

The purpose of this study was to define the normal anatomic vertebral angulation that occurs at the thoracolumbar junction. CT lateral scout tomographs were used in this study for two reasons. The first is that the lateral scout CT images do not have any geometric distortion. Although a fan beam is used to obtain the image, the reconstruction algorithm corrects any anteroposterior magnification. The validity of this was confirmed using the dry bone phantom. The second reason is that for most patients in this age group, thoracolumbar CT was incidental to the primary reason for examining the patient. In contrast,



**Fig. 5**—Histogram shows distribution of endplate angle measurements at T12 and L1 in 200 consecutive patients.



**Fig. 6**—Histogram compares distribution of endplate angle measurements at T12 in men and women.

radiographs and MR images of the spine almost always are of symptomatic patients.

One may argue that the sample examined is not reflective of a normal population because all patients included had an indication for a CT examination. However, by selecting patients between the ages of 25 and 35 years, it is unlikely that any would be affected by osteoporosis. We believe that in the normal population of women, vertebral morphology changes little until the menopause [1]. The lower age limit of 25 was defined to ensure that bone growth was complete, and the incidence of osteoporosis is rare in the 35-year age group. Careful selection of the criteria for CT and exclusion of abnormal findings should also have excluded those with abnormal vertebral morphology due to disease.

Selecting 2 SD is a conventional statistical description of a normal range in a normal Gaussian distribution of data. However, it has been previously highlighted that this is an artificial definition, which will mean that a few healthy patients will fall outside this range [7].

A potential selection bias may have occurred because patients were excluded for poor quality of the lateral tomographs. This process

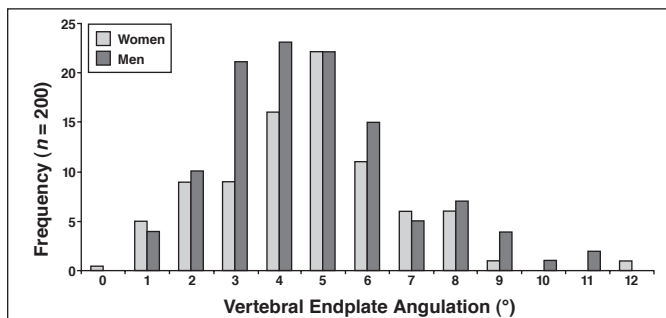
would tend to select out patients with larger body mass indexes but whose vertebral morphology may be considered entirely normal.

Despite these limitations, this sample is likely to be as close to normal as can be achieved in an analysis using clinical data. To perform normal population sample studies—that is, drawn directly from the community without any symptoms or known conditions—with radiography or MRI would not be financially or ethically viable.

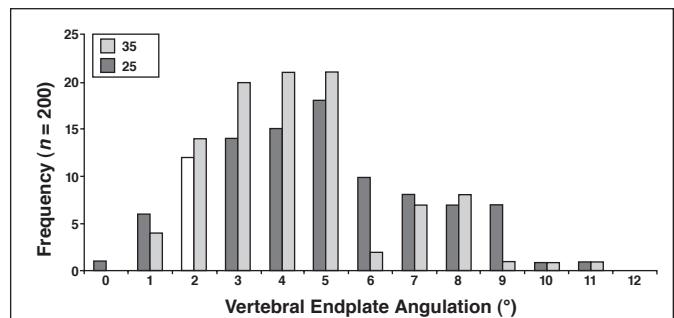
The range of values for angulation of the T12 and L1 vertebrae were derived from the mean of the two observers' measurements of 200 cases. The very good intraclass coefficient for the intra- and interobserver variability suggests that this is a reliable and reproducible set of measurements. The interobserver correlation may be slightly higher than in normal clinical practice because 15 observations were repeated, and because there was a greater than 3° difference between the observers. This repetition was performed to minimize any measurement errors attributable to mistakes in counting vertebral levels and transcribing measurements. These are part of variance between observers, but the repeat measurements were performed so that

most of the interobserver variability could be attributed to measurement errors only.

The results of this study differ from previous published data in several areas. Previous studies have mostly included women because of their greater propensity for osteoporosis [1, 8–10], and have included fewer patients [1, 8–12], and have used different methods to evaluate vertebral body morphology [1, 8, 9, 11, 12]. Most previous methods involved measuring the relative heights of the anterior, middle, and posterior aspects of vertebral bodies [1, 8, 9, 11, 12]. Panjabi et al. [13, 14] showed similar results for vertebral angulation at T12 ( $4.0^\circ \pm 1.11^\circ$ ) but a slightly higher degree of angulation at L1 ( $6.7^\circ \pm 1.61^\circ$ ). However, their results were derived from the dry bones of only 12 individuals (mean age, 46.3 years), some of whom had significant comorbidity at death (six patients had cancer). We would argue that our sample is a better and more normal population sample. The advantages of our approach are that we have included a larger number of patients, have included both men and woman, and by using CT we have reduced as much as possible the influence of geometric distortion. Angulation of endplates is not dependent on



**Fig. 7**—Histogram compares distribution of endplate angle measurements at L1 in men and women.

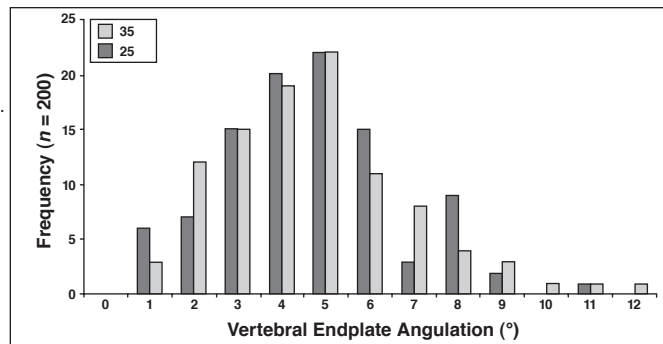


**Fig. 8**—Histogram compares distribution of endplate angle measurements at T12 in patients 25 and 35 years old.



## Vertebral Angulation at the Thoracolumbar Junction

**Fig. 9**—Histogram compares distribution of endplate angle measurements at L1 in patients 25 and 35 years old.



size of the vertebra, whereas absolute height measurements may vary with sex and height of individual.

The mean angle was slightly greater in men than women at T12 and the converse was true at L1. However, neither of these is statistically significant. With mean differences at T12 of 0.4° and at L1 of 0.02°, the normal ranges for both levels and sexes are therefore very similar. These findings differ from previously published data in which men appear to consistently have a greater degree of vertebral wedging than women [15]. This discrepancy may be explained by the relatively young age of our sample group. Although wedging in the women is constant before menopause, postmenopausal women have an increased incidence relative to men of a similar age. The similarity in the measurements of vertebral wedging in our two age groups supports a recently published cadaveric study, which suggested that vertebral morphology is constant with increasing age [16].

One question that arises after taking data from a source with limited geometric distortion is whether the values for normal range of angulation can be translated into conventional radiographic images of the lumbar spine. The trigonometric modeling of the parallax effect on endplate angulation suggests that the maximum alteration is an apparent reduction in endplate angle of 0.26°. This calculation suggests that, for practical purposes, the measures derived from the lateral CT radiograph are applicable to conventional radiography of the thoracolumbar junction. This finding is supported by a previous study that concluded that cephalograms and CT scano-

grams are comparable for depicting angular relations of structures [17].

The results of this study indicate that measuring vertebral endplate angulation at T12 and L1 using electronic calipers is a reliable and reproducible technique. In routine practice, this would be quick and easy to do if there was any concern about the amount of angulation of the thoracolumbar vertebra after visual inspection. The sum of the normal statistical range of endplate angulation at T12 and L1 (2 SD = 8.75°), estimated geometric distortion (0.26°), and the mean interobserver differences (1.2°) is 10.2°. We would suggest that 10° might be a useful rule of thumb for the maximal statistically "normal" endplate angulation measurement from a lateral radiograph, and that measurements outside the normal range may indicate a fracture or collapse even if cortical or trabecular disruption is not visible. The converse, however, is not true; an endplate angle of less than 10° does not exclude a fracture.

In conclusion, measurement of vertebral body endplate angulation using the method described in this article is a reliable and reproducible technique. Assuming that our study sample is a normal population, the range of normal angulation is approximately 0–9° for T12 and L1. For practical purposes, a vertebral angulation of 10° or more could be considered to be outside the normal range.

### References

- Gallagher JC, Hedlund LR, Stoner S, Meeger C. Vertebral morphometry: normative data. *Bone Miner* 1988; 4:189–196
- Holmes JF, Miller PQ, Panacek EA, Lin S, Home NS, Mower WR. Epidemiology of thoracolumbar

spine injury in blunt trauma. *Acad Emerg Med* 2001; 8:866–872

- Ali RM, Green DW, Patel TC. Scheuermann's kyphosis. *Curr Opin Pediatr* 1999; 11:70–75
- Tribus CB. Scheuermann's kyphosis in adolescents and adults: diagnosis and management. *J Am Acad Orthop Surg* 1998; 6:36–43
- Jensen KK, Tougaard L. A simple x-ray method for monitoring progress of osteoporosis. *Lancet* 1981; 2:19–20
- Black DM, Cummings SR, Stone K, Hudes E, Palermo L, Steiger P. A new approach to defining normal vertebral dimensions. *J Bone Miner Res* 1991; 6:883–892
- Leidig G, Minne HW, Sauer P, et al. A study of complaints and their relation to vertebral destruction in patients with osteoporosis. *Bone Miner* 1990; 8:217–229
- Davies KM, Recker RR, Heaney RP. Normal vertebral dimensions and normal variation in serial measurements of vertebrae. *J Bone Miner Res* 1989; 4:341–349
- McCloskey EV, Spector TD, Eyres KS, et al. The assessment of vertebral deformity: a method for use in population studies and clinical trials. *Osteoporos Int* 1993; 3:138–147
- Ryan PJ, Fogelman I. Osteoporotic vertebral fractures: diagnosis with radiography and bone scintigraphy. *Radiology* 1994; 190:669–672
- Dai L. Radiographic measurements of the heights of vertebral bodies in thoracic and lumbar spine. *Chin Med Sci J* 1996; 11:117–119
- Hurxthal LM. Measurement of anterior vertebral compressions and biconcave vertebrae. *Am J Roentgenol Radium Ther Nucl Med* 1968; 103: 635–644
- Panjabi MM, Goel V, Oxland T, et al. Human lumbar vertebrae. quantitative three-dimensional anatomy. *Spine* 1992; 17:299–306
- Panjabi MM, Takata K, Goel V, et al. Thoracic human vertebrae. quantitative three-dimensional anatomy. *Spine* 1991; 16:888–901
- Lauridsen KN, De Carvalho A, Andersen AH. Degree of vertebral wedging of the dorso-lumbar spine. *Acta Radiol Diagn (Stockh)* 1984; 25:29–32
- Masharawi Y, Salame K, Mirovsky Y, et al. Vertebral body shape variation in the thoracic and lumbar spine: characterization of its asymmetry and wedging. *Clin Anat* 2008; 21:46–54
- Chidiac JJ, Shofer FS, Al-Kutoubi A, Laster LL, Ghafari J. Comparison of CT scanograms and cephalometric radiographs in craniofacial imaging. *Orthod Craniofac Res* 2002; 5:104–113



# Defining a reference range of acetabular inclination and center-edge angle of the hip in asymptomatic individuals

Lucy A. Fowkes · Elia Petridou ·  
Christopher Zagorski · Amanjot Karuppih ·  
Andoni P. Toms

Received: 27 August 2010 / Revised: 26 December 2010 / Accepted: 17 January 2011 / Published online: 1 February 2011  
© ISS 2011

## Abstract

**Purpose** Acetabular morphology is an important predictor of the severity of osteoarthritis and survival of hip prostheses but there is limited data on the normal range of acetabular measurements on plain radiographs. The aim of this project was to determine the statistically normal ranges of acetabular inclination (AI) and center-edge angle (CEA).

**Method** One hundred coronal CT localizers (50 men and 50 women aged 20–30 years) were included in this study. All the patients underwent CT examination for thoracic or intra-abdominal indications. Patients with pelvic disease, fractures, history of serious trauma, or previous pelvic surgery were excluded. One pair of independent observers measured the AI and pelvic tilt (PT), and a further pair measured the center-edge angle (CEA), using electronic calipers on a high-resolution PACS workstation.

**Results** AI and CEA measurements were obtained for 200 hips. There was very good intra-class correlation between the observers ( $r=0.7$ – $0.8$ ). The mean AI was  $38.8^\circ$  (2SD  $32.1$ – $45.5^\circ$ ). That in men was  $38.0^\circ$  (2 SD  $31.8$ – $44.1^\circ$ ) and  $39.6^\circ$  (2 SD  $32.7$ – $46.8^\circ$ ) in women, which was statistically significantly different ( $p<0.001$ ). The mean CEA measurement for all patients was  $36.3^\circ$  (SD  $13.8^\circ$ ), for men  $37.7^\circ$  (SD  $10.8^\circ$ ) and for women  $34.9^\circ$  (SD  $11.4^\circ$ ) with a statistically significant gender difference ( $p<0.001$ ). The

mean pelvic tilt measurement (sacro-coccygeal-pubic symphysis) was  $38.3$  mm (2 SD  $18.3$ – $58.3$  mm) with a significant gender difference ( $p<0.001$ ).

**Conclusions** The results of this study define reference ranges of two common measures of acetabular morphology and confirm statistically significant differences between men and women.

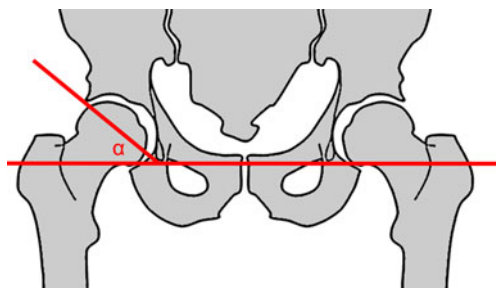
**Keywords** Acetabular inclination · Center-edge angle · Pelvic tilt

## Introduction

The etiology of hip osteoarthritis (OA) has been widely debated since the observation of premature OA in hips with previously undetected slipped upper femoral epiphysis (SUFE) [1]. It is generally accepted that abnormal acetabular development can lead to OA in patients in their third or fourth decade of life. Abnormal acetabular development can take a number of forms, which can be categorized as those that result in poor coverage of the femoral head or those that result in too much cover. Inadequate acetabular coverage of the femoral head, generally described as acetabular dysplasia or developmental dysplasia of the hip (DDH), most commonly leads to inadequate superolateral cover of the hip and is a common cause of hip arthroplasty in young patients who subsequently develop OA [2]. Acetabular overcoverage is one cause of femoroacetabular impingement (FAI) [3–5]. Subtypes of FAI can be differentiated [6, 7] and graded, and also used to prognosticate on the probability of developing early OA [8]. Despite the fact that abnormal acetabular anatomy may be an important predictor in the development of early OA, the normal range of acetabular orientation and morphology is imperfectly understood.

L. A. Fowkes · E. Petridou · C. Zagorski · A. Karuppih ·  
A. P. Toms  
Department of Radiology,  
Norfolk and Norwich University Hospital,  
Norwich, Norfolk NR4 7UY, UK

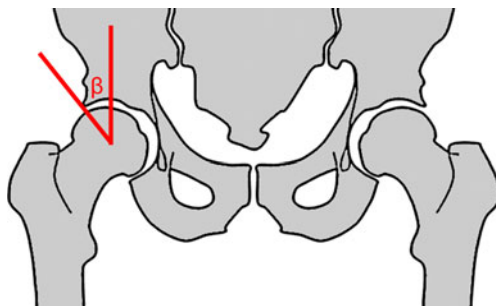
L. A. Fowkes (✉)  
Norwich Radiology Academy, Cotman Centre,  
Colney Lane,  
Norfolk NR4 7UB, UK  
e-mail: lucy.fowkes@nnuh.nhs.uk



**Fig. 1** Diagram demonstrating the measurement of acetabular inclination ( $\alpha$ ) from the inter-tear drop baseline

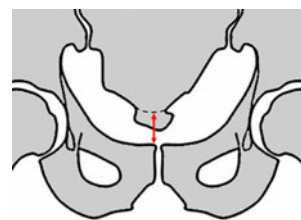
Until recently, conventional radiographs of the pelvis and hips have been the main tools for obtaining measures of acetabular orientation. Despite some evidence that computed tomography (CT) may be a more accurate tool because measurements can be obtained with reference to a true anatomical plane [9], there is also evidence that conventional antero-posterior (AP) radiographs are valid for deriving measures of acetabular orientation with the pelvis in a postural orientation [10]. This is probably more relevant to the biomechanics of the hip than the true anatomical plane. From the AP radiograph of the pelvis the acetabular angle or acetabular index (AI) [11] and the center-edge angle (CEA) [12] can be obtained as well as measures of acetabular orientation and cover in the sagittal plane. Acetabular version can be measured from a lateral projection of the hip [13].

The understanding of what constitutes a normal range of acetabular orientation values in disease-free hips is limited. The principle limiting factor is that it is not ethically acceptable to irradiate the pelvis of a large number of young healthy volunteers to obtain this data. To date, measurements have been obtained from cadaveric [14–16] and synthetic [10] pelvises. A “normal” range for AI was first described as varying from 33–38°, irrespective of gender, however this data was obtained from radiographs in patients over 60 years old. Confusingly, the same study considered that 39–42° represented the “upper limit of normal” [11]. Another study determined the normal range of AI to be 25–41° but it included patients up to the age of 85, the inclusion criteria for normality being entirely radiographic [17].



**Fig. 2** Diagram demonstrating the measurement of the centre-edge-angle ( $\beta$ )

**Fig. 3** Diagram demonstrating the measurement of the sacro-coccygeal to pubis distance (arrow) for assessing pelvic tilt

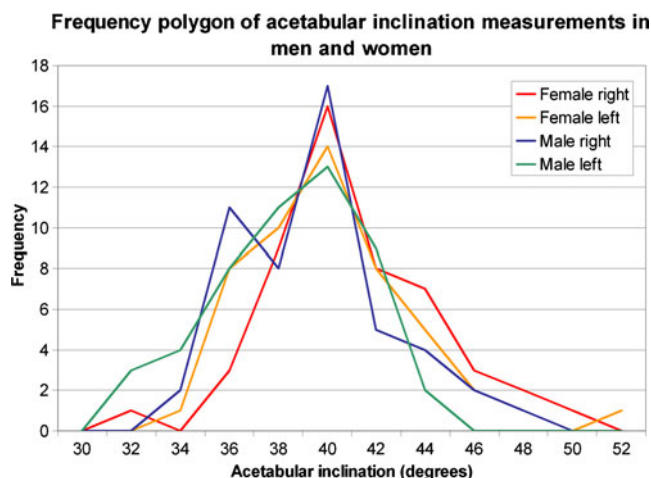


The centre-edge angle of Wiberg (CEA) is often used to express femoral head coverage [12] and values of 20–40° are deemed normal [18]. This range was derived from a combination of analysis of 300 radiographs of disease-free adult hips and radiography of a cadaveric pelvis. Unfortunately, the data was based upon a largely male (74%) population and 84% were over the age of 40 [18]. These measurements are also dependent upon pelvic position [3, 19] and this is unaccounted for in the ranges provided.

An AP projection of the pelvis is obtained during abdominopelvic CT planning in the form of a CT localizer. As it is incidental to the indication for the CT study, the tomogram can be used to determine a “normal” range of acetabular orientation and morphology. The aim of this study was to determine the statistically normal or reference ranges (i.e., the arithmetical mean with 2 standard deviations,  $\bar{x} \pm 2SD$ ) of AI and CEA derived from incidental pelvic CT radiography.

## Materials and methods

At our institution, local research ethics approval is not required for retrospective anatomical studies using PACS but the research is carried out under the research governance arrangements within the department of radiology that adhere to the principles of Good Clinical Practice in Research. One hundred coronal CT localizers (50 men and 50 women aged 20–30 years) were included in this



**Fig. 4** Frequency polygon of acetabular inclination

**Table 1** Descriptive statistics measurements comparing acetabular inclination (AI) and center-edge angle (CEA) in right and left hips in males and females\*Paired *t* test

Measure	Gender	Side	Mean (2 SD)	Difference	*Significance
Acetabular inclination	Male	Right	38.52 (6.14)	0.96	<i>p</i> =0.053
		Left	37.56 (5.96)		
	Female	Right	39.95 (6.8)	0.63	<i>p</i> =0.17
		Left	39.32 (6.96)		
Center-edge angle	Male	Right	37.23 (10.52)	1.0	<i>p</i> =0.15
		Left	38.23 (11.14)		
	Female	Right	34.78 (10.52)	0.34	<i>p</i> =0.60
		Left	35.12 (12.36)		

study. Inclusion criteria included a minimum age of 20 years, to ensure completion of puberty/epiphyseal growth plate closure, and a maximum of 30 years to minimize the effects associated with age-related degenerative changes. Studies were identified sequentially from the hospital PACS. All the patients underwent either abdomino-pelvic or chest and abdomino-pelvic CT examination for thoracic or intra-abdominal indications. Patients with soft tissue or bony malignancy, a presenting complaint localized to the pelvis, history of serious trauma, or previous pelvic surgery were excluded until 50 cases for each sex had been identified. One pair of independent observers measured the AI and pelvic tilt (PT), one of these observers with the aid of another calculated the foramen obturator index (FOI) and another pair performed center-edge angle (CEA) measurements. There were five independent observers in total. One of the observers was a consultant musculoskeletal radiologist with 9 years of experience and the other four were radiology trainees under his supervision. Measurements were made using electronic calipers on a high-resolution PACS workstation (Barco, Kortrijk, Belgium).

AI was defined as the angle between a line joining the lateral edge of the acetabular roof and the inferior aspect of the pelvic “teardrop”, and an inter-teardrop line [11] (Fig. 1). The CEA was measured, in accordance with Wiberg’s description [12], as the angle formed by the intersection of a vertical line through the center of the femoral head and a line passing from the center of the femoral head to the lateral sourcil (Fig. 2). PT was estimated from the coronal scout by measuring the distance between the superior border of the pubic symphysis to the

middle of the sacrococcygeal joint (Fig. 3). This measurement taken from the anteroposterior (AP) perspective has been shown to correlate most strongly with PT [14, 20]. FOI was used to assess pelvic rotation and was calculated by dividing the maximum horizontal width of the right obturator foramen by that of the left [21].

Descriptive statistics were obtained for AI, CEA, FOI, and PT and the inter-rater reliability of the observers’ measurements were tested by calculating intra-class correlation coefficients (SPSS 16.0) using average measures in a two-way mixed-effect model with 95% confidence intervals calculated according to the technique described by McGraw and Wong [22]. Pearson correlation coefficients were used to calculate the correlation of PT with AI and CEA. Correlation of FOI with AI and CEA was performed by calculating a ratio of AI and CEA by dividing the measurement for the right hip by the measurement of the left hip. This ratio was compared to FOI using a Pearson correlation coefficient.

## Results

### Patient demographics

The mean age of the patients whose images were used was  $26.3 \pm 5.6$  years (2 SD), with no statistically significant difference between the gender groups (*p*<0.05).

The most common indication for CT examination was abdominal pain (41%), followed by malignancy staging (30%), urological symptoms (15%) and other miscellaneous causes.

**Table 2** Descriptive statistics measurements comparing acetabular inclination (AI) and center-edge-angle (CEA) in right and left hips in males and females

Measure	Side	Gender	Mean (2 SD)	Difference	*Significance
Acetabular inclination	Right	Male	38.52 (6.14)	1.43	<i>p</i> =0.03
		Female	39.95 (6.8)		
	Left	Male	37.56 (5.96)	1.36	<i>p</i> <0.01
		Female	39.32 (6.96)		
Center-edge angle	Right	Male	37.23 (10.52)	2.45	<i>p</i> =0.02
		Female	34.78 (10.52)		
	Left	Male	38.23 (11.14)	3.11	<i>p</i> <0.01
		Female	35.12 (12.36)		

\*Students *t* test

## Normal ranges

The mean AI was  $38.8^\circ$  (2 SD  $32.1\text{--}45.5^\circ$ ) but for men was  $38.0^\circ$  (2 SD  $31.8\text{--}44.1^\circ$ ) and  $39.6^\circ$  (2 SD  $32.7\text{--}46.8^\circ$ ) in women, which was statistically significantly different ( $p < 0.001$ ) (Fig. 4). The intraclass correlation co-efficient (ICC) was  $R=0.67$  (95% CI:  $0.57\text{--}0.75$ ) (Figs. 6a, 7a, b), which corresponds to a very good/substantial inter-rater reliability [23]. There was no statistically significant difference between right and left hips (paired  $t$  test:  $p=0.75$ ) (Fig. 9a) (Tables 1 and 2).

The mean CEA measurement for all patients was  $36.3^\circ$  (95% CI:  $24.9\text{--}47.8^\circ$ ), for men  $37.7^\circ$  (95% CI:  $26.9\text{--}48.5^\circ$ ) and for women  $34.9^\circ$  (95% CI:  $23.5\text{--}46.3^\circ$ ) with a statistically significant gender difference ( $p < 0.001$ ) (Fig. 5). This was performed by a different pair of observers and the ICC was  $R=0.83$  (95% CI:  $0.77\text{--}0.87^\circ$ ) (Figs. 6b, 7c, d), which corresponds to a very good inter-rater reliability [23]. There was no statistically significant difference between the right and left hip (paired  $t$  test:  $p=0.07$ ) (Fig. 9b) (Tables 1 and 2).

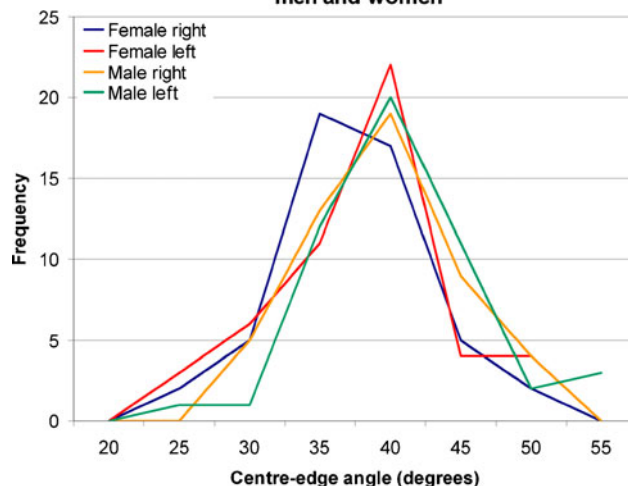
PT was only determined from 48 scouts (48%) due to incomplete visualization of the necessary landmarks. The mean PT measurement (sacro-coccygeal-pubic symphysis) was  $31.3\text{ mm}$  ( $\pm 2$  SD:  $2.8\text{--}59.8\text{ mm}$ ), for men:  $23.9\text{ mm}$  and for women:  $37.5\text{ mm}$ . This was again, statistically significant ( $p < 0.001$ ). The ICC for these measurements was  $R=0.94$  (95% CI:  $0.90\text{--}0.97$ ). There was no apparent correlation between PT and AI measurements (Pearson correlation coefficient  $R=0.014$ , (95%CI:  $-0.27$  to  $0.29$ )).

Landmarks from which the FOI could be calculated were demonstrated on 83 scout images (83%). The mean values for men and women were  $1.07$  ( $\pm 2$  SD:  $0.71\text{--}1.43$ ) and  $1.03$  ( $\pm 2$  SD:  $0.70\text{--}1.36$ ), respectively (Fig. 8). Correlation coefficients for comparison of FOI with right/left hip measurement ratios of AI and CEA were  $r=0.03$  (95% CI:  $-0.18\text{--}0.25$ ) and  $r=0.11$  (95% CI:  $-0.11\text{--}0.32$ ), respectively (Fig. 10a, b).

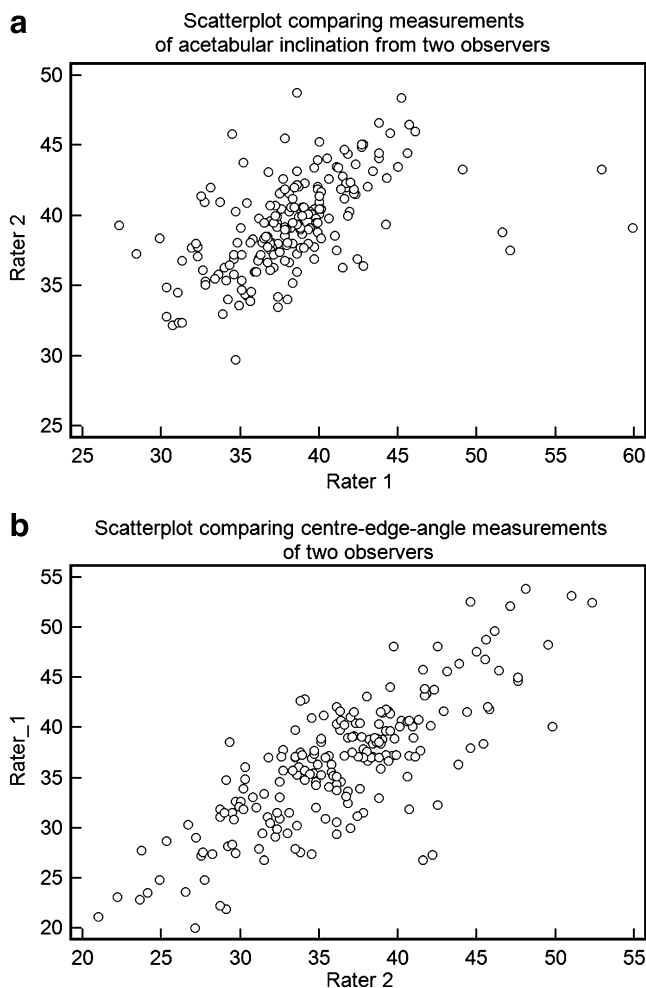
## Discussion

The aim of this study was to determine the statistically normal ranges of two common measures of acetabular morphology. The normal range of AI had previously been quoted as  $25\text{--}40^\circ$  [17] and  $33\text{--}38^\circ$  [11] with a “normal upper limit” of  $39\text{--}42^\circ$ . Our findings of a mean of  $38.8^\circ$  and subsequent 2 SD range of  $32.1\text{--}45.5^\circ$  suggest the range is wider and the upper limit higher than that suggested by earlier work. Previous work has failed to identify a direct correlation between AI and the development of OA [8]. However, this particular study involved the review of radiographs from patients with established unilateral hip

**Frequency polygon of centre-edge angle measurements in men and women**

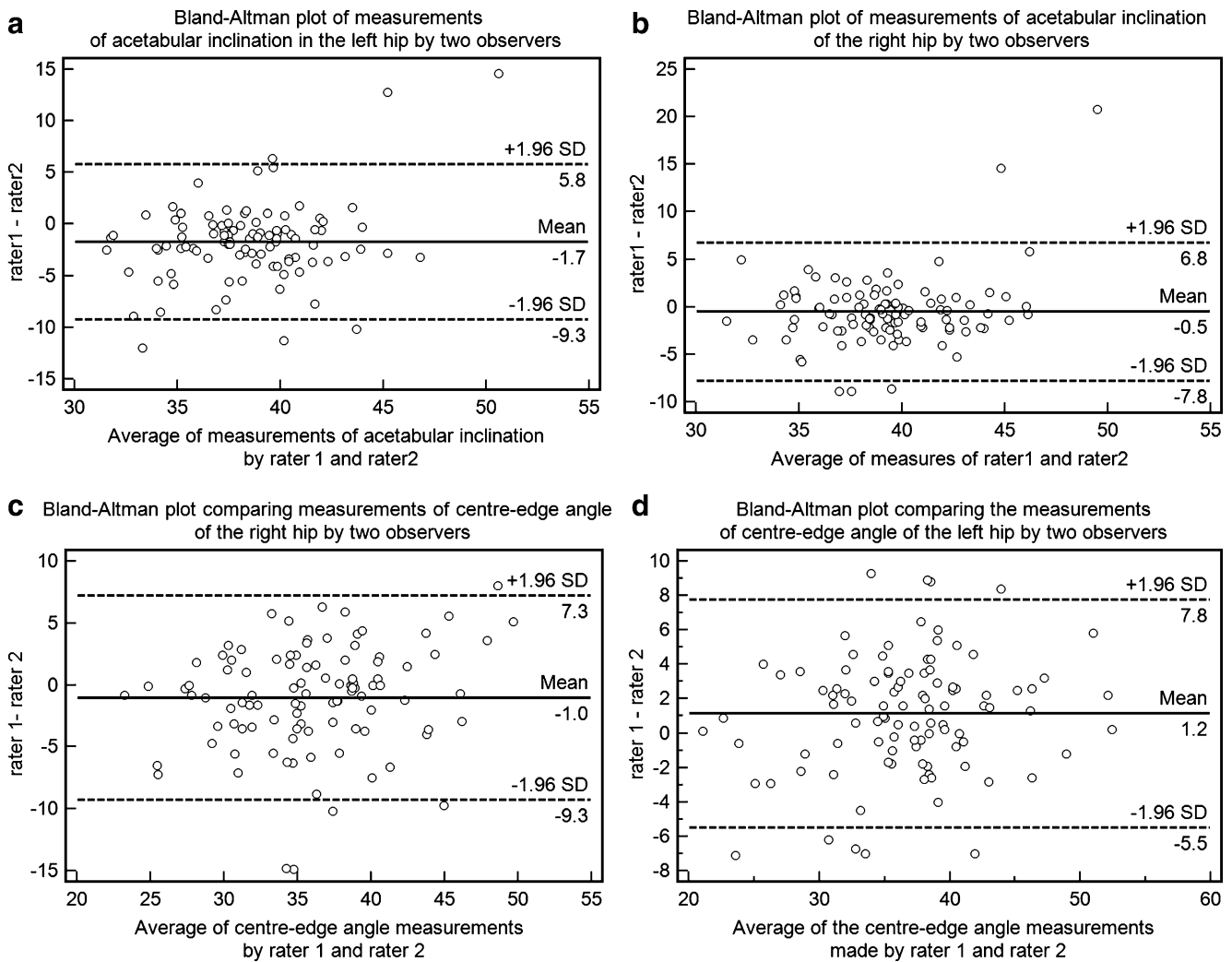


**Fig. 5** Frequency polygon of acetabular inclination



**Fig. 6** a Scatterplot of inter-rater reliability for AI and b CEA





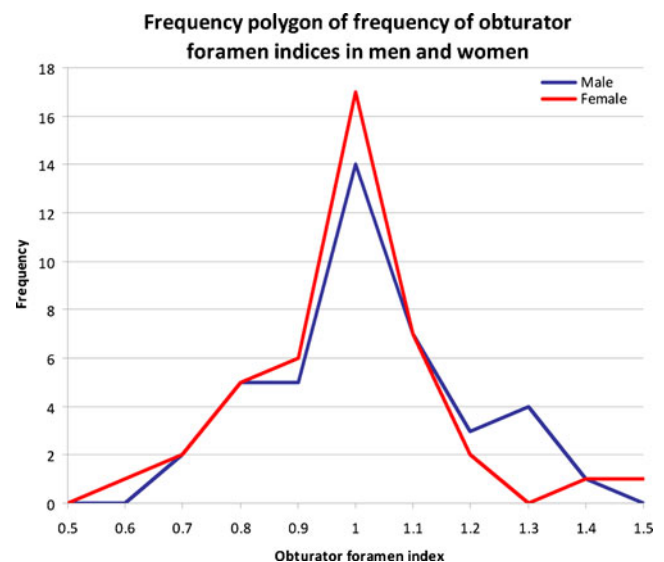
**Fig. 7** Bland-Altman plots for inter-rater reliability for acetabular inclination measurements of the left hip **a** and the right hip **b** and centre-edge-angle measurements in the right **c** and left hip **d**

disease and did not consider that pathology in one hip may precede a change in the mechanical force exerted upon the contralateral side and its subsequent radiological appearance and measurements [24].

The statistically significant different values of AI between genders reflect the differing pelvic conformations. The dimensions of the whole pelvis are greater in the male, but the female pelvic cavity is larger as it is adapted for parturition. Subsequently, the male acetabulum is larger than that in the female due to a wider anterolateral pelvic wall [25].

Interestingly, the reliability of the measures for CEA appears to be better than for AI. This may be related to the use of the inter-teardrop distance as the baseline from measuring AI. Previous work has demonstrated decreases in inter-rater reliability when the teardrop width and various femoral head parameters were used [26].

Pelvic positioning is known to affect most radiographic measures of acetabular morphology [11, 14]. A more



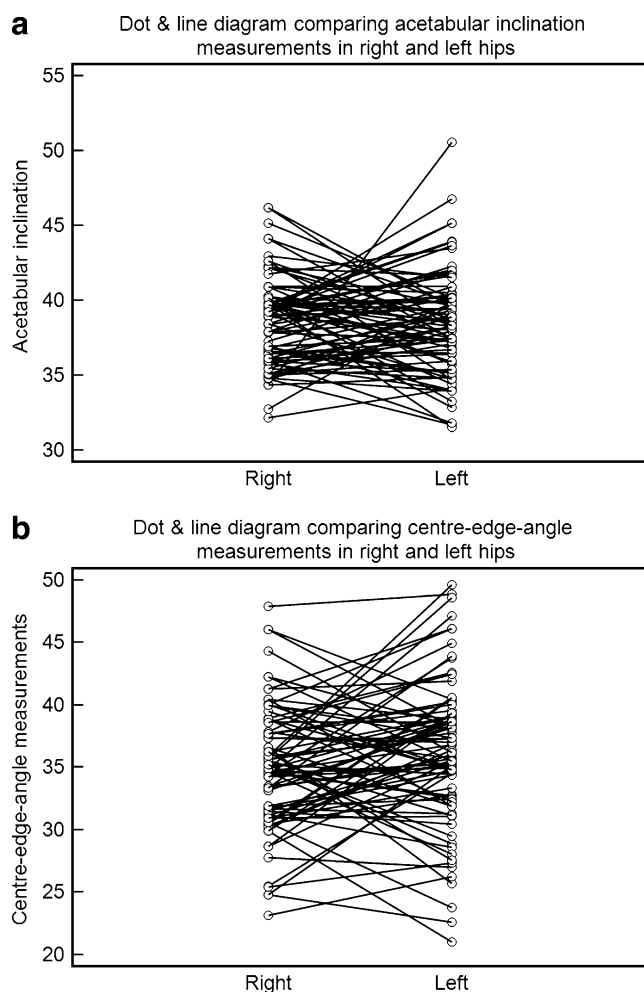
**Fig. 8** Frequency polygon of FOI

accurate representation than that used in this study would have been to measure this on the lateral projection as recommended by a number of authors [27, 28]. Although attempted, this was not possible in our study. The CT scout tomograms use the minimal dose possible to obtain landmarks from which to prescribe the CT examination. This meant that the landmarks required for pelvic tilt measurement could not be demonstrated clearly on any of these studies.

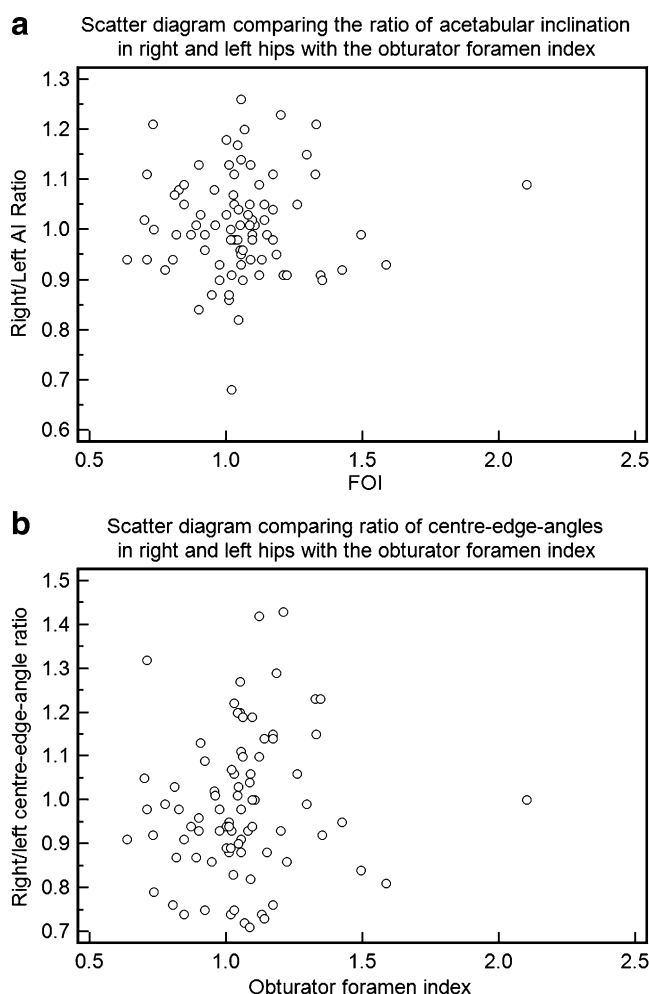
Similar problems affected the measurement of PT from the AP scout; with only 48 (26 female and 22 male) measurements obtained. This was due to superimposition and incomplete demonstration of the necessary anatomical landmarks. There was 100% agreement between the independent observers with regard to which cases were viable for measurement. Interestingly, the PT measurements that were obtained demonstrated poor correlation with AI and CEA. This would make anatomical sense. The orientation of the acetabular cup should be independent of pelvic tilt if a

constant postural orientation with the femoral head is to be obtained. However, although the sample size is adequate to demonstrate a correlation coefficient of  $R=0.5$  (moderate) or above, the large proportion of failed measurements means that selection bias was probably a strong influence. This may also account for the very high intra-class correlation value for the PT measurements. The likelihood of PT acquisition did not appear to be sex-related. The AP-PT measurement was greater in women than men, reflecting their more ovoid pelvic shape. This difference was found to be statistically significant ( $p<0.001$ ) in keeping with previous data [14], despite the fact that our data was acquired from a younger population. It again serves to reinforce the morphological differences between genders.

While the AP acquisition of PT is shown to correlate most strongly with that measured on the lateral view [9, 20] and is a reliable indicator of relative changes in PT, it is not accurate enough to determine absolute pelvic position. This is attributed to variations in sacral anatomy [9].



**Fig. 9** Dot and line diagrams demonstrating poor correlation of AI **a** and CEA **b** between right and left hips



**Fig. 10** Scatterplots demonstrating poor correlation between the ratio of right and left hip measurements of AI **a** and CEA **b** with FOI

Various methods to assess PT have been described; those based on AP radiographs [13, 24], use of an inclinometer [29]; trigonometric [15, 30] and 3D CT reconstruction models [31]. However, to date, the latter model is restricted to research use and thus cadaveric pelvises only. Ultimately, a lateral radiograph remains the “gold standard”, and this remains vulnerable to intersubject variability, assuming consistent positioning of anatomical landmarks.

The FOI in providing a measure of pelvic rotation demonstrated that, as is to be expected, the lateral tilt of the pelvis varied about a mean in a parametric distribution with small standard deviations (Fig. 8), suggesting that the sample was a not unreasonable representation of a true anteroposterior projection. Correlation of FOI with AI was “poor” (Fig. 10a) and with CEA “fair” (Fig. 10b). This suggests that there is not a strong correlation between pelvic rotation and the measurements of AI and CEA in this sample.

It must be stressed that the normal ranges provided by this and other studies are based upon the relationship of the acetabulum and femur in the supine position. Therefore, they cannot necessarily be assumed to represent their dynamic, weight-bearing situation. Although a study evaluating the effect of supine, sitting, and standing positions upon the pelvis in relation to the optimal placement of prosthetic acetabular components has suggested that it is reasonable to consider the supine position as equivalent to its functional one [31].

There are other potential confounding factors. The population selected is not truly representative of healthy individuals as they are symptomatic. For instance, many of them may have had pain that may have had a further influence on position of the pelvis. However, in the absence of being able to irradiate the pelvises of healthy asymptomatic volunteers, the CT scout tomogram can be a useful tool for deriving normative anatomical data [32].

## Conclusions

The reference range for AI in healthy individuals, derived from CT scout tomograms, is 32.1 to 45.5°. That of CEA is 22.5 to 50.1°. There is a statistically significant difference in the means, of 1.6° for AI and 2.8° for CEA, between men and women.

**Conflict of interest** None.

## References

- Murray DW. The definition and measurement of acetabular orientation. *J Bone Jt Surg Br.* 1993;75(2):228–32.
- Giori NJ, Trousdale RT. Acetabular retroversion is associated with osteoarthritis of the hip. *Clin Orthop Relat Res.* 2003;417:263–9.
- Tönnis D, Heinecke A. Acetabular and femoral anteversion: relationship with osteoarthritis of the hip. *J Bone Jt Surg Am.* 1999;81(12):1747–70.
- Wagner S, Hofstetter W, Chiquet M, Mainil-Varlet P, Stauffer E, Ganz R, et al. Early osteoarthritic changes of human femoral head cartilage subsequent to femoro-acetabular impingement. *Osteoarthritis and Cartilage / OARS.* 2003;11(7):508–18.
- Leunig M, Beck M, Dora C, Ganz R. Femoroacetabular impingement: trigger for the development of coxarthrosis. *Orthopade.* 2006;35(1):77–84.
- Ganz R, Leunig M, Leunig-Ganz K, Harris WH. The etiology of osteoarthritis of the hip: an integrated mechanical concept. *Clin Orthop Relat Res.* 2008;466(2):264–72.
- Steppacher SD, Tannast M, Werlen S, Siebenrock KA. Femoral morphology differs between deficient and excessive acetabular coverage. *Clin Orthop Relat Res.* 2008;466(4):782–90.
- Vossinakis IC, Georgiades G, Kafidas D, Hartofilakidis G. Unilateral hip osteoarthritis: can we predict the outcome of the other hip? *Skeletal Radiol.* 2008;37(10):911–6.
- Tannast M, Langlotz U, Siebenrock KA, Wiese M, Bernsmann K, Langlotz F. Anatomic referencing of cup orientation in total hip arthroplasty. *Clin Orthop Relat Res.* 2005;436:144–50.
- Haenle M, Heitner A, Mittelmeier W, Barbano R, Scholz R, Steinhäuser E, et al. Assessment of cup position from plain radiographs: impact of pelvic tilting. *Surg Radiol Anat.* 2007;29(1):29–35.
- Sharp IK. Acetabular dysplasia: the acetabular angle. *J Bone Jt Surg Br.* 1961;43(2):268.
- Wiberg G. Studies on dysplastic acetabula and congenital subluxation of the hip joint. *Acta Chir Scand.* 1939;83:1–132.
- Lewinnek GE, Lewis JL, Tarr R, Compere CL, Zimmerman JR. Dislocations after total hip-replacement arthroplasties. *J Bone Jt Surg Am.* 1978;60(2):217–20.
- Siebenrock KA, Kalbermatten DF, Ganz R. Effect of pelvic tilt on acetabular retroversion: a study of pelves from cadavers. *Clin Orthop Relat Res.* 2003;407:241–8.
- Nagao Y, Aohi H, Ishii SJ, Masuda T, Beppu M. Radiographic method to measure the inclination angle of the acetabulum. *J Orthop Sci Jap Orthop Assoc.* 2008;13(1):62–71.
- Kalteis TA, Handel M, Herbst B, Grifka J, Renkawitz T. In vitro investigation of the influence of pelvic tilt on acetabular cup alignment. *J Arthroplasty.* 2009;24(1):152–7.
- Stulberg SD, Harris WH. 1974. Acetabular dysplasia and development of osteoarthritis of the hip. In *The Hip: Proceeding of the Second Open Scientific Meeting of the Hip Society.* St Louis, Mo: CV Mosby, pp. 83–93
- Armbuster TG, Guerra Jr J, Resnick D, Goergen TG, Feingold ML, Niwayama G, et al. The adult hip: an anatomic study. Part I: the bony landmarks. *Radiology.* 1978;128(1):1–10.
- Lembeck B, Mueller O, Reize P, Wuelker N. Pelvic tilt makes acetabular cup navigation inaccurate. *Acta Orthop.* 2005;76(4):517–23.
- Tannast M, Murphy SB, Langlotz F, Anderson SE, Siebenrock KA. Estimation of pelvic tilt on anteroposterior X-rays—a comparison of six parameters. *Skeletal Radiol.* 2006;35(3):149–55.
- Tönnis D. Normal values of the hip joint for the evaluation of X-rays in children and adults. *Clin Orthop.* 1976;119:39–47.
- McGraw K, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Meth.* 1996;1(1):30–46.
- Altman DG, 1990. *Practical Statistics for Medical Research* 1st ed., Chapman and Hall/CRC
- Kojima A, Nakagawa T, Tohkura A. Simulation of acetabular coverage of femoral head using anteroposterior pelvic radiographs. *Arch Orthop Trauma Surg.* 1998;117(6–7):330–6.

25. Sandring S, Healy J, Johnson D, Williams A, Ellis H (eds.) 2004. Pelvic girdle and lower limb. In Gray's Anatomy: The Anatomical Basis of Clinical Practice. 39th Ed. Churchill Livingstone, p 1430
26. Pederson DR, Lamb CA, Dolan MA, Ralston BS, Weinstein SL, Morcuende JA. Radiographic measurements in developmental dysplasia of the hip: reliability and validity of a digitizing program. *J Ped Orthop*. 2004;24(2):156–60.
27. Richards PJ, Pattison JM, Belcher J, DeCann RW, Anderson S, Wynn-Jones C. A new tilt on pelvic radiographs: a pilot study. *Skeletal Radiol*. 2009;38(2):113–22.
28. Tannast M, Zheng G, Anderegg C, Burckhardt K, Langlotz F, Ganz R, et al. Tilt and rotation correction of acetabular version on pelvic radiographs. *Clin Orthop Relat Res*. 2005;438:182–90.
29. Anda S, Svenningsen S, Grontvedt T, Benum P. Pelvic inclination and spatial orientation of the acetabulum. A radiographic, computed tomographic and clinical investigation. *Acta Radiol*. 1990;31(4):389–94.
30. Sanders G, Starvakas P. A technique for measuring pelvic tilt. *Phys Ther*. 1981;61(1):49–50.
31. Nishihara S, Sugano N, Nishii T, Ohzono K, Yoshikawa H. Measurements of pelvic flexion angle using three-dimensional computed tomography. *Clin Orthop Relat Res*. 2003;411:140–51.
32. Crawford MB, Toms AP, Shepstone L. Defining normal vertebral angulation at the thoracolumbar junction. *AJR Am J Roentgenol*. 2009;193(1):W33–7.



# Does running cause metatarsophalangeal joint effusions? A comparison of synovial fluid volumes on MRI in athletes before and after running

Amy-Rose Kingston · Andoni P. Toms ·  
Subhadip Ghosh-Ray · Shelley Johnston-Downing

Received: 29 September 2008 / Revised: 21 November 2008 / Accepted: 26 December 2008 / Published online: 30 January 2009  
© ISS 2009

## Abstract

**Objective** The metatarsophalangeal joints (MTPJ) are the only joints that bear weight directly through synovium. The purpose of this study was to determine whether there is an association between synovial stresses during running and increases in volume of joint fluid.

**Materials and methods** This was a prospective case controlled study (nine healthy athlete volunteers acting as own controls). High-resolution coronal 3D T2W magnetic resonance imaging of the MTPJs were obtained following 24 h rest and after a 30-min run. The volume of joint fluid in each MTPJ ( $n=90$ ) was measured by two independent observers using an automated propagating segmentation tool.

**Results** The median volume of synovial fluid in the MTPJs at rest was 0.018 ml (inter-quartile range (IQ) range 0.005–0.04) and after running 0.019 ml (IQ range 0.005–0.04,  $p=0.34$ , 99% confidence interval (CI), 0.330.35). The volume of fluid in the MTPJs of the great toes was substantially larger than other toes (0.152 ml at rest, 0.154 ml after exercise,  $p=0.903$ ). Median volumes decrease from second to fifth MTPJs (0.032–0.007 ml at rest and 0.035–0.004 ml after exercise). Subset analysis for each toe revealed no significant difference in volumes before and after running ( $p=0.39$  to  $p=0.9$ ). The inter-rater reliability for observer measurements was good with an intra-class correlation of 0.70 (95% CI, 0.60 to 0.78).

**Conclusion** It appears to be normal to find synovial fluid, particularly in the MTPJs of the great toes, of athletes at rest and after running. There does not appear to be an association between moderate distance running and an increase in the volume of synovial fluid.

**Keywords** Metatarsophalangeal · Joint · Effusion · MRI · Athlete

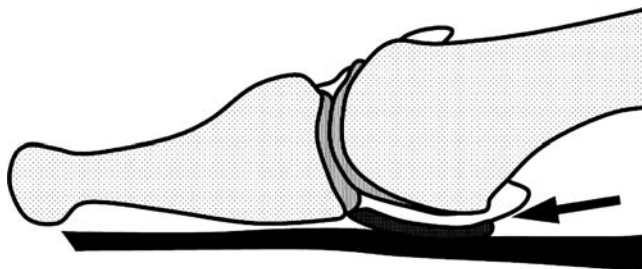
## Introduction

The metatarsophalangeal joints (MTPJ) are condyloid synovial joints whose function is critical to generating speed during running [1]. However, they are uniquely vulnerable because, of all the weight-bearing joints, they alone absorb significant amounts of energy through the synovium. During the stance phase of running, 20 J are absorbed by the metatarsal heads [2] through the synovium on the plantar aspect of the joint (Fig. 1). It has been demonstrated, in laboratory studies, that synovium is mechanosensitive [3, 4] with hyaluron being rapidly secreted [5] after only short episodes of mild stretching [6] in order to protect the articular surfaces. It would therefore be reasonable to suspect that stretching of this synovium might also lead to an increase in secreted synovial fluid. The presence of this physiological joint fluid might then be mistaken for a pathological joint effusion.

Magnetic resonance imaging (MRI) has been used to describe stress-related changes in bone marrow in the lower limb following abnormal biomechanical loading [7], team sports [8] and endurance running [9]. It is normal to demonstrate synovial fluid in the hip [10], ankle [11] and tendon sheaths around the ankle [12] in asymptomatic

This study was funded by a research grant from Action Arthritis.

A.-R. Kingston · A. P. Toms (✉) · S. Ghosh-Ray ·  
S. Johnston-Downing  
Department of Radiology,  
Norfolk & Norwich University Hospital Trust,  
Colney Lane,  
Norwich, Norfolk NR4 7UY, UK  
e-mail: andoni.toms@nnuh.nhs.uk



**Fig. 1** Diagrammatic representation of a mid-sagittal section through the first metatarsophalangeal joint demonstrating the plantar joint recess where synovium lines the dorsal surface of the plantar plate (arrow)

volunteers. In comparison, it is uncommon to demonstrate fluid on MRI in the non-weight-bearing glenohumeral joint [11]. However, any association between joint effusions in weight-bearing joints and exercise is far from clear. A comparison of professional marathon runners and asymptomatic volunteers demonstrated ankle effusions in 34% of the endurance athletes but also in 18% of the controls [13]. Indeed, the association between synovitis and joint effusions is the matter of some debate. It is not surprising that joint effusions can be demonstrated in the small joints of patients with rheumatoid arthritis who do not clinically have synovitis [14], but studies have also failed to demonstrate a correlation between effusions demonstrated on MRI of the temporomandibular joints and histopathological measures of synovitis [15].

The first aim of this project was to measure the incidence and volume of fluid in the MTPJs of asymptomatic endurance runners by segmenting volumes of fluid demonstrated on high-resolution T2W MRI. Automated volume segmentation has been demonstrated to be an accurate tool for measuring volumes of organs [16–18] and disease [19–23]. In particular, MRI has proved to be a reliable technique for measuring joint effusions. Segmenting volumes of fluid from T2W MR images is an accurate technique for measuring absolute volumes of fluid [24, 25] with reasonable inter-observer reliability [25]. The second aim of the project was to test the null hypothesis that running was not associated with an increase in volume of synovial fluid in the MTPJs in this group of volunteers.

## Materials and methods

This was a prospective cross-sectional and comparative study of MTPJ effusions before and after running. Governance approval was obtained on 15th June 2006, and ethical approval was obtained on the 12th December 2006. Approval was given for a pilot study to recruit up to ten volunteers.

## Subjects

Nine asymptomatic endurance runners from the University of East Anglia running club were recruited for the study. Informed consent from each volunteer was obtained. The criteria for inclusion were that volunteers must be able to run for between 30 and 45 min and must be aged between 18 and 30. Exclusion criteria included a history of foot trauma, foot surgery, inflammatory joint disease or contraindications to MRI. The nine participants consisted of three men and six women, with an average age of 24 and a range from 19 to 30 years. Heights ranged from 5 ft 5 in. to 6 ft 3 in. with an average height of 5 ft 7 in. The mean body mass index was 22.6, and all volunteers were in the normal range (19–26.2). The exercise history revealed that the volunteers ran between two and nine times per week. The duration of each run varied from 30 min to 1 h and 30 min. The volunteers had been running for between 6 and 17 years.

## MRI

Participants were asked not to exercise for 24 h prior to their first MRI. Both feet were imaged with a high-resolution fluid-sensitive MR sequence (Siemens Avanto 1.5 Tesla MRI machine, Erlangen, Germany. Coronal 3D T2W: TR 12.70, TE 6.30, slice thickness 0.5 mm, field of view 11×15 cm, matrix size 240×320). Participants were then asked to run for 30 min without stopping. The MRI was then repeated approximately 30 min after the runner's return.

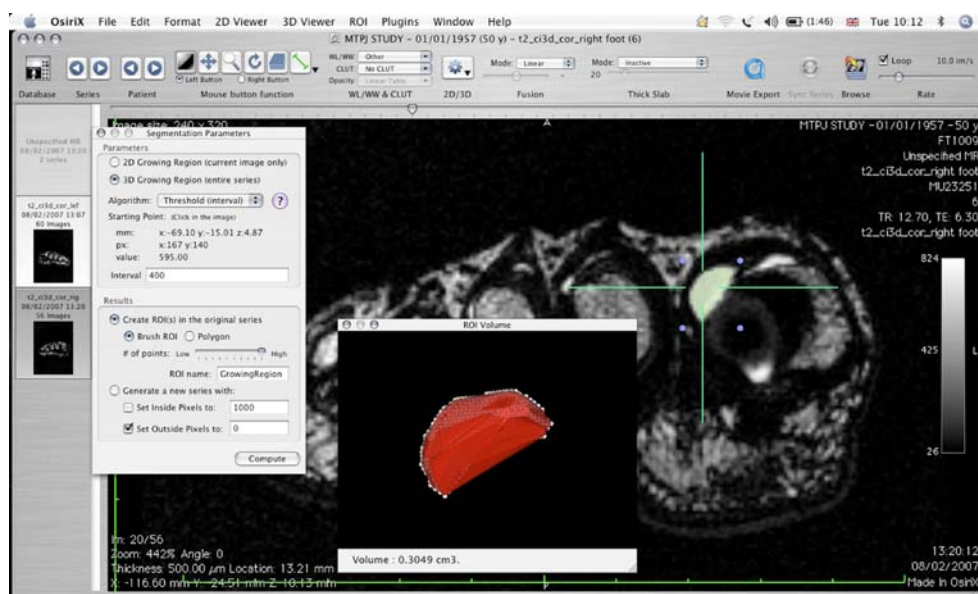
## Image analysis

The 18 MRI examinations were anonymised and assigned randomised identification numbers ([www.randomnumber.org](http://www.randomnumber.org)) so that the control and post-run MRI could not be identified. This was performed by a member of staff outside the research team.

The volume of fluid within each MTPJ was measured by two independent observers (consultant radiologist and radiology trainee). The volumes of fluid were obtained using an automated propagating 3D segmentation tool (Osirix version 2.6) on an AppleMac G4 Powerbook with a 23-in. high-resolution external monitor (Fig. 2). The observers defined the limiting values for the automated segmentation to achieve a best fit for each joint, and the total volume of synovial fluid for each toe was recorded in millilitres.

Wilcoxon signed rank tests of differences were applied as a non-parametric test of paired differences in the two arms of the study, and inter-observer reliability was tested using intra-class correlations (SPSS version 14.0) [26] as a measure of reliability.

**Fig. 2** Screenshot of volume measurements being obtained from a metatarsophalangeal joint using an automated propagating 3D segmentation tool (Osirix v 2.6)

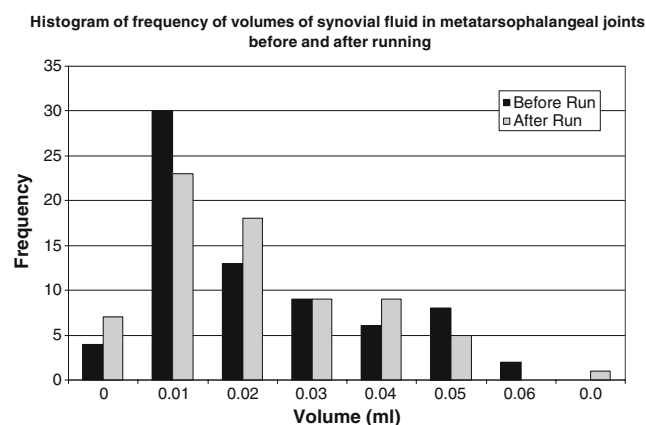


## Results

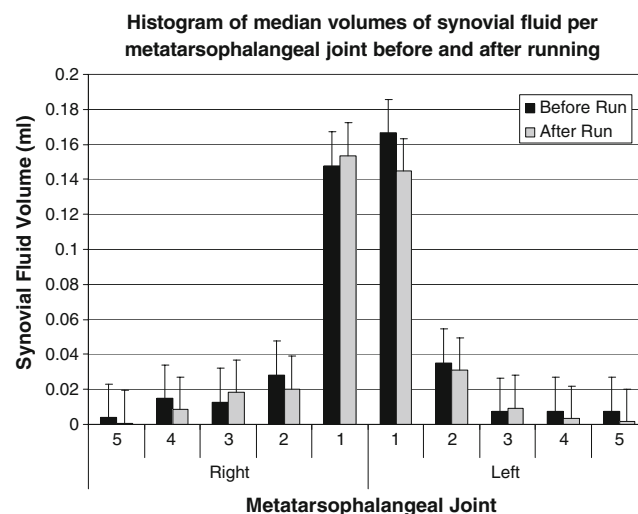
This distribution of volumes across the data set was non-parametric (Fig. 3). The median volume of synovial fluid in all 90 MTPJs before running was 0.0183 ml (inter-quartile range 0.005 to 0.043) and after running was 0.019 ml (inter-quartile range 0.005 to 0.04). Wilcoxon signed rank analysis of difference of the medians unsurprisingly demonstrated no significant difference between the volumes before and after running ( $p=0.341$ , 99% confidence intervals 0.329 to 0.353). The volume of fluid found in the metatarsophalangeal joints of the great toes was obviously much larger than in the other toes (Fig. 4), and therefore, subset analysis was performed on the volumes of fluid measured in each MTPJ (Table 1). The median synovial fluid volumes before and after running and the

test results of the differences in the medians (Wilcoxon signed rank test) were as follows: great toes, 0.152 ml before run, 0.154 ml after run,  $p=0.903$ ; second toes, 0.0325 ml before run, 0.035 ml after run,  $p=1.0$ ; third toes, 0.013 ml before run, 0.017 ml after run,  $p=0.39$ ; Fourth toes, 0.017 ml before run, 0.01 ml after run,  $p=0.429$ ; fifth toes, 0.007 ml before run, 0.004 ml after run,  $p=0.465$ . None of the subset analysis revealed any significant difference between the volumes of synovial fluid before and after running.

The standard errors of the means for the combined volumes before and after running for the first to fifth



**Fig. 3** Histogram demonstrating the non-parametric distribution of frequencies of MTPJ fluid volume measurements ( $\text{cm}^3$ ) per toe for the two observers (360 observations)



**Fig. 4** Histogram demonstrating the median volume of metatarsophalangeal fluid per toe (error bars represent the inter-quartile ranges). The median volumes per toe are based on the mean volume measured by the two observers

**Table 1** Table of analysis of differences in median volumes of MTPJ synovial fluid categorised by position of the toe

MTPJ	1st	2nd	3rd	4th	5th	All toes ( <i>n</i> =90)
Volume of synovial fluid before run <sup>a</sup>	0.15 (0.073 to 0.21)	0.033 (0.016 to 0.043)	0.013 (0.005 to 0.022)	0.017 (0.006 to 0.023)	0.007 (0.002 to 0.015)	0.018 (0.005 to 0.043)
Volume of synovial fluid after run <sup>a</sup>	0.15 (0.09 to 0.21)	0.035 (0.013 to 0.041)	0.017 (0.006 to 0.028)	0.010 (0.002 to 0.018)	0.004 (0 to 0.0114)	0.019 (0.005 to 0.04)
Standard error <sup>b</sup>	0.015	0.003	0.002	0.002	0.001	0.07
Difference P <sup>c</sup>	0.90 (0.89 to 0.91)	1.0	0.39 (0.38 to 0.40)	0.43 (0.42 to 0.44)	0.47 (0.45 to 0.47)	0.34 (0.33 to 0.35)

<sup>a</sup> Median volume in millilitre with inter-quartile range<sup>b</sup> Standard error calculated for mean volumes before and after running<sup>c</sup> Statistical significance of difference of non-parametric paired means with 99% confidence intervals (Wilcoxon signed rank test)

MTPJs were 0.015, 0.003, 0.002, 0.002 and 0.001 ml, respectively (Table 1). The intra-class coefficient (ICC) for 360 measurements (180 per observer) was good at 0.704 (95% confidence intervals, 0.602 to 0.779). Most of the observations were clustered fairly close together below 0.01 ml but demonstrated greater inter-observer variation in the larger volumes measured in the great toe MTPJs (Fig. 5).

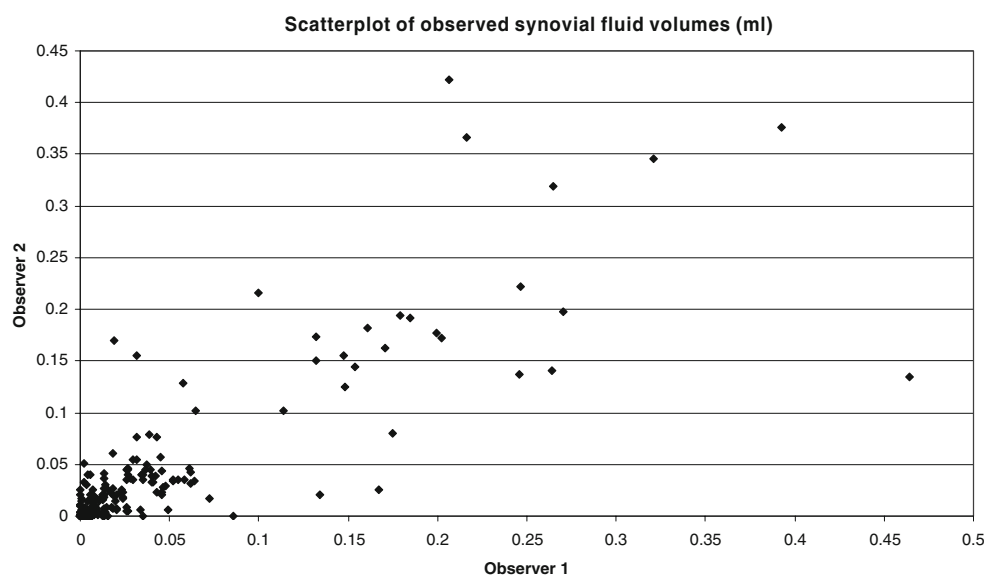
## Discussion

It is uncertain whether any rise in volume of MTPJ synovial fluid might be a normal mechanosensitive physiological response to synovial stretching or a pathological effusion resulting from joint trauma. Therefore, the term “effusion” has not been used to describe the joint fluid in these asymptomatic volunteers. The results of this study indicate that it is normal to find small amounts of joint fluid in the MTPJs of young asymptomatic athletes and that vigorous weight-bearing exercise appears to have little effect on the subsequent volume. By far, the largest volume of joint fluid occurs in the MTPJs of the great toes with substantially less in the other toes. The range of volumes is large with many joints demonstrating little or no synovial fluid.

The findings of this study echo similar findings describing joint effusions in the hips [10] and ankles [11] of normal volunteers as well as in the tendon sheaths around the ankle [12]. However, in contrast to published descriptions of bone marrow oedema adjacent to joints following athletic endeavour [8, 9] or abnormal biomechanics [7], the results of this study do not support the hypothesis that weight bearing on the synovium of the MTPJ results in an increased volume of synovial fluid.

It might be reasonable to assume that the synovial fluid measurements, in the range of volumes described in this paper and presenting in isolation, are normal findings in keeping with previous descriptions of joint and tendon sheath effusions elsewhere [7–9]. However, it could be argued that this synovial fluid is a feature of this group of athletes rather the normal population. The fact that there was no significant difference between the volumes of synovial fluid measured before and after running may also be the result of bias in the selection of the volunteers. Enthusiastic long-distance runners may have favourable MTPJ biomechanics when compared with the general population which provide them with some protection from the trauma of load bearing. A similar protective effect of training on the incidence of muscle injuries has previously been suggested [13]. The research question purposely did not include asymptomatic non-runners for several reasons. There was no primary aim to describe normative data in non-athletes. One of the aims of the study was to test the

**Fig. 5** Scatter plot comparing the volumes of fluid per foot measured by the two independent observers. This demonstrates a roughly linear correlation between the two observers but with some outlying measurements at the larger volumes



null hypothesis that running was not associated with an increase in synovial fluid volume, and therefore it was felt that the intervention, in other words the run, would be more reliably reproduced in experienced runners. Non-runners would be at risk of injury if not adequately prepared for a run and this was probably not ethically acceptable.

There is some indirect evidence that running might be responsible for joint effusions. A slightly increased volume in ankle effusions following marathon running has been demonstrated when compared to normal controls [13]. However, no comparison was made with the volume of fluid before running, and therefore, the possibility of selection bias persists. The length of run in our study may not have been sufficient to induce an effusion in this group of well-trained athletes, although laboratory studies do suggest that as little 10 min of synovial stretch is all that is required to stimulate a mechanosensitive response [6]. It may be the case that running can cause a physiological, or possibly pathological, increase in joint fluid in those who are not used to running.

The length of rest before the first MRI (at least 24 h) and the timing of the second MRI (within the first hour of completing the run) may also have an influence on the presence of synovial fluid. Clinical experience with pathological effusions would suggest that the rest period and the time to MR after running might both be too short. However, the pathophysiology of traumatic effusions and normal physiological variations in joint fluid are different. Traumatic effusions often continue to increase in size over the first 24 h as blood in the joint induces a synovitis, whereas there is evidence that synovium responds to stress, in order to protect the joint, within 20 min of the stimulus [5]. The volunteers in this study underwent MR examination after a minimum

of 30 min rest and therefore at least 1 h after starting exercise. Therefore, this was not an unreasonable interval after which to image, but it may have been more sensitive if the second MR was delayed for 3 or 4 h after the run. There is no evidence in the literature of how quickly normal synovial fluid is reabsorbed. Our anecdotal experience from MR arthrography is that fluid can be absorbed rapidly from a joint in the first hour after injection, and therefore, 24 h is probably adequate for fluid that is not the result of a pathological effusion. Therefore, clinical experience of the resolution of pathological effusions should probably not be used to inform the timing of the rest period, but again, a rest period of greater than 24 h may have proved more effective.

The number of volunteers ( $n=9$ ) in this pilot study is relatively small, although the number of toes is large ( $n=90$ ). The study has a 90% power to detect a significant difference ( $p=0.05$ ) of 0.08 ml in the MTPJs of the great toes before and after running with a significance ( $n=18$ ). In the second to fifth MTPJs, the study has a 90% power to demonstrate a significant difference ( $p=0.05$ ) in synovial fluid volume of approximately 0.004 ml ( $n=72$ ; one-sample  $t$  test power calculation—R 2.4.1, the R Foundation for Statistical Computing).

This represents a change in volume in the order of 25–50%. While it may be possible to demonstrate smaller significant differences with larger numbers of volunteers, this is unlikely to be a clinically useful finding.

There are a number of possible sources of error in the study. The volumes of fluid measured were very small. To minimise standard measurement errors, a very high-resolution thin slice select matrix was used which resulted in a theoretical acquired voxel measuring 0.00125 ml, although with interpolation, this is even smaller on the final image. This resulted in acceptable



standard errors for the larger volumes which obviously increased as the mean volume of the MTPJ fluid decreased from first to fifth toe. The accuracy of the measurements also depends on both the robustness of the segmentation algorithm in Osirix® [27] and the observers' decision to include a given starting point and the choice of threshold values for segmentation.

Inter-observer reliability is a much more likely cause of error in this study. While the ICC of 0.704 is considered good, it can be seen from the scatter plot (Fig. 5) that the smaller volume measurements correlate better than the larger. The volumes of synovial fluid in the first MTPJs were often complex shapes which is the probable cause of the increased inter-observer error. The range of pixel values for fluid varies from slice to slice, and as the effusions at the first MTPJs usually extended the greatest from anterior to posterior, this variation can result in "bleeding" of the segmented volume into surrounding tissues. The opposite can also occur, and the volume of fluid can be under-segmented.

## Conclusion

It appears to be normal to find synovial fluid in the MTPJs of asymptomatic athletes following 24 h rest. The largest volumes of synovial fluid are present in the MTPJs of the great toes; these are significantly larger than volumes in the second to fifth toes. This study demonstrates no association between volumes of synovial fluid in the MTPJs and moderate distance running.

## References

1. Krell JB, Stefanyshyn DJ. The relationship between extension of the metatarsophalangeal joint and sprint time for 100 m Olympic athletes. *J Sports Sci* 2006; 24: 175–180.
2. Stefanyshyn DJ, Nigg BM. Mechanical energy contribution of the metatarsophalangeal joint to running and sprinting. *J Biomech* 1997; 30: 1081–1085.
3. Momberger TS, Levick JR, Mason RM. Hyaluronan secretion by synoviocytes is mechanosensitive. *Matrix Biol* 2005; 24: 510–519.
4. Momberger TS, Levick JR, Mason RM. Mechanosensitive synoviocytes: a  $\text{Ca}^{2+}$ -PKC $\alpha$ -MAP kinase pathway contributes to stretch-induced hyaluronan synthesis in vitro. *Matrix Biol* 2006; 25: 306–316.
5. Coleman PJ, Scott D, Ray J, Mason RM, Levick JR. Hyaluron secretion into the synovial cavity of rabbit knees and comparison with albumin turnover. *J Physiol* 1997; 503: 645–656.
6. Dowthwaite GP, Ward AC, Flannely J, et al. The effect of mechanical strain on hyaluronan metabolism in embryonic fibrocartilage cells. *Matrix Biol* 1999; 18: 523–532.
7. Schweitzer ME, White LM. Does altered biomechanics cause marrow edema? *Radiology* 1996; 198: 851–853.
8. Major NM, Helms CA. MR imaging of the knee: findings in asymptomatic collegiate basketball players. *AJR Am J Roentgenol* 2002; 179: 641–644.
9. Schueller-Weidekamm C, Schueller G, Uffmann M, Bader TR. Does marathon running cause acute lesions of the knee? Evaluation with magnetic resonance imaging. *Eur Radiol* 2006; 16: 2179–2185.
10. Moss SG, Schweitzer ME, Jacobson JA, et al. Hip joint fluid: detection and distribution at MR imaging and US with cadaveric correlation. *Radiology* 1998; 208: 43–48.
11. Schweitzer ME, Magbalon MJ, Fenlin JM, Frieman BG, Ehrlich S, Epstein RE. Effusion criteria and clinical importance of glenohumeral joint fluid: MR imaging evaluation. *Radiology* 1995; 194: 821–824.
12. Schweitzer ME, van Leersum M, Ehrlich SS, Wapner K. Fluid in normal and abnormal ankle joints: amount and distribution as seen on MR images. *AJR Am J Roentgenol* 1994; 162: 111–114.
13. Lohman M, Kivisaari A, Vehmas T, Kallio P, Malmivaara A, Kivisaari L. MRI abnormalities of foot and ankle in asymptomatic, physically active individuals. *Skelet Radiol* 2001;30:61–66.
14. Luukkainen RK, Saltyshov M, Koski JM, Huhtala HS. Relationship between clinically detected joint swelling and effusion diagnosed by ultrasonography in metatarsophalangeal and talocrural joints in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2003; 21: 632–634.
15. Segami N, Suzuki T, Sato J, Miyamaru M, Nishimura M, Yoshimura H. Does joint effusion on T2 magnetic resonance images reflect synovitis? Part 3. Comparison of histologic findings of arthroscopically obtained synovium in internal derangements of the temporomandibular joint. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2003; 95: 761–766.
16. Betrouni N, Dewalle AS, Puech P, Vermandel M, Rousseau J. 3D delineation of prostate, rectum and bladder on MR images. *Comput Med Imaging Graph* 2008; 32: 622–630.
17. Tutschek B, Sahn DJ. Semi-automatic segmentation of fetal cardiac cavities: progress towards an automated fetal echocardiogram. *Ultrasound Obstet Gynecol* 2008; 32: 176–180.
18. Wijesooriya K, Weiss E, Dill V, et al. Quantifying the accuracy of automated structure segmentation in 4D CT images using a deformable image registration algorithm. *Med Phys* 2008; 35: 1251–1260.
19. Aristophanos M, Penney BC, Pelizzari CA. The development and testing of a digital PET phantom for the evaluation of tumor volume segmentation techniques. *Med Phys* 2008; 35: 3331–3342.
20. Dehmshki J, Amin H, Valdivieso M, Ye X. Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. *IEEE Trans Med Imaging* 2008; 27: 467–480.
21. Fabel M, von Tengg-Kobligh H, Giesel FL, et al. Semi-automated volumetric analysis of lymph node metastases in patients with malignant melanoma stage III/IV—a feasibility study. *Eur Radiol* 2008; 18: 1114–1122.
22. Kollorz EK, Hahn DA, Linke R, Goecke TW, Hornegger J, Kuwert T. Quantification of thyroid volume using 3-D ultrasound imaging. *IEEE Trans Med Imaging* 2008; 27: 457–466.
23. Street E, Hadjiiski L, Sahiner B, et al. Automated volume analysis of head and neck lesions on CT scans using 3D level set segmentation. *Med Phys* 2007; 34: 4399–4408.
24. Heuck AF, Steiger P, Stoller DW, Gluer CC, Genant HK. Quantification of knee joint fluid volume by MR imaging and CT using three-dimensional data processing. *J Comput Assist Tomogr* 1989; 13: 287–293.
25. Ostergaard M, Stoltenberg M, Henriksen O, Lorenzen I. The accuracy of MRI-determined synovial membrane and joint effusion volumes in arthritis. A comparison of pre- and post-aspiration volumes. *Scand J Rheumatol* 1995; 24: 305–311.
26. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 2: 420–428.
27. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 2004; 17: 205–216.

# A Reliability Study of Measurement Tools Available on Standard Picture Archiving and Communication System Workstations for the Evaluation of Hip Radiographs Following Arthroplasty

Sanjay R. Patel, MRCS, FRCR, Andoni P. Toms, FRCS, FRCR, Javed M. Rehman, FRCR, and James Wimhurst, FRCS(Orth)

*Investigation performed at the Norwich Radiology Academy and the Departments of Radiology and Orthopaedics, Norfolk and Norwich University NHS Foundation Trust, Norwich, United Kingdom*

**Background:** Conventional radiography is the primary imaging tool for routine follow-up of total hip replacements, but the reliability of this method has been questioned. The aim of this study was to assess the reliability of commonly used measurements of the position of hip prostheses on postoperative radiographs with use of tools available on all standard picture archiving and communication system workstations.

**Methods:** Fifty anteroposterior pelvic and lateral hip radiographs that were made after a unilateral total hip arthroplasty were included in this study. Acetabular inclination, lateral offset, lower-limb length, center of rotation, and femoral stem angle were independently assessed by two observers. Intraclass correlation coefficients were calculated for each measurement.

**Results:** The results demonstrated excellent reliability for acetabular angle ( $r = 0.95$ ), lower-limb length ( $r = 0.91$ ), and lateral offset ( $r = 0.95$ ) measurements and good reliability for center of rotation ( $r = 0.73$ ) and lateral femoral stem angle ( $r = 0.68$ ) measurements.

**Conclusions:** The position of total hip replacements can be reliably assessed with use of simple electronic tools and standard radiology workstations.

**Clinical Relevance:** Simple electronic tools on picture archiving and communication system workstations can be reliably used to measure alignment of total hip replacement prostheses at routine work stations. These measurements are reproducible and may avoid the need for expensive software and templates.

The postoperative radiograph is the mainstay for analyzing the alignment of total hip replacements in clinical practice. As radiography has moved from hardcopy film to digital techniques, the availability of templating programs and computer-aided techniques has increased both for preoperative planning<sup>1-8</sup> and intraoperative evaluation<sup>9-11</sup>. Several techniques and applications have been developed to correct geometric distortion in radiographs<sup>12-16</sup>. Roentgen stereophotogrammetric analysis is the most accurate technique available for measuring small amounts of movement in a prosthesis on serial radiographs,

but it is time-consuming and relatively expensive and not indicated for routine clinical care<sup>17-19</sup>. For many clinicians using picture archiving and communication systems, DICOM (Digital Imaging and Communications in Medicine) browsers, which provide a standardized image file format, have become a standard tool for viewing postoperative radiographs. These workstations typically include a set of simple electronic tools (such as window center and width, pan, zoom, and angle and distance calipers) and have no specific orthopaedic applications. Some of these have been evaluated for preoperative planning<sup>20</sup> but not, to

**Disclosure:** None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

our knowledge, as tools for examining radiographs after hip arthroplasty.

The aim of this study was to measure the reliability of simple generic tools in a picture archiving and communication system workstation for use in measuring the alignment and position of hip prostheses on postoperative radiographs.

### Materials and Methods

Fifty anteroposterior and lateral radiographs were selected sequentially from the picture archiving and communication system archive. This sample size of fifty was derived from a correlation coefficient sample size calculation assuming a type-I error ( $\alpha$ ) of 0.05, a type-II error ( $\beta$ ) of 0.1, and a correlation coefficient of 0.7, which is generally considered to be significant in biological systems. This required a sample size of thirty-four, and therefore fifty was selected to provide a safe margin.

The inclusion criterion was a unilateral total hip replacement with use of the same total hip prosthesis (Exeter V40; Stryker, Newbury, United Kingdom). All of the radiographs were made on a computerized radiography system (FCR XG5000, model CR-IR 362; Fujifilm, Tokyo, Japan). The anteroposterior hip radiograph was made with the patient supine and was centered on the pubic symphysis with use of a film focus distance of 115 cm. The lateral radiographs were made with the patient supine with the contralateral hip flexed and externally rotated with the x-ray beam angled at 45° inferomedial to superolateral, through the hip joint. The first postoperative radiograph for each patient was selected. Inclusion criteria were that each radiograph had to be centered, straight (equal sized obturator foramina), and include the proximal one-third of the femora<sup>21</sup>.

The measurements performed on the anteroposterior radiograph were the acetabular angle, lower limb length, lateral offset, center of rotation, and

varus or valgus stem angle, and the measurement performed on the lateral radiograph was the anteroposterior femoral stem angle (femoral anteversion). The measurements were done on diagnostic Picture Archiving and Communication System (PACS) workstations (version 2.1.2.1, Centricity PACS; GE Healthcare Systems, Chalfont St. Giles, United Kingdom) with use of high-resolution monitors (three megapixels, MFGD 3420; Barco, Kortrijk, Belgium). Each measurement was performed by two independent observers (a consultant and a radiology trainee) who were blinded to the other measurements until all data were collated. The technique for each measurement is described below.

The interteardrop line (A) was drawn by connecting the inferior margins of the right and left teardrops with the line measurement tool (Fig. 1). By drawing a second line (X) connecting the superolateral and inferomedial margins of the acetabular cup, the acetabular angle was derived from the resulting electronic goniometer angle (sometimes referred to as the Cobb angle tool on PACS<sup>22</sup>) (Fig. 1).

To assess the effect of the hip replacement on lower limb length, a dotted line (B2) was drawn perpendicular to the interteardrop line (A) to the most prominent point of the lesser trochanter on the side of the prosthesis. This line (B2) can be kept perpendicular by using the angle tool, which gives a continuous measurement and allows the line to the lesser trochanter to be adjusted until it is 90° to the interteardrop line. The same line (B1) is drawn to the contralateral lesser trochanter, keeping the angle at 180° (parallel to line B2). The effect on lower limb length on the side of the total hip replacement was defined as the length of the line B1 subtracted from the length of the line B2 in millimeters (Fig. 1). For the lateral offset, lines B and B1 were drawn perpendicular to line A starting from both teardrops (Fig. 2). Lateral offset of the total hip replacement was defined as line C subtracted from line C1 (C1 was always the side of the total hip replacement) in millimeters.

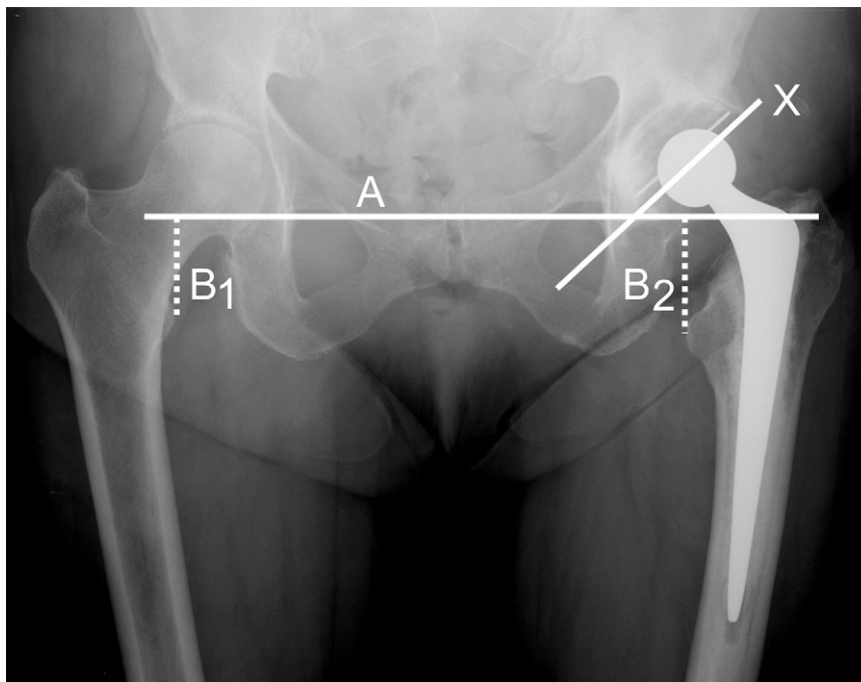


Fig. 1  
Anteroposterior radiograph of the pelvis demonstrating the technique for measurement of acetabular inclination and lower-limb length. The interteardrop line (A) is drawn between the inferior aspect of the medial acetabular teardrops. Acetabular inclination is measured as the angle subtended by a line connecting the superolateral and inferomedial corners of the acetabular implant (X) and the interteardrop line. Lower-limb length is measured as the length of a line (B1) drawn perpendicular from the interteardrop line (A) to the lesser trochanter of the uninvolved hip, subtracted from the same line (B2) drawn to the lesser trochanter of the involved hip.



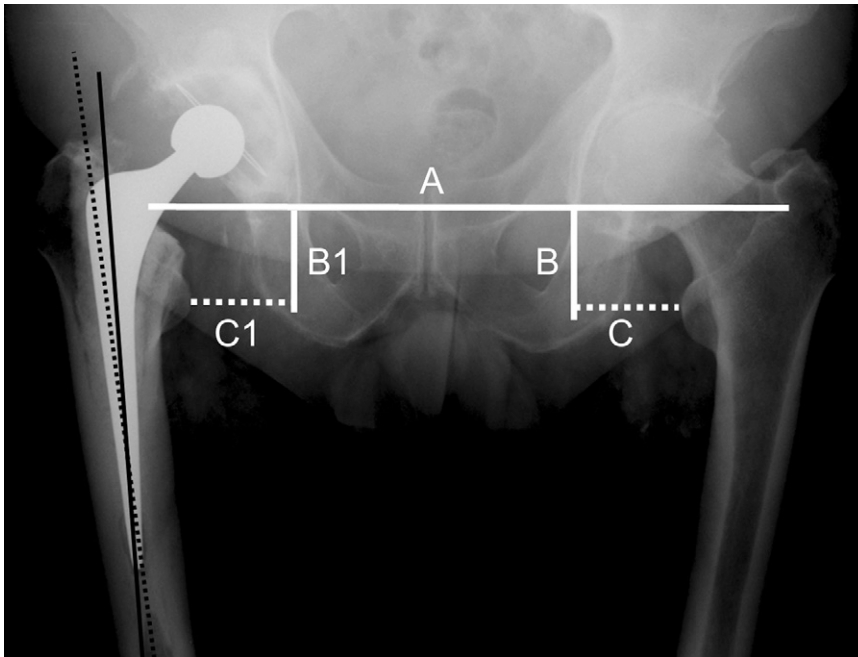


Fig. 2

Anteroposterior radiograph of the pelvis demonstrating the technique for measurement of lateral offset and anteroposterior stem angle. For lateral offset, lines (B and B1) were drawn perpendicular to the interteardrop line from each teardrop. A second line (C) was drawn parallel to the interteardrop line to the lesser trochanter of the uninvolved hip. A similar line (C1) was drawn to the lesser trochanter of the involved hip. The lateral offset was measured as the length of line C subtracted from line C1. Femoral stem angle was measured by obtaining an angle subtended by two lines: one drawn through the center of the shaft of the femoral component (solid black line) and a second drawn down the center of the medullary cavity of the proximal femoral diaphysis (dotted black line).

At this stage, all lines were deleted except the interteardrop line (A). The center of rotation was measured with use of the circular region-of-interest tool and the line tool. With use of the region-of-interest tool, a circle was expanded and fitted to the acetabulum. The corner markers of the region of interest were used to draw two perpendicular lines that bisected the circle and therefore defined the center (O). A third line (H) was drawn from the center of the circle (O) to the inferior margin of the teardrop, where an angle ( $\alpha$ ) was measured with use of the angle (Cobb) tool (Fig. 3). The x and y coordinates of the center of rotation could then be calculated with use of simple trigonometry from the angle  $\alpha$  and the hypotenuse H.

The femoral stem angle was measured on both the anteroposterior and lateral radiographs by drawing a dotted black line along the proximal femoral diaphysis equidistant to the cortex on either side. The angle subtended by a second line (the solid black line) drawn along the long axis of the femoral component was defined as the stem angle, and varus or valgus alignment or anteversion or retroversion was recorded (Fig. 2).

### Statistical Analysis

Descriptive statistics for each measurement were obtained and reliability was calculated with use of intraclass correlation coefficients<sup>23</sup>.

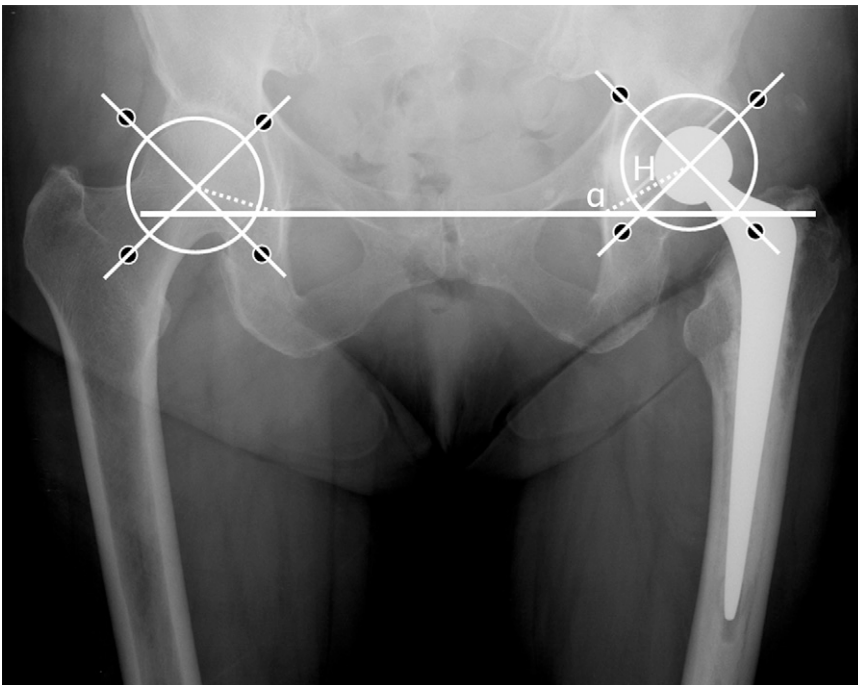


Fig. 3

Anteroposterior radiograph of the pelvis demonstrating the technique for measurement of the center of rotation. A circular region of interest was drawn to fit the femoral head and acetabular cup. The region-of-interest tool provides corner markers from which two lines are drawn that define the center of rotation. The length of a line (H) drawn from the inferior aspect of the acetabular teardrop to the center of rotation and the angle ( $\alpha$ ) subtended by the interteardrop line are recorded.

**TABLE I** Tabulated Summary of Results Demonstrating the Descriptive Statistics and Intraclass Correlation Coefficients for Each Measurement

Measurement	Mean (and Standard Deviation)	Mean Difference Between Observers (and Standard Deviation)	Intraclass Correlation Coefficient (R) (95% Confidence Intervals)
Acetabular inclination ( <i>deg</i> )	44 ± 7.1	1.8 ± 2.4	0.95 (0.92-0.97)
Lower-limb length ( <i>mm</i> )	3.9 ± 11.1	4.2 ± 4.8	0.91 (0.85-0.95)
Lateral offset ( <i>mm</i> )	-3.0 ± 13.2	3.6 ± 4.1	0.95 (0.92-0.97)
Center of rotation in distance ( <i>mm</i> )	44 ± 5.0	2.7 ± 2.5	0.86 (0.75-0.92)
Center of rotation in angle ( <i>deg</i> )	30.9 ± 5.3	4.6 ± 3.5	0.73 (0.53-0.85)
Anteroposterior stem angle ( <i>deg</i> )	-1.4 ± 2.0	1.1 ± 1.0	0.85 (0.74-0.91)
Lateral stem angle ( <i>deg</i> )	-1.5 ± 3.6	2.5 ± 3.2	0.68 (0.43-0.81)

**Source of Funding**

There was no external source of funding.

**Results**

The results are summarized in Table I. The mean acetabular angle (and standard deviation) measured  $44^\circ \pm 7.1^\circ$ . The intraclass correlation coefficient (*r*) was 0.95 (95% confidence interval (CI): 0.92 to 0.97) (Figs. 4 and 5). The mean effect on lower-limb length was  $3.9 \pm 11.1$  mm (*r* = 0.91; 95% CI: 0.85 to 0.95). The mean lateral offset measured  $-3.0 \pm 13.2$  mm (*r* = 0.95; 95% CI: 0.92 to 0.97). The center of rotation comprised two measurements: distance (from

the inferior portion of the teardrop) and angle (to the inter-teardrop line). The mean distance measured  $44 \pm 5.0$  mm (*r* = 0.86; 95% CI: 0.76 to 0.92), and the mean angle measured  $30.09^\circ \pm 5.3^\circ$  (*r* = 0.73; 95% CI: 0.53 to 0.85). The mean femoral stem angle on the anteroposterior radiographs measured  $-1.4^\circ \pm 2.0^\circ$  (*r* = 0.85; 95% CI: 0.74 to 0.91), and the mean femoral stem angulation on the lateral radiographs measured  $-1.5^\circ \pm 3.6^\circ$  (*r* = 0.68; 95% CI: 0.43 to 0.81) (Figs. 6 and 7).

The mean difference in the measurements between two observers for acetabular angle was  $1.8^\circ \pm 2.4^\circ$ . The mean lower-limb length had a  $4.2 \pm 4.8$ -mm variation in measurement. The

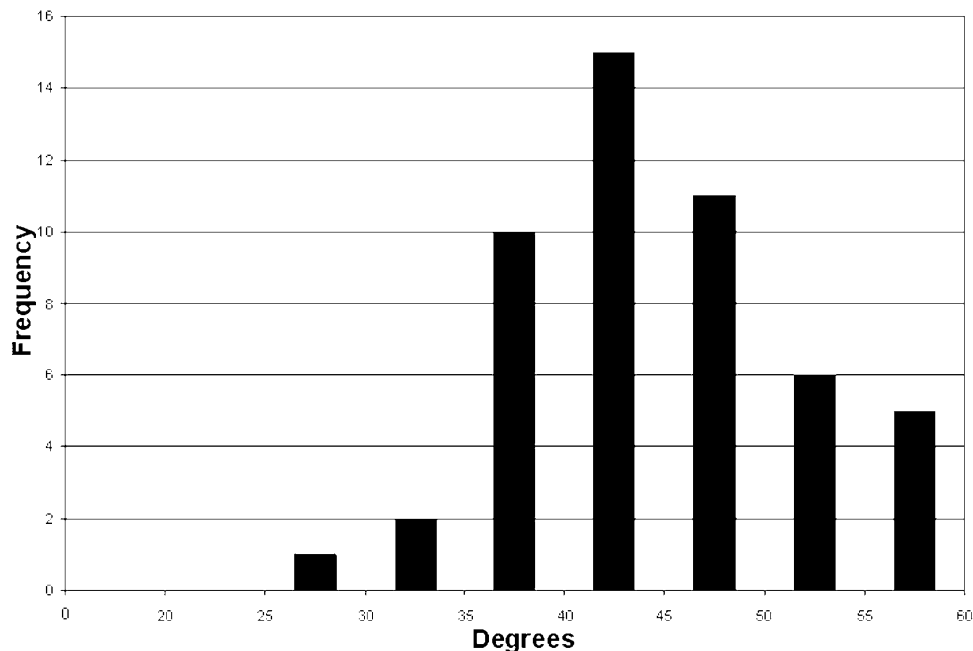
**Histogram of frequency of mean acetabular angle measurements**

Fig. 4

Frequency histogram demonstrating the normal distribution of acetabular inclination measurements.

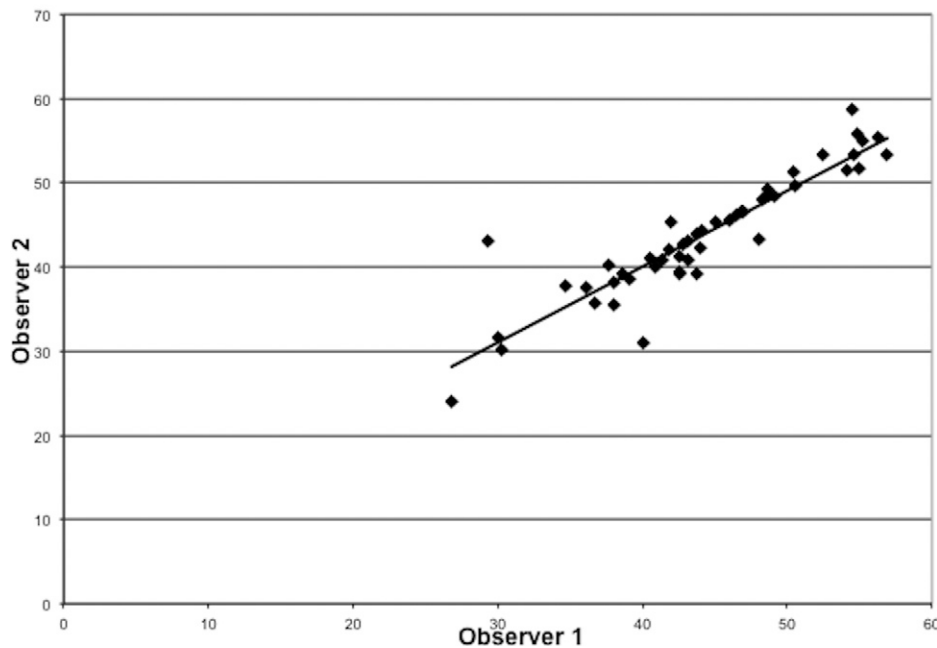
**Scatterplot of observations of acetabular angulation (degrees)**

Fig. 5  
Scatterplot demonstrating the strongest inter-rater reliability of all of the measurements: acetabular inclination ( $r = 0.95$ ).

mean difference for lateral offset was  $3.6 \pm 4.1$  mm. The inter-observer variations for center of rotation for distance and for angle were  $2.7 \pm 2.5$  mm and  $4.6^\circ \pm 3.5^\circ$ , respectively. The

femoral stem angle in anteroposterior and lateral radiographs showed a difference in measurement, with means of  $1.1^\circ \pm 1.0^\circ$  and  $2.5^\circ \pm 3.2^\circ$ , respectively.

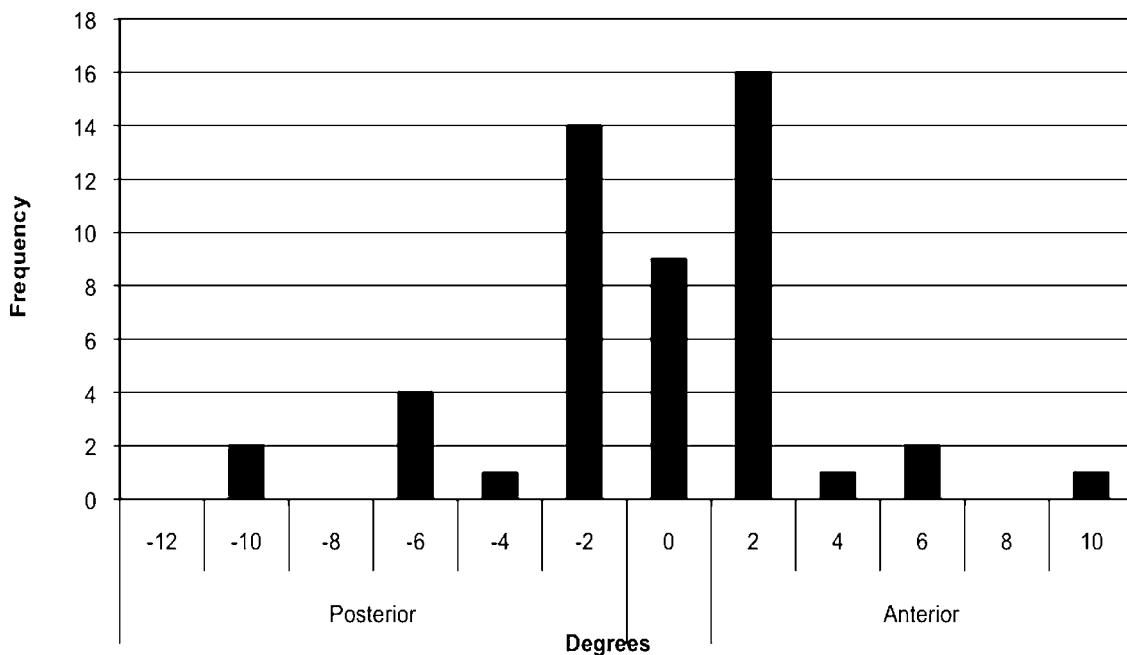
**Histogram of frequency measurements of lateral radiograph stem angle**

Fig. 6  
Frequency histogram demonstrating the distribution of femoral stem angle measurements on the lateral radiographs.

### Scatterplot of observations of lateral stem angle measurements (degrees)

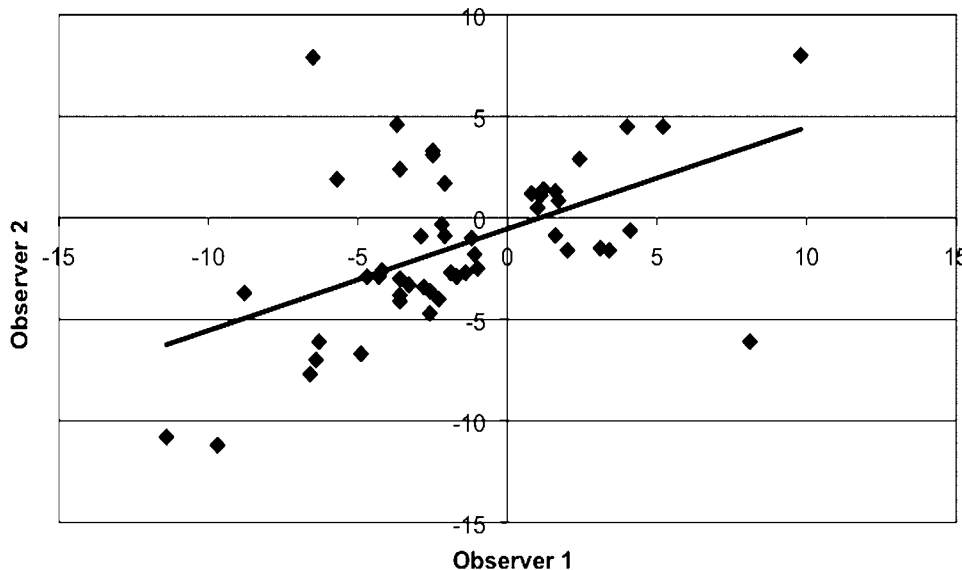


Fig. 7  
Scatterplot demonstrating the lowest interrater reliability of all of the measurements: femoral stem angle on the lateral radiograph ( $r = 0.68$ ).

## Discussion

Roentgen stereophotogrammetric analysis with tantalum bead bone-labeling is considered to be the gold standard for micromovement measurement of total hip replacements<sup>24,25</sup> and can demonstrate loosening within four months<sup>24,26,27</sup>. However, for many orthopaedic surgeons, the main radiographic tool for follow-up of total hip replacement remains the conventional postoperative radiograph, at least as a first-line investigation. The accuracy of this method does not compare with roentgen stereophotogrammetric analysis; however, it does provide the clinician with a first impression that there might be movement in an implant, and there is evidence to support this impression<sup>28-32</sup>. While there is variability in patient positioning between radiographs, our results indicate that the process of obtaining standard measures of hip position, from conventional radiographs on picture archiving and communication system workstations, is reliable.

Acetabular cup inclination is one of the most important parameters influencing long-term outcome of total hip replacements<sup>28,33-35</sup>. In our study, the mean acetabular cup inclination was  $44^\circ$ , which is within the commonly accepted normal range ( $33^\circ$  to  $50^\circ$ ), and therefore the data are comparable with those in previously published studies<sup>29,36-38</sup> (Figs. 4 and 6). Previous reliability studies of acetabular cup inclination, limited to the use of the Cohen kappa coefficient, in dysplastic hips and in femoroacetabular impingement have been disappointing<sup>39</sup>. It is currently considered that for continuous scale data, intraclass correlation coefficients are the better test of reliability<sup>40</sup>. In our study, the intraclass correlation coefficient for reliability of acetabular cup inclination was 0.95, which by any method of interpretation is excellent correlation<sup>41,42</sup>. Similar reliability results have been reported for measurements of vertical acetabular inclination (an intraclass correlation coefficient of 0.96) by two

orthopaedic surgeons<sup>43</sup>, suggesting that the results of our study are not anomalous.

The technique described in our study to measure lateral offset was devised to make use of the tools available on a picture archiving and communication system workstation, and these measurements appear to be reliable with excellent intraclass correlation coefficients. The mean difference between observers was 3.6 mm, which is a substantial improvement on that reported in previous studies, which noted a variation in lateral offset measurement ranging from +20 mm to -15 mm<sup>44</sup>.

The mean lower-limb-length discrepancy in our study was 3.9 mm, which is comparable with previous published data<sup>45,46</sup>. In our study, using picture archiving and communication system workstation tools, we demonstrated an excellent reliability with an intraclass correlation coefficient of 0.91, which is a substantial improvement on results from a previous study with a reported intraclass correlation coefficient of 0.62<sup>47</sup>.

A number of different methods have been described for measuring the center of rotation<sup>33,48</sup>. We used two parameters: first, the angle and, second, the distance of the center of rotation. To our knowledge, no previous studies have measured the reliability of center-of-rotation measurements. The results of our study demonstrate good reliability ( $r = 0.86$  and  $0.73$  for center of rotation in distance and in angle, respectively) but not the excellent reliability demonstrated with other measurements. The main reason for this is probably the increased number of steps required to obtain the measures, which means that variations between observers are additive at each step.

To our knowledge, the interobserver reliability of femoral stem-shaft angle measurements has been reported only once previously, in a study of resurfacing arthroplasties, in which a mean interobserver difference of  $1.49^\circ \pm 2.28^\circ$  (95% CI:  $-4.47^\circ$  to  $4.47^\circ$ ;  $r = 0.855$ ) was demonstrated but no reliability statistics

were performed<sup>49</sup>. These results were similar to those in our study, with a mean difference of 1.1° in anteroposterior and 2.5° in lateral radiographs (Figs. 5 and 7).

The main limitation of our study is that radiographs of only one type of prosthesis (Exeter) were used, and it may be that the results are not generalizable to all total hip replacements. A further limitation, of clinical application rather than study design, is that the interrater reliability is only one factor affecting variability on hip measurements on conventional radiographs. The most important application of the measurements is the changes over time, and these can be influenced by patient positioning and radiographic technique. However, this study does demonstrate that interobserver variation need not be a large component of this temporal variability. This study demonstrates good reliability, but not validity, of the measurements described.

Picture archiving and communication system workstations are now available in most hospitals in the Western world and have, as a standard feature, electronic calipers for measuring

length and angles on conventional radiographs. These simple electronic tools can be used reliably for measuring the postoperative position of total hip replacements in a serial manner and may avoid the need for expensive software and templates. ■

Sanjay R. Patel, MRCS, FRCR  
Andoni P. Toms, FRCS, FRCR  
Norwich Radiology Academy, Cotman Centre,  
Colney Lane, Norwich NR4 7UB, United Kingdom.  
E-mail address for S.R. Patel: sanjay25anusha@gmail.com

Javed M. Rehman, FRCR  
James Wimbush, FRCS(Orth)  
Departments of Radiology (J.M.R.) and Orthopaedics (J.W.),  
Norfolk and Norwich University NHS Foundation Trust,  
Colney Lane, Norwich NR4 7UB, United Kingdom

## References

- González Della Valle A, Comba F, Taveras N, Salvati EA. The utility and precision of analogue and digital preoperative planning for total hip arthroplasty. *Int Orthop*. 2008;32:289-94.
- Barrack RL, Burnett SJ. Preoperative planning for revision total hip arthroplasty. *J Bone Joint Surg Am*. 2005;87:2800-11.
- Viceconti M, Chiarini A, Testi D, Taddei F, Bordini B, Traina F, Toni A. New aspects and approaches in pre-operative planning of hip reconstruction: a computer simulation. *Langenbecks Arch Surg*. 2004;389:400-4.
- Viceconti M, Lattanzi R, Antonietti B, Paderni S, Olmi R, Sudanese A, Toni A. CT-based surgical planning software improves the accuracy of total hip replacement preoperative planning. *Med Eng Phys*. 2003;25:371-7.
- Lattanzi R, Viceconti M, Zannoni C, Quadrani P, Toni A. Hip-Op: an innovative software to plan total hip replacement surgery. *Med Inform Internet Med*. 2002;27:71-83.
- Lattanzi R, Baruffaldi F, Zannoni C, Viceconti M. Specialised CT scan protocols for 3-D pre-operative planning of total hip replacement. *Med Eng Phys*. 2004;26:237-45.
- Sugano N, Ohzono K, Nishii T, Haraguchi K, Sakai T, Ochi T. Computed-tomography-based computer preoperative planning for total hip arthroplasty. *Comput Aided Surg*. 1998;3:320-4.
- Hofmann AA, Bolognesi M, Lahav A, Kurtin S. Minimizing leg-length inequality in total hip arthroplasty: use of preoperative templating and an intraoperative x-ray. *Am J Orthop*. 2008;37:18-23.
- Woolson ST, Harris WH. A method of intraoperative limb length measurement in total hip arthroplasty. *Clin Orthop Relat Res*. 1985;194:207-10.
- Renkawitz T, Tingart M, Grifka J, Sendtner E, Kalteis T. Computer-assisted total hip arthroplasty: coding the next generation of navigation systems for orthopedic surgery. *Expert Rev Med Devices*. 2009;6:507-14.
- Jianguo Z, Ming Y, Qingming Y, Chengtao W. Computer-assisted femoral head resurfacing. *J Med Eng Technol*. 2009;33:9-17.
- Oddy MJ, Jones MJ, Pendegrass CJ, Pilling JR, Wimbush JA. Assessment of reproducibility and accuracy in templating hybrid total hip arthroplasty using digital radiographs. *J Bone Joint Surg Br*. 2006;88:581-5.
- Conn KS, Clarke MT, Hallett JP. A simple guide to determine the magnification of radiographs and to improve the accuracy of preoperative templating. *J Bone Joint Surg Br*. 2002;84:269-72.
- Bayne CO, Krosin M, Barber TC. Evaluation of the accuracy and use of x-ray markers in digital templating for total hip arthroplasty. *J Arthroplasty*. 2009;24:407-13.
- The B, Verdonchot N, van Horn JR, van Ooijen PM, Diercks RL. Digital versus analogue preoperative planning of total hip arthroplasties: a randomized clinical trial of 210 total hip arthroplasties. *J Arthroplasty*. 2007;22:866-70.
- White SP, Bainbridge J, Smith EJ. Assessment of magnification of digital pelvic radiographs in total hip arthroplasty using templating software. *Ann R Coll Surg Engl*. 2008;90:592-6.
- Decking J, Schuetz U, Decking R, Puhl W. The migration of femoral components after total hip replacement surgery: accuracy and precision of software-aided measurements. *Skeletal Radiol*. 2003;32:521-5.
- Ilchmann T. Radiographic assessment of cup migration and wear after hip replacement. *Acta Orthop Scand Suppl*. 1997;276:1-26.
- Onsten I, Berzins A, Shott S, Sumner DR. Accuracy and precision of radiostereometric analysis in the measurement of THR femoral component translations: human and canine in vitro models. *J Orthop Res*. 2001;19:1162-7.
- Hossain M, Lewis J, Sinha A. Digital pre-operative templating is more accurate in total hip replacement compared to analogue templating. *Eur J Orthop Surg Traumatol*. 2008;18:577-80.
- Massin P, Schmidt L, Engh CA. Evaluation of cementless acetabular component migration. An experimental study. *J Arthroplasty*. 1989;4:245-51.
- Gstoettner M, Sekyra K, Walochnik N, Winter P, Wachter R, Bach CM. Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. *Eur Spine J*. 2007;16:1587-92.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420-8.
- Mjöberg B. Loosening of the cemented hip prosthesis. The importance of heat injury. *Acta Orthop Scand Suppl*. 1986;221:1-40.
- Ryd L. Roentgen stereophotogrammetric analysis of prosthetic fixation in the hip and knee joint. *Clin Orthop Relat Res*. 1992;276:56-65.
- Derbyshire B, Prescott RJ, Porter ML. Notes on the use and interpretation of radiostereometric analysis. *Acta Orthop*. 2009;80:124-30.
- de Bruin PW, Kaptein BL, Stoel BC, Reiber JH, Rozing PM, Valstar ER. Image-based RSA: Roentgen stereophotogrammetric analysis based on 2D-3D image registration. *J Biomech*. 2008;41:155-64.
- Müller O, Reize P, Trappmann D, Wülker N. Measuring anatomical acetabular cup orientation with a new X-ray technique. *Comput Aided Surg*. 2006;11:69-75.
- Wan Z, Malik A, Jaramaz B, Chao L, Dorr LD. Imaging and navigation measurement of acetabular component position in THA. *Clin Orthop Relat Res*. 2008;467:32-42.
- Mayr E, Kessler O, Prassl A, Rachbauer F, Krismer M, Nogler M. The frontal pelvic plane provides a valid reference system for implantation of the acetabular cup: spatial orientation of the pelvis in different positions. *Acta Orthop*. 2005;76:848-53.
- Penney GP, Edwards PJ, Hipwell JH, Slomczykowski M, Revie I, Hawkes DJ. Postoperative calculation of acetabular cup position using 2-D-3-D registration. *IEEE Trans Biomed Eng*. 2007;54:1342-8.
- Cahir JG, Toms AP. (iii) CT and MRI of hip replacements. *Orthop Trauma*. 2009;23:101-8.
- Wan Z, Boutary M, Dorr LD. The influence of acetabular component position on wear in total hip arthroplasty. *J Arthroplasty*. 2008;23:51-6.
- Hirakawa K, Mitsugi N, Koshino T, Saito T, Hirasawa Y, Kubo T. Effect of acetabular cup position and orientation in cemented total hip arthroplasty. *Clin Orthop Relat Res*. 2001;388:135-42.
- Little NJ, Busch CA, Gallagher JA, Rorabeck CH, Bourne RB. Acetabular polyethylene wear and acetabular inclination and femoral offset. *Clin Orthop Relat Res*. 2009;467:2895-900.

- 36.** Murray DW. The definition and measurement of acetabular orientation. *J Bone Joint Surg Br.* 1993;75:228-32.
- 37.** Yoon YS, Hodgson AJ, Tonetti J, Masri BA, Duncan CP. Resolving inconsistencies in defining the target orientation for the acetabular cup angles in total hip arthroplasty. *Clin Biomech (Bristol, Avon).* 2008;23:253-9.
- 38.** Sarmiento A, Latta LL. A radiographic review of 135 total hip Charnley arthroplasties followed between 15 and 35 years. *Acta Chir Orthop Traumatol Cech.* 2006;73:145-50.
- 39.** Clohisy JC, Carlisle JC, Trousdale R, Kim YJ, Beaule PE, Morgan P, Steger-May K, Schoenecker PL, Millis M. Radiographic evaluation of the hip has limited reliability. *Clin Orthop Relat Res.* 2009;467:666-75.
- 40.** Obuchowski NA, Lieber ML. Statistics and methodology. *Skeletal Radiol.* 2008;37:393-6.
- 41.** Altman DG. *Practical Statistics for Medical Research.* London: Chapman and Hall London; 1991.
- 42.** Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74.
- 43.** Park SW, Park JH, Han SB, Choi GW, Song DI, An ES. Are portable imaging intraoperative radiographs helpful for assessing adequate acetabular cup positioning in total hip arthroplasty? *J Korean Med Sci.* 2009;24:315-9.
- 44.** Krishnan SP, Zaghoul AM, Carrington RWJ, Jeffery RM, Garlick N. The effect of variation in medial offset correction on functional outcome after total hip arthroplasty—a pilot study. *J Orthop.* 2007;4:e9.
- 45.** Konyves A, Bannister GC. The importance of leg length discrepancy after total hip arthroplasty. *J Bone Joint Surg Br.* 2005;87:155-7.
- 46.** Ranawat CS. The pants too short, the leg too long! *Orthopedics.* 1999;22:845-6.
- 47.** Sayed-Noor AS, Hugo A, Sjöden GO, Wretenberg P. Leg length discrepancy in total hip arthroplasty: comparison of two methods of measurement. *Int Orthop.* 2009;33:1189-93.
- 48.** Yoder SA, Brand RA, Pedersen DR, O’Gorman TW. Total hip acetabular component position affects component loosening rates. *Clin Orthop Relat Res.* 1988;228:79-87.
- 49.** Davis E, Gallie P, Olsen M, Schemitsch E, Waddell J. The accuracy of radiographs in measuring femoral neck-shaft angulation and femoral component angulation following hip resurfacing arthroplasty. *J Bone Joint Surg Br.* 2009;91 (Suppl II):222.



## Original Paper

# Optimization of metal artefact reduction (MAR) sequences for MRI of total hip prostheses

A.P. Toms<sup>a,\*</sup>, C. Smith-Bateman<sup>a</sup>, P.N. Malcolm<sup>a</sup>, J. Cahir<sup>a</sup>, M. Graves<sup>b</sup><sup>a</sup> Department of Radiology, Norfolk & Norwich University Hospital Trust, Norwich, Norfolk NR4 7UY, UK<sup>b</sup> University Department of Radiology, Addenbrooke's Hospital, Cambridge, UK

## ARTICLE INFORMATION

## Article history:

Received 26 September 2009

Accepted 1 December 2009

**AIM:** To describe the relative contribution of matrix size and bandwidth to artefact reduction in order to define optimal sequence parameters for metal artefact reduction (MAR) sequences for MRI of total hip prostheses.

**METHODS AND MATERIALS:** A phantom was created using a Charnley total hip replacement. Mid-coronal T1-weighted (echo time 12 ms, repetition time 400 ms) images through the prosthesis were acquired with increasing bandwidths (150, 300, 454, 592, and 781 Hz/pixel) and increasing matrixes of 128, 256, 384, 512, 640, and 768 pixels square. Signal loss from the prosthesis and susceptibility artefact was segmented using an automated tool.

**RESULTS:** Over 90% of the achievable reduction in artefacts was obtained with matrixes of  $256 \times 256$  or greater and a receiver bandwidth of approximately 400 Hz/pixel or greater. Thereafter increasing the receiver bandwidth or matrix had little impact on reducing susceptibility artefacts. Increasing the bandwidth produced a relative fall in the signal-to-noise ratio (SNR) of between 49 and 56% for a given matrix, but, in practice, the image quality was still satisfactory even with the highest bandwidth and largest matrix sizes. The acquisition time increased linearly with increasing matrix parameters.

**CONCLUSION:** Over 90% of the achievable metal artefact reduction can be realized with mid-range matrixes and receiver bandwidths on a clinical 1.5 T system. The loss of SNR from increasing receiver bandwidth, is preferable to long acquisition times, and therefore, should be the main tool for reducing metal artefact.

© 2010 The Royal College of Radiologists. Published by Elsevier Ltd. All rights reserved.

## Introduction

Total hip replacements (THRs) are no longer a contra-indication to magnetic resonance imaging (MRI). Techniques for metal artefact reduction have been described and refined to the point where MRI of THRs is becoming routine in some radiology departments. The principal indication for

MRI is a clinical presentation of pain associated with a THR that either appears normal at radiography or has radiographic changes that do not account for the symptoms. In this setting MRI can be particularly useful for diagnosing infection, tendon avulsions, bursitis, metal hypersensitivity reactions, and early bone marrow changes.<sup>1–5</sup>

Artefacts from metal can be most readily reduced by increasing the amplitude of the frequency encoding (FE) gradient, so that it is as large as possible in comparison to the susceptibility-induced gradients produced in the tissue by the metal implant. Modified spin-echo sequences employing view-angle tilting techniques, sometimes called metal artefact reduction sequences (MARS), may also be

\* Guarantor and correspondent: Department of Radiology, Norfolk & Norwich University Hospital Trust, Colney Lane, Norwich, Norfolk NR4 7UY, UK. Tel.: +44 1603 286104; fax: +44 1603 286077.

E-mail address: [andoni.toms@nnuh.nhs.uk](mailto:andoni.toms@nnuh.nhs.uk) (A.P. Toms).



used, although they may not be available on all MRI machines and they introduce some blurring in the images.<sup>6,7</sup>

As mentioned above, artefacts are most easily reduced by increasing the amplitude of the FE gradient. The user does not generally have direct control over this amplitude, but it can be increased by either widening the receiver bandwidth or decreasing the pixel size.<sup>8–12</sup> The problem is that some of these parameters incur penalties that might adversely affect image quality. Increasing the matrix size, for a fixed field of view, increases the acquisition time, which might result in movement artefact, and increasing the receiver bandwidth decreases the signal-to-noise ratio (SNR). It is predictable that there will be a diminishing return in artefact suppression as these parameters are increased. Therefore, the aim of the present study is to measure the effect of increasing receiver bandwidth and matrix size on artefact suppression, to measure the interaction between these two parameters, and to assess the risks to image quality from this approach to metal artefact suppression.

## Materials and methods

The present study used a phantom, which by not using human participants did not require research ethics approval at our institution. A phantom was constructed by embedding the femoral component of a Charnley total hip prosthesis (De Puy International Ltd., Leeds, UK) in 8 kg of solid vegetable fat (Fig. 1). All images were acquired on a 1.5 T Siemens Magnetom Avanto (Siemens Medical Systems, Erlangen, Germany) using the following protocols. A single T1-weighted image [echo time (TE) 9–13 ms, repetition time (TR) 500 ms, field of view (FOV) 500 mm, section thickness 5 mm] through the mid-coronal plane of the prosthesis was acquired at bandwidths of 150, 300, 454, 592, and 781 Hz/pixel using each of the following matrix sizes:  $128 \times 128$ ,  $256 \times 256$ ,  $384 \times 384$ ,  $512 \times 512$ ,  $640 \times 640$ , and  $768 \times 768$  pixels.

The volume of artefact was measured using an automated segmentation tool (OsiriX v2.6).<sup>13</sup> A region of interest (ROI) measuring  $5 \text{ cm}^2$  was selected from an area representing the background fat signal unaffected by the susceptibility artefact. Two observers measured the ROI independently. The thresholds were defined as being three standard deviations above and below the mean pixel value for the ROI (Fig. 2). All pixel values segmented outside this range were considered to be a combination of signal void and susceptibility artefact from the prosthesis. The area of artefact was calculated by subtracting the cross-sectional area of normal background signal from the total surface area of the section through the phantom. Inter-observer reliability for the ROI measurements was calculated using intra-class correlations (ICC).<sup>14</sup> SNR was calculated by dividing the mean signal value for the same ROI (taken from the fat of the phantom) by the standard deviation of signal values measured from an ROI (approximately  $200 \text{ cm}^2$ ) obtained from the signal void of

the surrounding air. The acquisition time for each image was recorded.

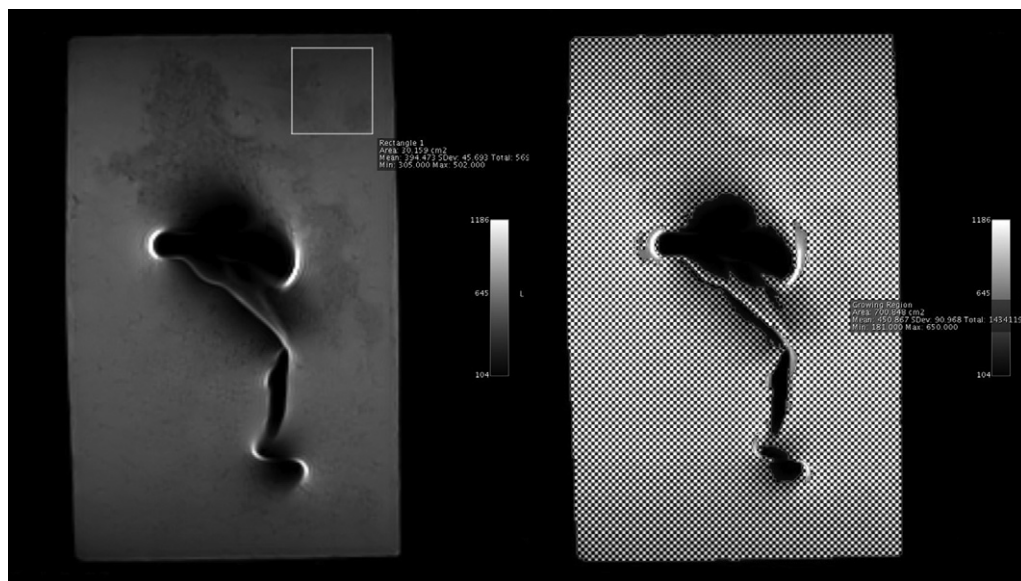
## Results

The mean background signal intensity varied from 460 to 524 (mean SD = 21.6). The ICC for the two observers was  $r = 0.96$  (95% confidence intervals 0.92–0.98, SPSS version 14.0 SPSS Inc, Chicago, IL, USA). There was a clear relationship between the receiver bandwidth with the largest increases in signal intensity occurring between 150 and 300 Hz/pixel (Table 1). With a matrix of  $128 \times 128$  increasing the bandwidth produced a progressive reduction in the area of the susceptibility artefact (Fig. 3). With larger matrix sizes, some early reduction in the artefact was demonstrated with the change from 150 to 300 Hz/pixel, but thereafter improvement in the artefact was equivocal. With a matrix of  $256 \times 256$  or larger any increases in bandwidth above approximately 400 Hz/pixel had no effect of artefact reduction. SNR decreased progressively with increases in bandwidth, but the decrease was also related to matrix size. The largest falls in SNR (from 1548 down to 200) occurred as the matrix increased from  $128^2$  to  $384^2$ ; above a matrix of  $384^2$  pixels the SNR remained below 200 (Fig. 4). Increasing the matrix size from  $128^2$  to  $384$  resulted in 80% of the possible artefact reduction. Further increases



**Figure 1** Hip prosthesis phantom made from a Charnley femoral component set in 8 kg of solid vegetable fat.





**Figure 2** T1-weighted MRI image through the mid-coronal plane of a Charnley THR demonstrating the ROI measurement of mean background signal intensity (a) and subsequent segmentation (reticulated area) of normal background signal from susceptibility artefact (b).

in matrix had little effect on artefact reduction but incurred significant increases in acquisition time (Fig. 5).

## Discussion

Artefacts can be effectively reduced by increasing the amplitude of the FE gradient. For most MR operators this can be achieved by either widening the receiver bandwidth or decreasing the pixel size, i.e., increasing the resolution for a fixed FOV in the FE direction. The FE gradient amplitude  $G_{FE}$  is given by

$$G_{FE} = \frac{2 \cdot \pi \cdot BW}{\gamma \cdot \Delta x}$$

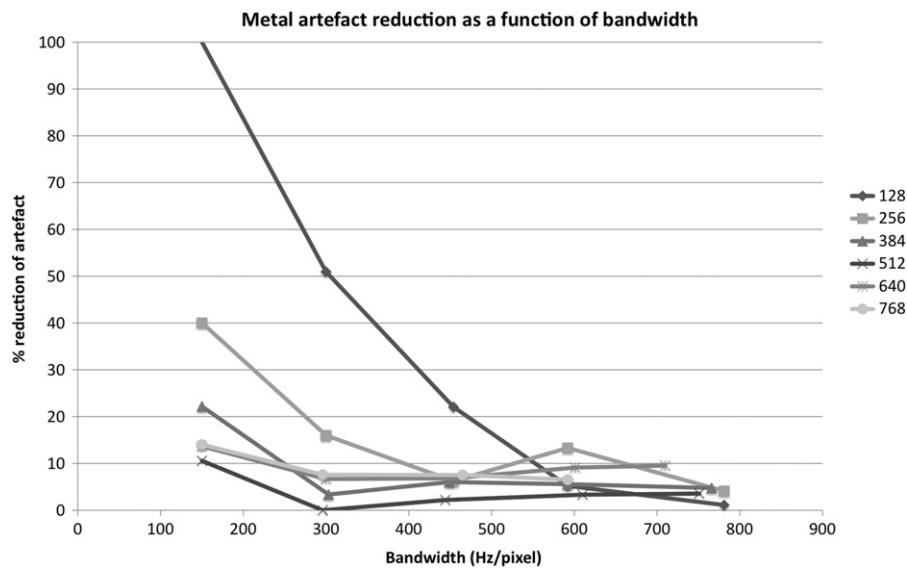
where BW is the bandwidth per pixel,  $\gamma$  is the gyro-magnetic ratio and  $\Delta x$  is the pixel size in the FE direction. Hence increasing the bandwidth (BW) or decreasing  $\Delta x$  will increase  $G_{FE}$ . The consequence of increasing the amplitude by either widening the bandwidth or decreasing the pixel size is a reduction in image SNR. For example, doubling the bandwidth, e.g., from 122 to

244 Hz/pixel, or halving  $\Delta x$ , e.g., by changing the FE resolution from 256 to 512, without changing the FOV, will reduce the SNR by  $1/\sqrt{2}$ , i.e. a reduction of 29%. Similarly, a factor of four change in either parameter will reduce the SNR by  $1/\sqrt{4}$ , i.e., a reduction of 50%. Although both methods have the advantage of increasing  $G_{FE}$ , it should be noted that if the FE resolution increases, the bandwidth should not be permitted to decrease. Depending on the system's vendor, the bandwidth can be set directly (in Hz/pixel), but other systems require specification of the full receiver bandwidth, e.g.,  $\pm 15.6$  kHz and the bandwidth has to be calculated manually, i.e., if the FE resolution were 256 then the bandwidth would be  $2 \times 15,600/256 = 122$  Hz/pixel. Note that if  $\Delta x$  is simply halved by increasing the FE resolution to 512 and the full receiver bandwidth is not changed, then the bandwidth becomes  $2 \times 15,600/512 = 61$  Hz/pixel and there is no increase in the  $G_{FE}$ , as the  $\Delta x$  has halved and the bandwidth has halved. In this situation it is also necessary to double the full receiver bandwidth to  $\pm 31.25$  kHz to ensure that the bandwidth is

**Table 1**  
The reduction in metal artefact and SNR with increasing bandwidth and matrix size.

Receiver bandwidth (Hz/pixel)	Matrix											
	128 × 128		256 × 256		384 × 384		512 × 512		640 × 640		768 × 768	
	Percent artefact	SNR	Percent artefact	SNR	Percent artefact	SNR	Percent artefact	SNR	Percent artefact	SNR	Percent artefact	SNR
150	100	1	41	0.35	20	0.19	2	0.12	0	0.09	0	0.07
300	66	0.71	27	0.25	15	0.14	10	0.09	10	0.06	7	0.05
454	49	0.58	24	0.2	15	0.11	7	0.07	10	0.05	7	0.04
592	29	0.5	22	0.18	12	0.1	7	0.06	10	0.04	10	0.03
781	22	0.44	20	0.16	15	0.08	7	0.06	12	0.04		
Time (s)	36		71		104		139		173		207	

SNR, signal-to-noise ratio.

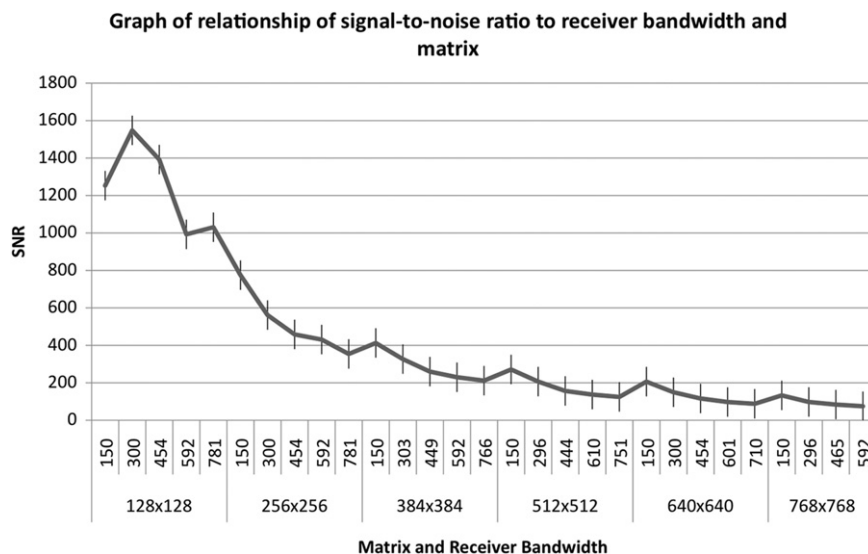


**Figure 3** Chart demonstrating reduction in metal artefact as a function of bandwidth for fixed matrixes.

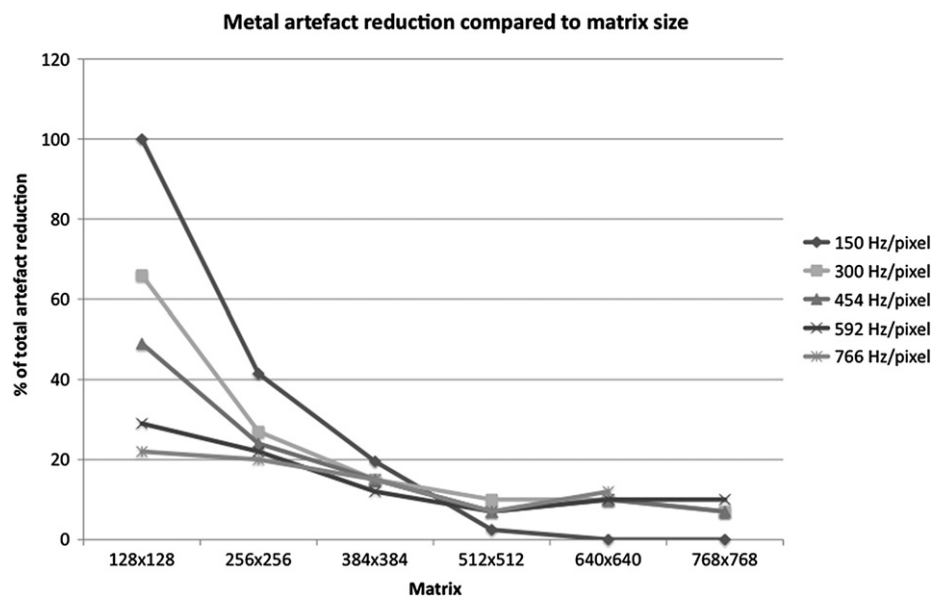
maintained at 122 Hz/pixel. However, if the full receiver bandwidth is doubled without changing the FE resolution, then the  $G_{FE}$  would double. The most important fact to remember is that if the FE resolution is increased, the BW must not change, or if the Dx is fixed, the bandwidth must be increased to increase the  $G_{FE}$ .

The results of the present study indicate that although increasing matrix size and receiver bandwidth will both reduce the susceptibility artefact, there is a diminishing return in artefact reduction as both of these parameters are increased. Most of the achievable metal artefact reduction (80–90%) occurs with bandwidths and matrixes somewhere in the middle of the available ranges for this standard clinical 1.5 T MRI machine, e.g., matrix  $300^2$  pixels and receiver bandwidth of 450 Hz/pixel. Increasing the parameters further does not result in a significant improvement

in artefact. Increasing the receiver bandwidth still further results in a dramatic reduction in SNR, but still produces images that are perfectly acceptable for diagnostic studies (Fig. 6). Conversely, as the resolution of the phase-encoding direction of the matrix increases, it does so in proportion to the acquisition time. The largest matrix used took over 3 min to acquire a single section, although this does not really reflect clinical practice. Acquisition times would be significantly reduced by decreasing the matrix in the phase-encoding direction and acquiring multiple sections simultaneously. Therefore, because the loss of SNR with increasing bandwidth is acceptable, and increasing the bandwidth alone can result in over 90% of achievable artefact reduction (Fig. 3), this should be the primary tool for metal artefact reduction. It is not necessary to consider changing the matrix (or voxel size), with its significant time



**Figure 4** Chart demonstrating relationship of SNR to increasing receiver bandwidth for given matrix sizes.



**Figure 5** Chart demonstrating the relationship of metal artefact reduction to matrix size.

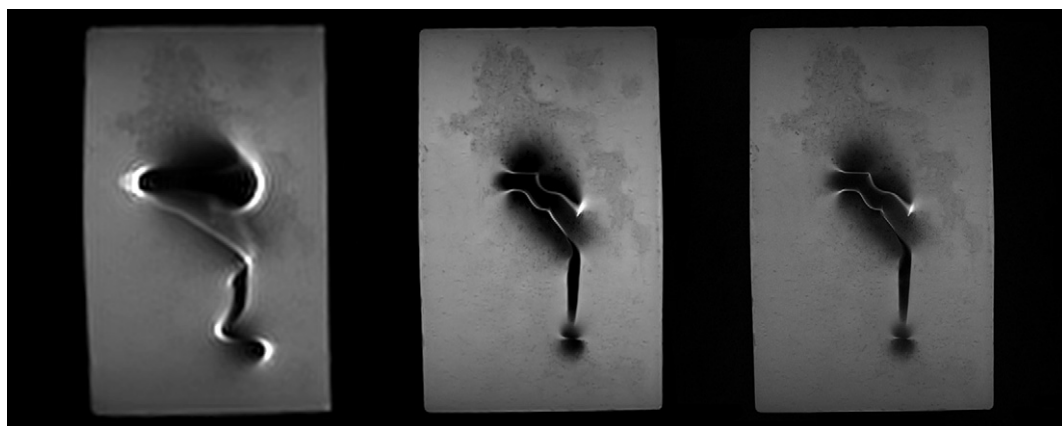
penalty, to control susceptibility artefact, but rather the matrix size should be determined by the resolution requirements of the study.

It has been previously suggested that decreasing the voxel size does not have a real effect on susceptibility artefact but simply improves the appearance of the artefact to the observer.<sup>15</sup> This study suggests otherwise; matrix alone can be used to reduce susceptibility artefact when all other parameters are fixed (Fig. 5). However, it is only by decreasing  $\Delta x$  in the FE direction that amplitude is increased and, therefore, the metal artefact is reduced. Increasing the matrix in the phase-encoding direction simply improves the spatial resolution of the image in that direction.

The method used to segment the artefact is a limitation of the study. Pixel values in MRI are somewhat arbitrary; they are not directly related to any inherent property of tissue and, therefore, will vary as the MRI parameters for

each sequence change. This is illustrated by the mean background pixel values obtained from the ROI, which varied by up to 12% with increasing bandwidth. To compensate for this a simple correction factor was applied to the thresholds used to segment each image. However, this correction factor was based on changes in mean signal intensity in the background fat and did not take account of the spread of data. This means that the correction factor is an approximation rather than exact correction. The final segmentation is probably a close estimate rather than a true measure.

In conclusion, over 90% of achievable metal artefact reduction can be achieved with receiver bandwidths and matrix sizes in the middle of the ranges available on a standard 1.5 T MRI machine. Receiver bandwidth can be used alone to achieve maximal metal artefact reduction with a loss in SNR that appears to be acceptable. Matrix size



**Figure 6** Three T1-weighted images of the phantom obtained at the following matrix sizes and receiver bandwidths: (a) matrix  $128^2$ , bandwidth 150 Hz/pixel; (b) matrix  $384^2$ , bandwidth 449 Hz/pixel; and (c) matrix  $768^2$ , bandwidth 781 Hz/pixel, demonstrating that most of the visible artefact reduction takes place with increases in the lower end of matrix sizes and receiver bandwidths.

should be determined by spatial resolution requirements alone and should not be considered necessary as a tool for metal artefact reduction.

## References

1. Cahir JG, Toms AP, Marshall TJ, et al. CT and MRI of hip arthroplasty. *Clin Radiol* 2007;**62**:1163–71. discussion 1172–1173.
2. Pfirrmann CWA, Notzli HP, Dora C, et al. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology* 2005;**235**:969–76.
3. Potter HG, Nestor BJ, Sofka CM, et al. Magnetic resonance imaging after total hip arthroplasty: evaluation of periprosthetic soft tissue. *J Bone Joint Surg Am* 2004;**86-A**:1947–54.
4. Toms AP, Marshall TJ, Cahir J, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008;**63**:49–58.
5. White LM, Kim JK, Mehta M, et al. Complications of total hip arthroplasty: MR imaging—initial experience. *Radiology* 2000;**215**:254–62.
6. Lee MJ, Janzen DL, Munk PL, et al. Quantitative assessment of an MR technique for reducing metal artifact: application to spin-echo imaging in a phantom. *Skeletal Radiol* 2001;**30**:398–401.
7. Kolind SH, MacKay AL, Munk PL, et al. Quantitative evaluation of metal artifact reduction techniques. *J Magn Reson Imaging* 2004;**20**:487–95.
8. Chang SD, Lee MJ, Munk PL, et al. MRI of spinal hardware: comparison of conventional T1-weighted sequence with a new metal artifact reduction sequence. *Skeletal Radiol* 2001;**30**:213–8.
9. Eustace S, Goldberg R, Williamson D, et al. MR imaging of soft tissues adjacent to orthopaedic hardware: techniques to minimize susceptibility artefact. *Clin Radiol* 1997;**52**:589–94.
10. Eustace S, Jara H, Goldberg R, et al. A comparison of conventional spin-echo and turbo spin-echo imaging of soft tissues adjacent to orthopedic hardware. *AJR Am J Roentgenol* 1998;**170**:455–8.
11. Olsen RV, Munk PL, Lee MJ, et al. Metal artifact reduction sequence: early clinical applications. *RadioGraphics* 2000;**20**:699–712.
12. Suh JS, Jeong EK, Shin KH, et al. Minimizing artifacts caused by metallic implants at MR imaging: experimental and clinical studies. *AJR Am J Roentgenol* 1998;**171**:1207–13.
13. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 2004;**17**:205–16.
14. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;**86**:420–8.
15. Runge VM, Nitz WR, Schmeets SH, et al. *The physics of clinical MR taught through images*. Thieme Publishing Group; 2005.

# Imaging the femoral sulcus with ultrasound, CT, and MRI: reliability and generalizability in patients with patellar instability

Andoni P. Toms · John Cahir · Louise Swift ·  
Simon T. Donell

Received: 14 November 2008 / Accepted: 23 December 2008 / Published online: 30 January 2009  
© ISS 2009

## Abstract

**Objective** Recent advances in surgical intervention for patellar instability have led to a need for long-term radiological monitoring. The aim of this study is to determine whether or not magnetic resonance imaging (MRI) or ultrasound (US) can replace computed tomography (CT) as the standard of care for the evaluation of the femoral sulcus.

**Materials and methods** This was a prospective study comparing the reliability of CT, magnetic resonance (MR), and US for measuring the femoral sulcus in patients with patellar instability. Twenty-four patients were recruited to undergo a CT, MR, and US examination of each knee. Two observers independently measured femoral sulcus angles from subchondral bone and hyaline cartilage on two occasions. Intraclass correlations and generalizability coefficients were calculated to measure the reliability of each of the techniques. Thereafter, two observers measured the femoral sulcus angle from ultrasound images recorded by two independent operators to estimate interobserver and interoperator reliability.

**Results** Forty-seven knees were examined with CT and US and 44 with MRI. The sulcus angle was consistently smaller when measured from subchondral bone compared to cartilage (5–7°). Interobserver reliability for CT, MR, and US measurements from subchondral bone were 0.87, 0.80, and 0.82 and from cartilage 0.80, 0.81, and 0.50. Generalizability coefficients of measurements from subchondral bone for CT, MR, and US were 0.87, 0.76, and 0.81 and for cartilage 0.76, 0.73, and 0.05. Most of the variability in the US occurred at image acquisition rather than measurement.

**Conclusion** In patients with patellar instability, CT and MR are reliable techniques for measuring the femoral sulcus angle but US, particularly of the articular cartilage, is not. MR is therefore the most suitable tool for longitudinal studies of the femoral sulcus.

**Keywords** Femoral sulcus · CT · MR · US · Reliability

## Introduction

Trochlear dysplasia is becoming increasingly recognized as an important finding in patients with patellofemoral pathology, especially patellar instability. In 1964, Brattstrom [1] described the groove and defined three types of dysplasia; hypoplasia of the medial condyle (most frequent), aplasia of the medial condyle, and total dysplasia with a flat or convex femoral trochlea. Maldague and Malghem [2] were the first to use a strict (overlapping posterior femoral condyles) lateral plain radiograph of the knee to detect an insufficient trochlear depth. Dejour et al. [3] used the lateral plain radiograph to classify the dysplasia into three types depending on where the groove line of the sulcus crossed the medial femoral condylar line and reached the lateral condylar line (Fig. 1).

A. P. Toms (✉) · J. Cahir  
Departments of Radiology,  
Norfolk and Norwich University Hospital,  
Colney Lane,  
Norwich, Norfolk NR4 7UY, UK  
e-mail: andoni.toms@nnuh.nhs.uk

S. T. Donell  
Institute of Orthopedics,  
Norfolk and Norwich University Hospital,  
Colney Lane,  
Norwich, Norfolk NR4 7UY, UK

L. Swift  
School of Medicine, Health Policy and Practice,  
University of East Anglia,  
Norwich, Norfolk NR4 7TJ, UK





**Fig. 1** Lateral plain radiograph of the knee demonstrating trochlear dysplasia. Note the groove line passes anterior to the extension of the anterior femoral cortical line

Traditionally, the femoral sulcus angle, as measured on the tangential patella radiograph, has been used to assess trochlear dysplasia. However, the measured severity of the dysplasia depends critically on the angle of knee flexion. In dysplastic trochleas, the lesser the knee flexion, the higher is the sulcus angle [4] (Fig. 2).

The cartilaginous trochlear groove appears to be normal morphologically at birth [5] and is determined genetically more by cartilage rather than bony morphology [6]. Most patients, with trochlear dysplasia, present in adolescence towards the end of their growth spurt. In some, there is a strong family history. The dysplasia occurs at the level of the distal femoral physis. Either the dysplasia appears and worsens leading to patellar instability, or maltracking of the patella allows overgrowth of the distal femoral physis, and therefore dysplasia develops. When abnormal patella tracking is corrected before puberty, a more normal-looking groove may develop with time [7, 8]. To determine whether the origin of the dysplasia is genetically driven or secondary to patellar maltracking requires a longitudinal study with sequential imaging of the distal femur. The imaging technique ideally needs to be reproducible, risk free, and cheap and requires a suitable radiological tool for serial imaging of the femoral sulcus over periods of up to 15 years. The “gold standard” for imaging the femoral sulcus has, for many years, been computed tomography (CT), which has largely replaced conventional radiography [9] although there have been, to our knowledge, no previous comparisons of either ultrasound (US) or magnetic resonance imaging (MRI) with CT.

CT has been described as an accurate and reliable tool for assessing femoral sulcus morphology [10–12]. It does however carry a small radiation burden. While the mean effective dose of this technique is small and targets a

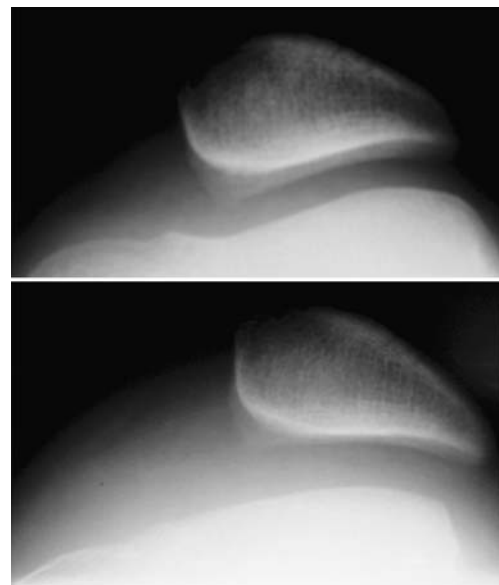
relatively radio-insensitive area, serial imaging with CT is undesirable if a reliable alternative can be found especially in children. US and MRI can both image the femoral sulcus without the need for ionizing radiation but neither technique has yet been fully validated for this application.

US of the developing sulcus has been described in children and adolescents [5] and has been validated in normal adults [13] using CT as the standard. US is an accurate method for measuring cartilage thickness in cadavers with normal sulcus morphology [14]. The use of US in patients with sulcal dysplasia has not, to our knowledge, been investigated. US would be the ideal tool for serial imaging of the femoral sulcus because it is cheap, widely available, and well tolerated by patients. There is, however, the question of whether the interoperator or intraoperator reliability in generating the necessary US images can match CT. MRI might be expected to be similarly reliable to CT although the improved conspicuity of the articular cartilage on both MRI and US might result in significantly different measures when compared to measurements taken from subchondral bone on CT.

The aim of this study was to determine whether or not US or MRI is sufficiently reliable to replace CT as alternative tools for imaging the femoral sulcus in subsequent longitudinal studies.

## Materials and methods

This was a prospective study of the reliability of US, magnetic resonance (MR), and CT to measure the femoral



**Fig. 2** Tangential patella (skyline) plain radiograph of trochlear dysplasia with the knee in 30° (a) and 20° (b) knee flexion (same patient as in Fig. 1)

sulcus angle in patients presenting with symptoms of patellar maltracking. Institutional Research Governance and Ethics approval was obtained prior to starting the study.

### Subjects

A total of 24 patients (14 females, ten males) were recruited from a single orthopedic outpatient clinic specializing in patellofemoral disorders. The inclusion criteria were patient presentations to the patellar instability clinic with a history of patellar dislocation or instability with clinical signs that include patellar apprehension, abnormal tilt, or patellar maltracking. Exclusion criteria included previous patellofemoral surgery and contraindications to MRI. All patients provided written informed consent. Ten patients presented with symptoms in both knees (eight females, two males). The 24 patients presented with patellar instability in a total of 30 knees in which there had been at least one dislocation in 19 knees. Two patients presented with anterior knee pain in four knees. The mean age at onset of symptoms was 19 years old (range 9 to 20), and the mean age at the time of imaging was 22 years old (range 12 to 44). Hypermobility syndrome was definitely present in six patients (two females, four males). Abnormal patellar tracking was found

clinically in 24 knees, including those whose principal presentation was anterior knee pain.

### Radiology

Patients attended the radiology department on two separate occasions. At the first visit, they underwent CT, US, and MRI examinations of the knee. At a second visit, not less than 2 weeks after the first visit, they were booked to attend for a further US examination.

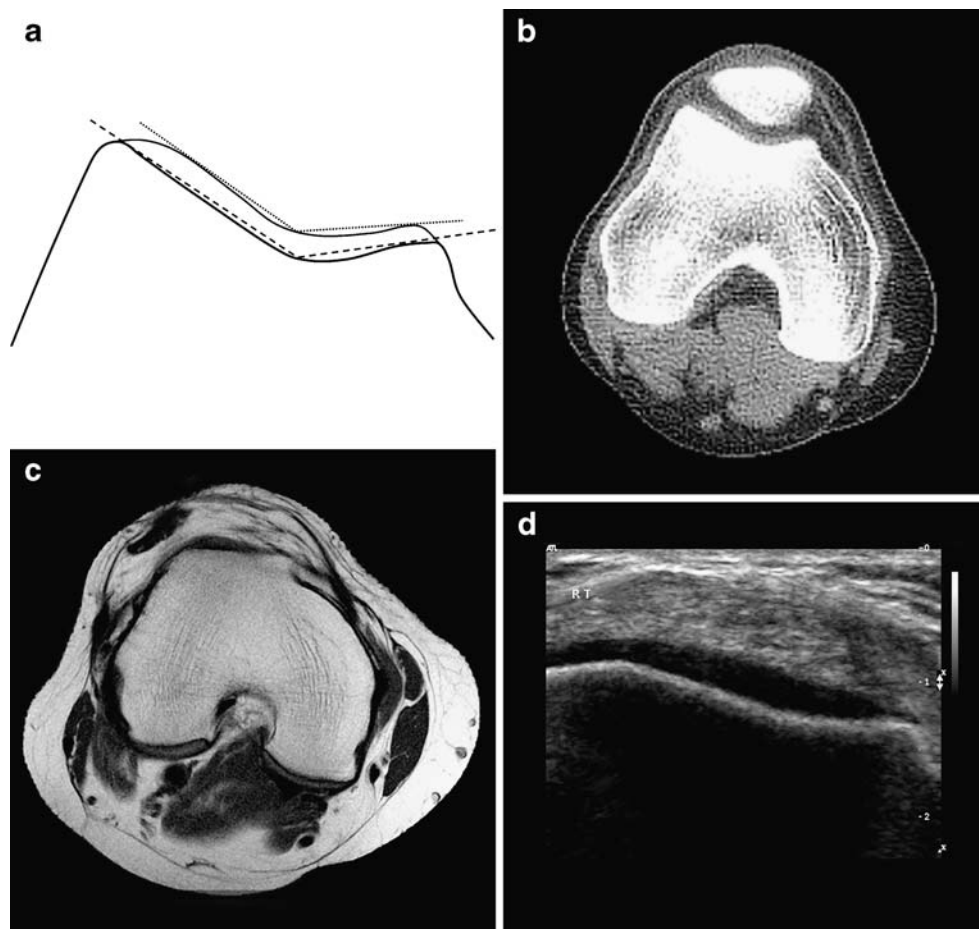
### CT

CT studies were performed on a GE lightspeed machine (General Electric, Milwaukee, WI, USA). Both knees were imaged together (standard of care) with 5-mm-thick contiguous axial images from the level of the midpoint of the higher of the two patellas to the level of the more distal tibial tuberosity. Data were reconstructed using a bone algorithm.

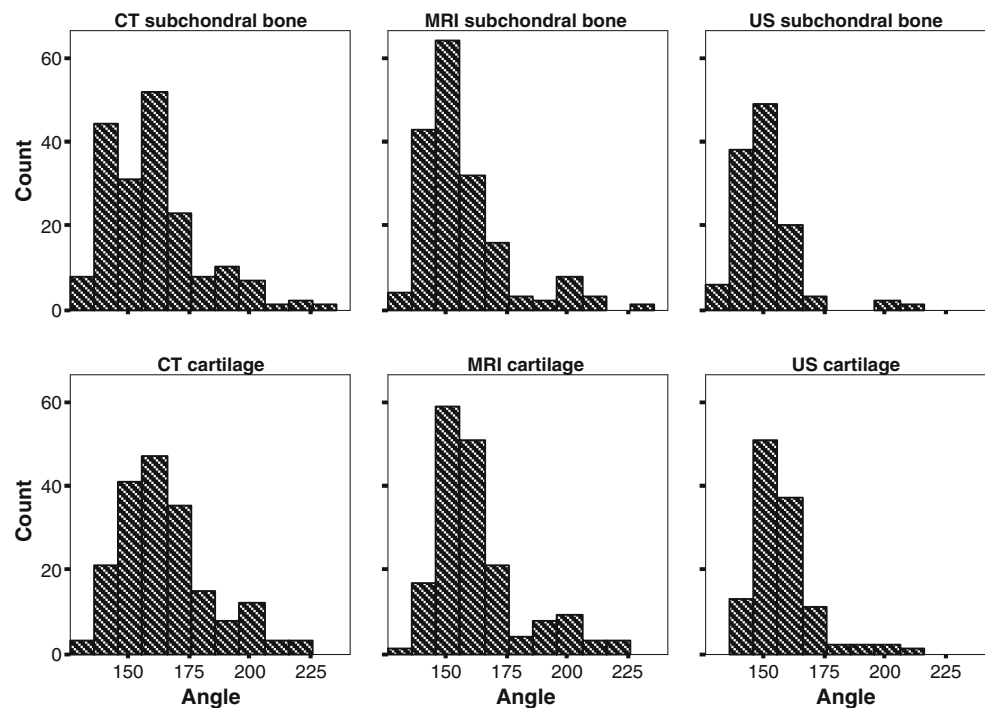
### Ultrasound

US examinations were performed with the knee flexed to 90°. The femoral sulcus was insonated anteriorly

**Fig. 3** **a** Diagrammatic representation of an axial view of the femoral sulcus indicating the three points identified for each measure of the sulcus angle: medial and lateral trochlea ridges and nadir of the sulcus. **b** Axial CT of the right knee at the level of the femoral epicondyles demonstrating a normal femoral sulcus. **c** Axial T2/PD (TE 45) weighted MR at the level of the femoral epicondyles demonstrating a hypoplastic femoral sulcus. **d** Axial US of the femoral sulcus of the right knee demonstrating a subchondral concavity but an articular cartilage convexity resulting in a femoral boss



**Fig. 4** Comparative histograms of the femoral sulcus angles measured with CT, MRI, and US



using a 5–12-MHz linear array transducer (ATL 5000 HDI, Wave Imaging Corp., Willoughby, OH, USA). Axial images of the sulcus were obtained at the level of the most anterior apex of the lateral femoral condyle [5] with the orientation of the transducer perpendicular, in two planes, to the long axis of the femur as judged visually by the operator. To measure interobserver variation of two operators, two musculoskeletal radiologists independently examined both knees at the first presentation. A third US examination (performed by one of two subspecialty musculoskeletal radiologists) was performed, no less than 2 weeks after the first examination, in order to assess intraobserver reliability. Each observer measured the femoral sulcus angle from all US images generated by the two operators.

## MR

In order to reduce bias, the MRI sequence was designed with equivalent matrix and slice thickness parameters to the CT protocol. Axial PD/T2 fast spin echo sequences (TE 45, TR 3000, FOV 19 cm, matrix 512×512, slice thickness 5 mm with no gap; Siemens Avanto 1.5 T, Siemens Medical Solutions) were acquired for each knee separately from midpoint of patella to tibial tuberosity.

## Sample size

By using the method for intraclass correlations (ICC) [15], a sample size of 43 knees was calculated to be sufficient to

provide a 95% confidence interval of half-width 0.1 for a classical ICC, assuming an ICC of 0.85.

## Measurements

All measurements were made by two independent observers reviewing images on 2K diagnostic workstations (GE Centricity, General Electric, and Barco) with frame setup for a single image per screen. For CT and MR, the sulcus angle was measured using an electronic protractor on the axial slice at the level where the femoral epicondyles were most widely separated on the medial to lateral axis. The femoral sulcus angle was measured from the deepest point of the sulcus to the most anterior portions of the medial and lateral trochlea margins (Fig. 3a). This was performed for the sulcus as defined by both the articular surface of the hyaline cartilage and the sulcus as defined by subchondral bone. All measurements were repeated by both observers at least 2 weeks after the first measurement.

**Table 1** Table summarizing the means and variances of femoral sulcus measurements

	CT		MRI		US	
	Bone	Cartilage	Bone	Cartilage	Bone	Cartilage
Mean	157.8	163.2	155.3	162.5	151.1	158.1
Variance	363.7	329.6	297.3	298.7	149.5	144.3
<i>n</i>	188	188	176	176	119	119



**Table 2** The table shows the estimated variance from each potential source of variability

	CT		MR		US	
	Bone	Cartilage	Bone	Cartilage	Bone	Cartilage
Components						
Patient	294.0	262.4	213.1	223.7	151.6	0
Knee	84.5	41.6	47.8	40.3	15.5	9.2
Time	0	0	0	0	0	0
Observer	0.6	0.3	1.0	0	0	0
Patient $\times$ time	0	2.6	14.6	24.6	3.7	75.3
Patient $\times$ observer	13.7	29.9	0	11.0	3.5	57.4
Knee $\times$ time	0	13.0	6.1	3.6	0	0
Knee $\times$ observer	6.3	1.0	42.7	23.0	0	10.7
Time $\times$ observer	0	0.14	0	0	0	5.6
Error	36.5	47.7	19.5	36.0	31.9	12.1
Coefficients						
Interobserver	0.87	0.80	0.82	0.81	0.83	0.50
Intraobserver	0.92	0.84	0.88	0.82	0.83	0.45
Generalizability coefficient	0.87	0.76	0.76	0.73	0.81	0.05

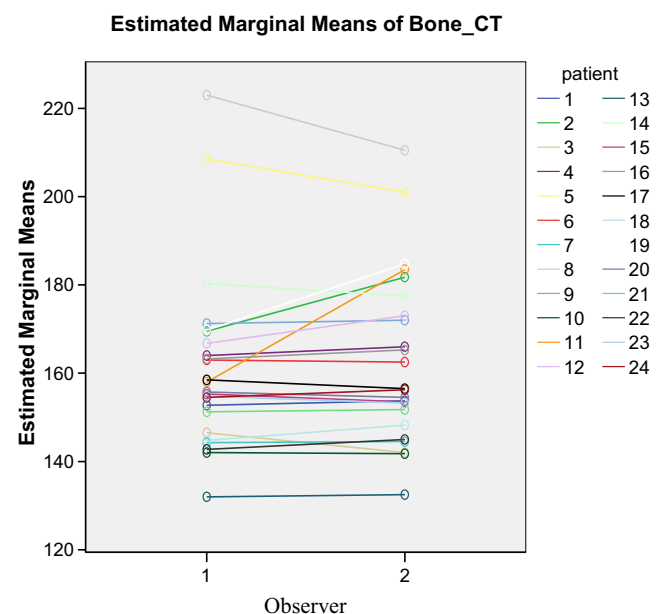
Variance in the measurements may be due to variability between observers (interobserver) and between observations by one observer (intraobserver). This variance may also be related to different femoral sulcus angles (i.e. specific to a “knee”) or, because two knees were examined in most patients, may be related to patient-specific factors (labeled “patient”). The variance due to combinations of these sources of variability has been calculated. For instance “Knee  $\times$  time” is a measure of how much variance can be attributed to repeated observations combined with the variance due to the shape of the knees themselves. Because the variance between the two knees of a single patient may be related, this is considered by examining “patient  $\times$  time” which is the variance specific to individual patients (or pairs of knees) between observations. The final three rows are calculated from the variance components and demonstrate the interobserver, intraobserver reliabilities and the generalizability coefficient, an overall measure of the reliability of any observer at any time

### Statistical analysis

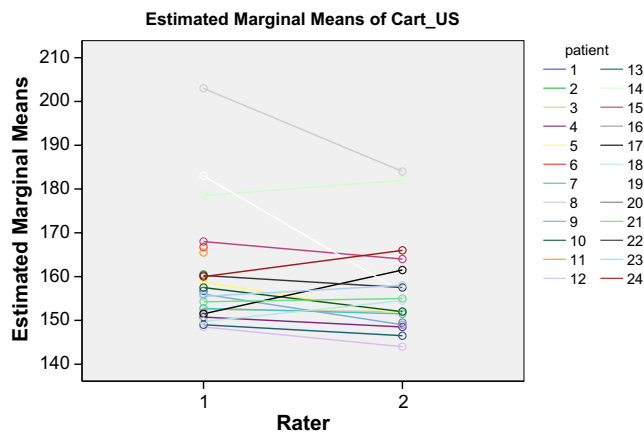
All analyses were performed using SPSS for windows version 14.0. The reliabilities of each of the three imaging techniques for measuring the sulcus angle from subchondral bone and articular cartilage were calculated. For each technique, the sulcus angles derived from subchondral bone and articular cartilage were compared. These were also compared to the CT from subchondral bone as the current gold standard. The femoral sulcus angles measured by ultrasound in the second study were further analyzed to determine how much variability arose from the operators obtaining the images compared to the observers measuring the angles alone.

Classical reliability theory breaks down variation in the data into a component due to subject, a component due to a single potential source of measurement error, for instance observer, and a residual or random error component and then calculates ICCs from these [16]. Our study, however, has two sources of measurement error, observer and time. With classical theory, the reliability of one source of error can only be calculated if the other source of error is held fixed. For instance, a separate estimate of intraobserver reliability would have to be calculated for each observer separately. Generalizability (GT) extends classical theory by allowing for more than one source of measurement error in the variance decomposition [15, 17]. Ratios of the

components can then be used to calculate interobserver and intraobserver reliabilities based on *all* the data and the generalizability coefficient, an overall estimate of the reliability of a measure at any time by any observer.



**Fig. 5** Graph illustrating reliability of femoral sulcus measurements from subchondral bone on CT. There is a clear patient effect for bone CT; the *lines* do not cross over much



**Fig. 6** Graph illustrating reliability of femoral sulcus measurements from articular cartilage on US. In comparison to Fig. 5, the lines cross much more often indicating increased variation between observers

All variance components were estimated using a mixed-model facility and restricted maximum likelihood (SPSS 14.0 for Windows). We allowed two-factor interactions in our model as well as main effects, i.e., we included a component for the difference between the difference in observers at first and second visit, as well as the main components for variance due to observers and variance due to time. Reliability coefficients were calculated so that they were analogous to correlations in that they measured absolute and not relative error [17].

## Results

### Descriptive statistics

Twenty-four subjects (12 males and 12 females, median age 17.5 range 10 to 41 years) and 47 knees were included in the study. One knee was excluded from the study because of a previous trochleoplasty at another hospital. All 47 knees had a CT. Out of 47 knees, 44 had an MR (three knees were not examined because of an oversight by the MR technicians). The planned three ultrasound examinations were performed on 27 knees, two on 18 knees, and only one ultrasound performed on the two knees of one patient.

The mean femoral sulcus angle measured on CT ( $n=188$ ) was  $157.8^\circ$  (SD  $19.1^\circ$ ) for subchondral bone and  $163.2^\circ$  (SD  $18.2^\circ$ ) for articular cartilage. The mean femoral sulcus angle measured on MR ( $n=176$ ) was  $155.3^\circ$  (SD  $17.2^\circ$ ) for subchondral bone and  $162.5^\circ$  (SD  $17.3^\circ$ ) for articular cartilage. The mean femoral sulcus angle measured on US ( $n=119$ ) was  $151.1^\circ$  (SD  $12.2^\circ$ ) for subchondral bone and  $158.1^\circ$  (SD  $12.0^\circ$ ) for articular cartilage (Fig. 4).

While the means and medians for all measures were broadly similar, US is markedly less variable for both bone and cartilage (Table 1, Fig. 4).

As a sensitivity analysis, classical ICCs [16] were used to estimate interobserver and intraobserver reliabilities for each time and each observer, respectively. As these ignore any possible correlation between knees on the same patient, we also calculated ICCs allowing separate components for patients and knee.

### Reliability

The most reliable measure of femoral sulcus angle comes from CT of subchondral bone with interobserver reliability of 0.87 and intraobserver reliability of 0.92 (Table 2; Fig. 5). The interobserver and intraobserver reliabilities were all above 0.8 except for the US measurement of the sulcus angle taken from hyaline cartilage (interobserver correlation coefficient of 0.50 and intraobserver correlation coefficient of 0.45; Fig. 6). The generalizability coefficient for the sulcus angle measured from subchondral bone on CT, MR, and US was 0.87, 0.76, and 0.81, respectively. The sulcus angle measurements taken from articular cartilage on CT and MRI had generalizability coefficients of 0.76 and 0.73, respectively.

The variance decomposition for sulcus angle measurements from US of cartilage (Table 2) suggests that patient effects vary from observer to observer and from time to time implying a very low generalizability coefficient. However, sensitivity analyses suggest that this result is dependent on the choice of model. For instance, including an additional variance component (subject  $\times$  observer  $\times$  time) gives a generalizability coefficient of 0.52. Further, when data from the first and second knees were analyzed separately, the generalizability coefficients were 0.64 and 0.39. The ranges

**Table 3** Table illustrating the range of observer correlation coefficients

	CT		MR		US	
	Bone	Cartilage	Bone	Cartilage	Bone	Cartilage
Interobserver	0.85–0.95	0.74–0.88	0.73–0.96	0.61–0.91	0.58–0.79	0.28–0.35
Intraobserver	0.68–0.93	0.73–0.86	0.75–0.88	0.68–0.89	0.62–0.91	0.23–0.65
Generalizability coefficient (knee 1, knee 2)	0.80, 0.91	0.72, 0.76	0.74, 0.71	0.70, 0.77	0.81, 0.63	0.64, 0.39

of interobserver and intraobserver reliabilities (Table 3) support the main results. As measurements on the knees of the same subject may be correlated, the ICCs produced are not independent of each other.

#### Comparison of bone and cartilage measurements

When sulcus angle was measured from articular cartilage, this produced a significantly higher mean angle measurement compared to the corresponding subchondral bone measurement for all techniques. The mean difference between subchondral bone and cartilage measurements was  $5.31^\circ$  for CT (95% CI 3.66–6.97),  $7.19^\circ$  for MR (95% CI 5.61–8.77), and  $6.92^\circ$  for US (95% CI 5.37–8.48). The corresponding Bland and Altman plots suggested that this bias did not differ for high/low measurements.

Classical ICCs between subchondral bone and cartilage measurements using the same technique and made by the same observer at the same time were 0.87 for CT and 0.84 for MRI but only 0.65 for US. Allowing for correlation between knees of the same subjects increased these slightly (0.88, 0.86, and 0.72). When scores can be adjusted by a fixed amount to allow for bias (SPSS's "consistent" ICCs), the corresponding ICCs were 0.87, 0.89, and 0.77. By using similar methods, both MRI measures are much more similar to CT measures from subchondral bone than are the US measures (Table 4).

#### US image and observer variability

There was more variability between the operators generating the US images than between observers measuring the femoral sulcus angle of those images. There is no evidence of a consistent observer effect; it varies from patient to patient (Table 5). The implied interobserver and interoperator reliabilities for subchondral bone measurements were both 0.77 whereas the corresponding figures for cartilage measurements were 0.77 and 0.75. Overall interobserver error (any image, any observer) was estimated at 0.72 for subchondral bone and 0.68 for cartilage measurements. Figures 7 and 8 illustrate the great variability in operator scores for each knee compared to interobserver differences. The few knees with larger discrepancies between observers appear to be those with higher scores.

#### Discussion

The mean sulcus angle for all three imaging techniques ranged from  $151^\circ$  to  $163^\circ$  with two standard deviations ranging from  $26^\circ$  to  $40^\circ$ . Previously described normal ranges of sulcus angles vary from  $140^\circ$  to  $143^\circ$  [11, 12, 18–20] with similar standard deviations to our population. The mean sulcus angles in patients with extensor mech-

**Table 4** Table illustrating the similarity of measures of the femoral sulcus angle with the current clinical standard: CT measurements from subchondral bone

	CT	MR		US	
	Cartilage	Bone	Cartilage	Bone	Cartilage
Classical ICC	0.87	0.81	0.71	0.39	0.32
ICC allowing for correlations between knees	0.88	0.88	0.79	0.61	0.62

anism malalignment or patellar maltracking have been reported to vary from  $150^\circ$  to  $161^\circ$  [12, 19, 21]. Therefore, the patient population in this study does appear to represent a cross section of patients with normal and dysplastic femoral sulci.

The results of this study suggest that the most reliable way to measure the femoral sulcus angle, in this population, is to use CT and to take measurements from subchondral bone. However, the femoral sulcus angle, defined by subchondral contours, can also be reliably measured with MR and US. When the sulcus angle measurements are taken from the surface of the articular cartilage, CT and MR are still reliable but US is not. It is not quite clear why this is the case.

There is a significant difference in the femoral sulcus measurements taken from subchondral bone when compared with those taken from articular cartilage. This on its own is not surprising. The depth of articular cartilage is greatest at the nadir of the sulcus and therefore the sulcus angle should be greater when measured from articular cartilage. This difference in measurements is least with CT (mean difference  $5^\circ$ ) and greater with US and MR (mean difference  $7^\circ$ ). This probably relates to the greater conspicuity of articular cartilage on US and MR compared with CT. These findings support evidence from cadaveric and arthrographic studies that suggest that articular cartilage morphology has a significant effect on trochlear morphology and that therefore it could be argued that measures

**Table 5** Table summarizing the variance components for ultrasound dataset

	Bone	Cartilage
Patient	137.5	109.2
Knee	6.5	26.0
Observer	0	0
Image	13.5	22.3
Patient $\times$ observer	10.4	14.6
Knee $\times$ observer	0	0
Observer $\times$ image	0.8	0.6
Error	32.1	26.3

should be based on cartilage rather than subchondral bone [22, 23]. While this may be true for future research, in clinical practice, a difference of  $5^\circ$  to  $7^\circ$  may not matter to the surgeon. This seems to be the case with MR and CT but not with US. All statistical methods demonstrated a low reliability for US measurement from cartilage but GT, in particular, suggested that the patient effect varies from time to time and observer to observer, implying that US is not at all useful for measuring the sulcus angle from cartilage.

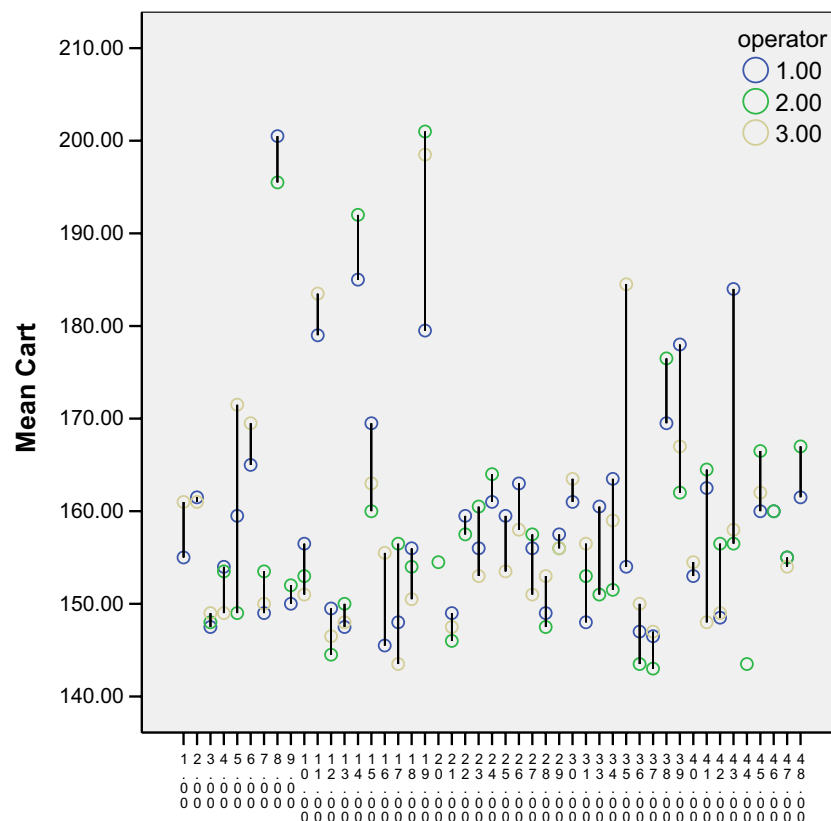
An explanation for the disparity in reliability between femoral sulcus angles measured with US from articular cartilage compared with subchondral bone might be the number of adolescents in the study group (12 between the ages of 10 and 17). It is possible that trochlear dysplasia in this age group predominantly affects the articular cartilage in such a way that the technique of standardizing the image acquisition, namely selecting a point level to the most anterior part of the lateral trochlear margin [24], is difficult to reproduce with ultrasound. Certainly, the variability appears to lie with the ultrasound examination rather than the measurements taken from the acquired images.

The finding that US is not reliable for measuring femoral sulcus angles is contrary to suggestions from previous studies

[5, 13]. However, where reliability studies have been performed, these have been in volunteers with normal trochlear morphology [13]. Once the sulcus becomes hypoplastic and flat or convex and even bossed, then the rules for localizing the ultrasound image are no longer applicable. There may be no lateral trochlear ridge. Figures 7 and 8 suggest an association of increased interobserver variability with trochlear dysplasia. It is possible that the reliability of US might be improved by changing the technique for localizing the optimal image. The focus of the trochlear dysplasia is at the level of the physis and this could be located by palpating the femoral epicondyles and drawing a line (perpendicular to the coronal plane) between the two and using this as the plane for the US image. Intuitively, this has more potential for variability than CT or MR.

MR does appear to be accurate at discriminating normal from dysplastic femoral sulci by defining dysplasia as a trochlear depth of more than 3 mm at a point 3 cm proximal to the joint line [25] but there is no evidence of the reliability of this technique.

Intraclass correlation coefficients [16] have become increasingly prevalent in the radiology literature and are the favored method for evaluating the reliability of single



**Fig. 7** Graph illustrating the association of observer variability (for three observations) with magnitude of femoral sulcus angle measures when measured from hyaline cartilage by US. The few knees with larger discrepancies between observers appear to be those with higher angle measurements

radiological observations by providing a measure of interobserver variability [26]. However, there are limitations to this statistical methodology when variability in the observation can be influenced by more than one factor. In this study, variability in the US measurements can be attributed to the patient, the US operator, and the observer. For assessing the reliability of more complex systems, generalizability theory can take account of these multiple factors and provide an overall reliability measure [15, 17].

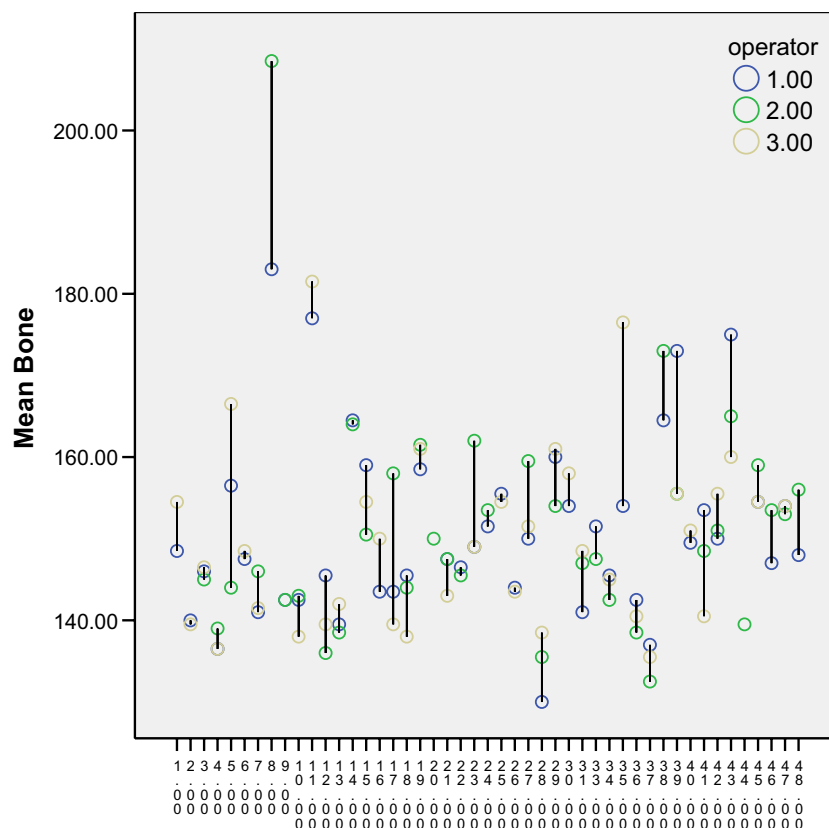
The sample size calculated for this study was sufficient to provide a 95% confidence interval of desired width for calculation of classical ICCs for interobserver or intraobserver reliability assuming that one knee is taken from each patient. This study achieved the desired sample size but two knees were used from each patient. Introducing such potential for within-patient correlation would increase the desired sample size for studies which estimate between-patient effect sizes. To examine the likely effect of this on the precision of the study, we can calculate sample size by treating the data as four highly correlated (maximum intraclass correlation 0.85) measurements on the same patient and in this case the sample size is enough to ensure a lower half-width of 0.11. In other words, our sampling method does not result in a huge loss in precision. Exact sample sizes for generalizability studies are not calculable but because sample size methods for studies of this

complexity are not available the use of generalizability theory, which makes use of all the data, should increase the precision of the study and offset any reduction due to the correlation between knees.

A further possible source of bias in the US measurements is that only 14 patients had all three US examinations. Ten patients did not attend for the second visit US examination (nine had two US examinations and one had a single US because the second operator was not available). Bias could be introduced if only those patients with definite trochlear dysplasia requiring intervention attended for the second visit thus skewing the population profiles for measures of interobserver and intraobserver variability. However, subset analysis of the mean sulcus angle for those who attended US once or twice revealed no significant difference between the two groups.

The CT and MR protocols employed in this study are somewhat archaic. Much more elegant examinations, with thinner slices and larger matrix sizes, can be performed with current technology. However, at the time of starting the study, our experience of imaging trochlear dysplasia was based on many years of the CT protocol as described and it was felt that we should test MR and US against the then current clinical standard. The MR sequence was tailored to eliminate any bias that might be attributable to differences in slice thickness or matrix size. It might be that with higher resolution MR

**Fig. 8** Graph illustrating the association of observer variability (for three observations) with magnitude of femoral sulcus angle measures when measured from subchondral bone by US. Again, the larger angle measurements are associated with more variability but the variability in the smaller angle measures is visibly less than the cartilage measurements





imaging of the femoral sulcus improved spatial resolution of the articular cartilage which could offer improvements in reliability of sulcal measurements.

For this study, the MR and CT were performed in full extension and the US examination was designed to try and produce a similar “axial” section through the sulcus. The measurable sulcus angle varies depending on the degree of knee flexion [12] and it has been suggested that optimal differentiation of the normal from abnormal sulcal morphology occurs with “axial” images with the knee in 10° of flexion [27]. It is possible that the position of the knee might have an effect on the reliability of measurements taken from subsequent images and therefore the results of this study are limited to the knee in extension.

## Conclusion

In patients with patellar instability, the reliability of US, particularly when measuring the femoral sulcus angle from articular cartilage, is poor. Therefore, US is probably not a suitable tool for longitudinal studies. Both CT and MR are reliable techniques for measuring the sulcus angle. Therefore, MR, without the radiation burden of CT, is probably the most suitable tool for performing repeated measurements during a longitudinal study of the developing femoral sulcus.

**Conflicts of interest** None.

**Acknowledgment** This study was supported by a research grant from Action Arthritis.

## References

- Brattstrom H. Shape of the intercondylar groove normally and in recurrent dislocation of patella: a clinical and X-ray-anatomical investigation. *Acta Orthop Scand* 1964;68: 1–148.
- Malghem J, Maldague B. Depth insufficiency of the proximal trochlear groove on lateral radiographs of the knee: relation to patellar dislocation. *Radiology* 1989;170: 507–510.
- Dejour H, Walch G, Neyret P, Adeleine P. Dysplasia of the femoral trochlea. *Rev Chir Orthop Reparatrice Appar Mot* 1990;76: 45–54.
- Davies AP, Bayer J, Owen-Johnson S, Shepstone L, Darrah C, Glasgow MM, et al. The optimum knee flexion angle for skyline radiography is thirty degrees. *Clin Orthop Relat Res* 2004;423: 166–171.
- Nietosvaara Y. The femoral sulcus in children. An ultrasonographic study. *J Bone Joint Surg Br* 1994;76: 807–809.
- Tardieu C, Glard Y, Garron E, et al. Relationship between formation of the femoral bicondylar angle and trochlear shape: independence of diaphyseal and epiphyseal growth. *Am J Phys Anthropol* 2006;130: 491–500.
- Beals RK, Buehler K. Treatment of patellofemoral instability in childhood with creation of a femoral sulcus. *J Pediatr Orthop* 1997;17: 516–519.
- Bonnard C, Nocquet P, Sollogoub I, Glorion B. Patellar instability in children. Result of transposition of the medial third of patellar tendon. *Rev Chir Orthop Reparatrice Appar Mot* 1990;76: 473–479.
- Laurin CA, Dussault R, Levesque HP. The tangential X-ray investigation of the patellofemoral joint: x-ray technique, diagnostic criteria and their interpretation. *Clin Orthop Relat Res* 1979;144: 16–26.
- Inoue M, Shino K, Hirose H, Horibe S, Ono K. Subluxation of the patella. Computed tomography analysis of patellofemoral congruence. *J Bone Joint Surg Am* 1988;70: 1331–1337.
- Martinez S, Korobkin M, Fondren FB, Hedlund LW, Goldner JL. Diagnosis of patellofemoral malalignment by computed tomography. *J Comput Assist Tomogr* 1983;7: 1050–1053.
- Biedert RM, Gruhl C. Axial computed tomography of the patellofemoral joint with and without quadriceps contraction. *Arch Orthop Trauma Surg* 1997;116: 77–82.
- Martino F, De Serio A, Macarini L, Rizzo A, Laforgia R, Rotondo A, et al. Ultrasonography versus computed tomography in evaluation of the femoral trochlear groove morphology: a pilot study on healthy, young volunteers. *Eur Radiol* 1998;8: 244–247.
- Modest VE, Murphy MC, Mann RW. Optical verification of a technique for in situ ultrasonic measurement of articular cartilage thickness. *J Biomech* 1989;22: 171–176.
- Streiner D, Norman G. Health measurement scales: a practical guide to their development and use. 3rd ed. New York: Oxford University Press; 2003, Chapter: Generalizability theory.
- Shrout E, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. *Psychol Bull* 1979;85: 420–428.
- Shavelson RJ, Webb NM, Rowley GL. Generalizability theory. *Am Psychol* 1989;44: 922–932.
- Alemparte J, Ekdahl M, Burnier L, Hernandez R, Cardemil A, Cielo R, et al. Patellofemoral evaluation with radiographs and computed tomography scans in 60 knees of asymptomatic subjects. *Arthroscopy* 2007;23: 170–177.
- Nietosvaara AY, Aalto KA. Ultrasonographic evaluation of patellar tracking in children. *Clin Orthop Relat Res* 1993;297: 62–64.
- Reikeras O, Hoiseth A. Patellofemoral relationships in normal subjects determined by computed tomography. *Skeletal Radiol* 1990;19: 591–592.
- McNally EG, Ostlere SJ, Pal C, Phillips A, Reid H, Dodd C. Assessment of patellar maltracking using combined static and dynamic MRI. *Eur Radiol* 2000;10: 1051–1055.
- Staubli HU, Durrenmatt U, Porcellini B, Rauschnig W. Anatomy and surface geometry of the patellofemoral joint in the axial plane. *J Bone Joint Surg Br* 1999;81: 452–458.
- Shih YF, Bull AM, Amis AA. The cartilaginous and osseous geometry of the femoral trochlear groove. *Knee Surg Sports Traumatol Arthrosc* 2004;12: 300–306.
- Hing CB, Shepstone L, Marshall T, Donell ST. A laterally positioned concave trochlear groove prevents patellar dislocation. *Clin Orthop Relat Res* 2006;447: 187–194.
- Pfirrmann CW, Zanetti M, Romero J, Hodler J. Femoral trochlear dysplasia: MR findings. *Radiology* 2000;216: 858–864.
- Obuchowski NA, Lieber ML. Statistics and methodology. *Skeletal Radiol* 2008;37: 393–396.
- Kujala UM, Osterman K, Kormanen M, Nelimarkka O, Hurme M, Taimela S. Patellofemoral relationships in recurrent patellar dislocation. *J Bone Joint Surg Br* 1989;71: 788–792.

# Grading the severity of soft tissue changes associated with metal-on-metal hip replacements: reliability of an MR grading system

Helen Anderson · Andoni Paul Toms · John G. Cahir ·  
Richard W. Goodwin · James Wimhurst · John F. Nolan

Received: 9 April 2010 / Revised: 4 June 2010 / Accepted: 4 July 2010 / Published online: 25 July 2010  
© ISS 2010

## Abstract

**Introduction** Metal-on-metal (MoM) soft tissue reactions or aseptic lymphocytic vasculitis-associated lesions (ALVAL) are being recognised using metal artefact reduction (MAR) MR with increasing frequency following the advent of second generation metal-on-metal bearings, but there is no standardised technique for reporting of MR appearances in this disease. The aim of this study was to measure the reliability of a grading system designed for scoring the severity of MoM disease on MRI.

**Materials and methods** MRI examinations of 73 hips in 59 patients were retrospectively selected and then anonymised, randomised and reviewed by three independent observers (musculoskeletal radiologists). Each MR examination was scored as either A: normal, B: infection, C1: mild MoM disease, C2: moderate MoM disease or C3: severe MoM disease according to pre-defined criteria. Kappa correlation statistics were used to compare the observations.

**Results** There was substantial agreement among all three observers; the correlation coefficient between the two most

experienced observers was  $\kappa=0.78$  [95% confidence intervals (CI): 0.68–0.88] and when compared with the least experienced observer coefficients were  $\kappa=0.69$  (95% CI: 0.57–0.80) and  $\kappa=0.66$  (95% CI: 0.54–0.78). The strongest correlation occurred for grades A, C2 and C3. The weakest correlations occurred for grades B and C1.

**Conclusion** The grading system described in this study is reliable for evaluating ALVAL in MoM prostheses using MR but is limited in differentiating mild disease from infection.

**Keywords** MR · Hip · Arthroplasty · Grade · Reliability

## Introduction

Modern metal-on-metal (MoM) hip replacements have been marketed as having a long life expectancy, with minimal wear over their lifetime, compared to traditional implants [1]. The most common type of total hip prosthesis comprises a metal head within a high density polyethylene acetabular cup (often backed by a metallic liner) whereas in MoM both sides of the articulation are metallic. They have been particularly targeted at a younger group of patients with the aim of reducing the number of operations they would require over their lifetime [1]. They also have the advantage of not generating polyethylene particle debris, which is a common cause of aseptic osteolysis in metal-on-polyethylene (MoP) bearings. While MoM bearings were some of the earliest designs for total hip arthroplasty, manufacturing tolerances at that time were not accurate enough and failure rates were unacceptably high [2]. However, some patients did not have implants that failed early and these often outlived MoP designs because the wear rates were considerably less [3–5].

H. Anderson · A. P. Toms · J. G. Cahir · R. W. Goodwin  
Department of Radiology,  
Norfolk and Norwich University Hospital,  
Colney Lane,  
Norwich, Norfolk NR4 7UY, UK

J. Wimhurst · J. F. Nolan  
Department of Orthopaedics,  
Norfolk and Norwich University Hospital,  
Colney Lane,  
Norwich, Norfolk NR4 7UY, UK

A. P. Toms (✉)  
Norwich Radiology Academy,  
Cotman Centre, Colney Lane,  
Norwich NR4 7UB, UK  
e-mail: andoni.toms@nnuh.nhs.uk

MoM bearings have been re-introduced to surgical practice because there has been an improvement in the understanding of bearing tribology (the study of friction in moving parts) and manufacturing techniques, but metallic ions generated by the corrosion or wear particles have long been a concern [6, 7]. There is evidence that the metallic ions can elicit an immunological response, and there is a theoretical link with oncogenesis [8]. Since 2005 there has been a steadily growing body of evidence of peri-prosthetic soft tissue complications linked to shedding of metallic particulate debris. Significant proportions of patients are presenting with complications that include perivascular lymphocytic infiltration, soft tissue necrosis and tendon rupture [9–11]. The exact mechanism of the disease process is not yet fully understood although it is thought it may be mediated by a type IV hypersensitivity reaction [9, 10].

In most cases the plain film appearances are normal and therefore MR, using metal artifact reduction (MAR) sequences, has emerged as the technique of choice for imaging these patients [12, 13]. Reported cases of MoM complications encountered at MR include fluid collections, periprosthetic soft tissue masses, proximal femoral bone marrow oedema, surrounding musculature and soft tissue oedema and necrosis, tendon avulsions and fractures [11, 14–18].

While MoM soft tissue disease has now been reported in a number of different hip replacements, there is no known correlation between the MR appearances and clinical, surgical or histopathological outcomes. Testing for such a correlation requires a reliable grading system for the MR findings of post operative MoM prostheses but there is, as yet, no published standardised method for reporting of these examinations.

The purpose of this paper is to present an MR grading system that can be used to score the severity of MoM disease and to measure the reliability of the scoring system.

## Materials and methods

This was a retrospective reliability study. A total of 73 MAR MRI examinations of hip replacements from 59 patients (40 women, age range 38–84 years and 19 men, 49–78 years) were selected for the study. The hip replace-

ments were implanted between 2002 and 2008. The implanted hips included in this study were hybrid (cemented stem, cementless cup) 28 mm articulation MoM ( $n=66$ ), and a variety of PoM hips (1 hybrid, 1 uncemented, 5 cemented). The MoP arthroplasties were included to ensure that there were adequate numbers of normal post-operative appearances and complications unrelated to the MoM bearings for the statistics. The observers were not able to identify which studies were MoM or MoP. All MR examinations were performed on a 1.5 T superconducting magnet (Siemens Magnetom Avanto, Siemens Medical Systems, Erlangen, Germany) with the parameters described in Table 1. The DICOM data for the MR examinations were anonymised using an in house anonymisation tool and labeled with a randomly assigned study number [19, 20]. Three musculoskeletal radiologists, blinded to the clinical history, type of prosthesis and radiographs, independently reviewed each MR examination. Two radiologists each had 5 years of experience at reporting MR of MoM hips. The third radiologist had little experience at reporting these examinations (less than 20 cases). All examinations were reviewed on a workstation running Osirix on a 24 inch monitor [21].

Each observer scored all of the MR examinations according to a grading system devised by the research team (Table 2). The categories were based on our local experience of how MR appeared to influence the management of patients with MoM disease. Categories A (normal; Fig. 1) and B (infection; Fig. 2) are generic and can be applied to any prosthesis. Category C was defined by features that were considered typical of MoM disease. Category C was sub-classified in to three further categories: mild (C1; Fig. 3), moderate (C2; Fig. 4) and severe (C3; Fig. 5), which correlate with our institution's decision to treat. For C1 (mild) this meant no immediate intervention but clinical review with a further MR in 1 year unless symptoms deteriorate. C2 (moderate) usually corresponds to a decision to revise the hip replacement electively and C3 (severe) is an indication to revise urgently.

Weighted kappa statistics were performed on the raters' scores to measure the reliability of this grading system (SPSS version 16.0) and interpreted using the method described by Landis and Koch (0.0–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 =

**Table 1** Table describing the MAR MR imaging parameters used in this study

	TE (ms)	TR (ms)	SL (mm)	FOV (cm)	Matrix	BW (Hz/pixel)
Coronal Turbo SE T1W	12	520	5	40×40	512×512	620
Coronal STIR	41	4,360	5	40×40	512×512	620
Axial Turbo SE T1W	12	540	6	30×30	512×512	620
Axial Turbo SE T2W	98	5,420	6	30×30	512×512	620
Sagittal Turbo SE T2W	99	6,120	5	40×40	512×512	620



**Table 2** Criteria for grading of MR examinations in patients with MoM hip replacements

Grade	Description	Criteria
A	Normal or acceptable	Normal post-op appearances including seromas and small haematomas
B	Infection	Fluid-filled cavity with high signal T2 wall; inflammatory changes in soft tissues; $\pm$ bone marrow oedema
C1	Mild MoM disease	Periprosthetic soft tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either less than 5 cm maximum diameter
C2	Moderate MoM disease	Peri-prosthetic soft tissue mass/fluid-filled cavity greater than 5 cm diameter or C1 lesion with either of following: (1) muscle atrophy or edema in any muscle other than short external rotators or (2) bone marrow edema: hyperintense on STIR
C3	Severe MoM disease	Any one of the following: (1) fluid-filled cavity extending through deep fasci, (2) a tendon avulsion, (3) intermediate T1W soft tissue cortical or marrow signal, (4) fracture

moderate agreement, 0.61–0.80 = substantial agreement, 0.81–1.00 = almost perfect agreement) [22].

## Results

There were a total of 219 scores from the three observers for the 73 hips. Of these, 28% of the scores were grade A, 7% grade B, 12% grade C1, 26% grade C2 and 27% were grade C3. In 65% of all observations, the MR examinations were graded as demonstrating soft tissue metal-on-metal disease.

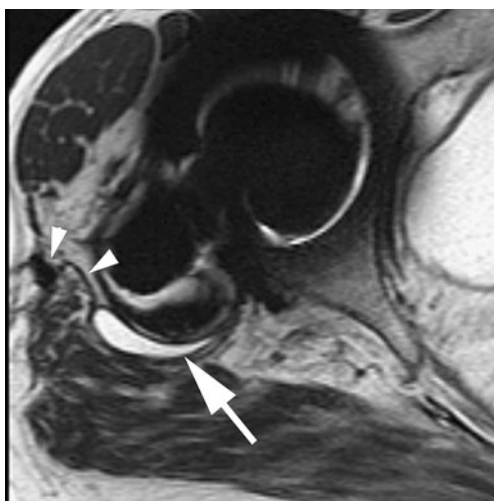
Weighted kappa correlation for the two more experienced musculoskeletal radiologists was  $\kappa=0.78$  (95% confidence intervals: 0.68–0.88). When the musculoskeletal radiologist with least experience of reporting MoM MR was compared with the more experienced two observers the kappa correlation coefficients were  $\kappa=0.69$  (95% confidence intervals: 0.57–0.80) and  $\kappa=0.66$  (95% confidence intervals: 0.54–0.78). Using the Landis and Koch method

for interpretation of kappa correlation statistics, these results indicate substantial agreement ( $\kappa=0.61$  to 0.80) among all three observers [22].

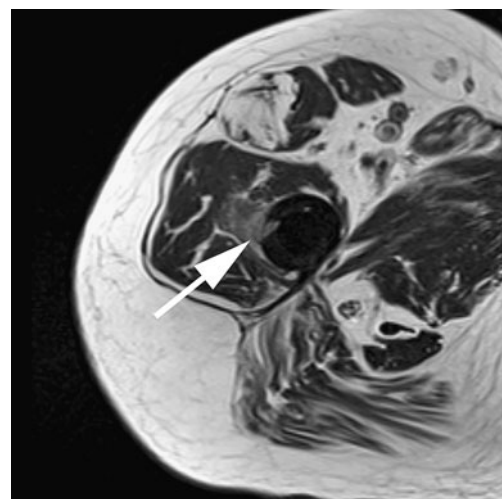
The strongest agreement appeared to be for the grade A, C2 and C3 categories and the most disagreement appeared to be for categories B and C1 (Table 3).

## Discussion

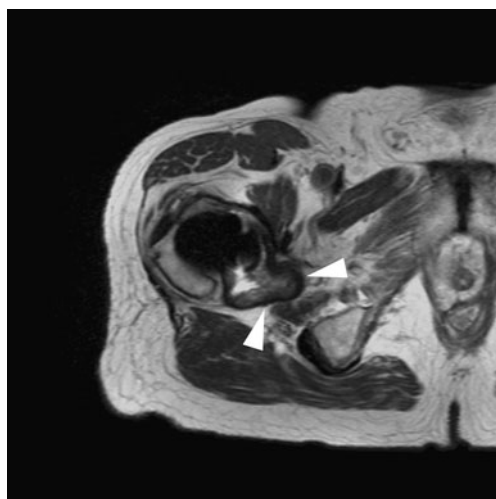
The results of this study suggest that the proposed MR grading system for MoM disease is reliable. The kappa correlation for all three radiologists can therefore be interpreted as demonstrating “substantial” agreement [22]. The weakest area of inter-rater reliability is in categories B and C1. This is not entirely surprising. One of the limitations of the dataset is the relatively small number of MR examinations of infected prostheses. This is not a common indication for MR in our institution and therefore the case material is limited to a handful of cases. Small volume peri-prosthetic soft tissue abnormalities in MoM



**Fig. 1** Grade A: axial T2W MR of a polyethylene-on-metal hip replacement demonstrating a thin walled fluid collection (arrow) in the plane of the surgical approach consistent with a small post-operative seroma

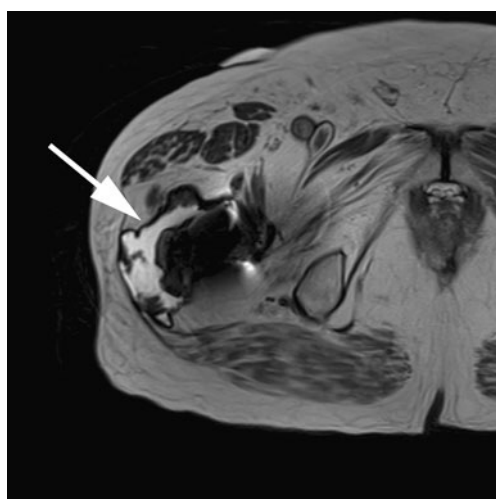


**Fig. 2** Grade B: axial T2W MR through the right femur demonstrating a cortical breach with inflammatory material extending in to the right vastus lateralis muscle

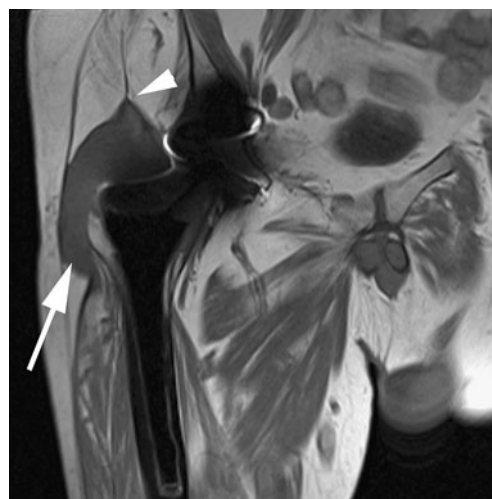


**Fig. 3** Grade C1: axial T2W MR at the level of the greater trochanter demonstrating a small fluid collection abutting the neck of the prosthesis medially (*arrowheads*). The thick ragged rim of the fluid sac is typical of a metal-on-metal soft tissue reaction. The low signal is caused by dephasing due to metal particles in the lining of the collection

disease can, in our experience, be non-specific in appearance with features that may overlap with infection or normal post-operative findings such as seromas. For practical purposes this is not a significant problem. Patients with mild MR grades of MoM disease (C1) do not, in our institution, require intervention but are followed up clinically and with serial MR. Those patients with infection do not have MR to make the diagnosis of infection but rather MR is typically limited to patients who are suspected of having infected prosthesis and are significant anaesthetic risks. These patients can sometimes be temporised by draining any localised abscess demonstrated on the MR, and treated with appropriate antibiotics. For the critical



**Fig. 4** Grade C2: axial T2W MR at the level of the greater trochanter demonstrating an MoM fluid collection (*arrow*) greater than 5 cm in diameter in contact with the femoral neck



**Fig. 5** Grade C3: coronal T1W MR through the right hip demonstrating a large fluid collection (*arrow*) with avulsion of the gluteus medius and minimus tendons (*arrowhead*) and atrophy of the muscle bellies

decisions that the MR is useful for, namely differentiating normal post-operative appearances from moderate or severe disease, the correlations between observers are strong and the grading system described in this paper is reliable.

While the grading system is reliable this does not mean that it is an accurate measure of disease. To the best of our knowledge there is, to date, no published correlation between the MR grade of MoM disease and either grading of operative findings, histological measures of severity or clinical scores. While many of our patients with moderate or severe disease have had their diagnoses confirmed with histological outcomes, these were not systematically correlated with MR findings in this study because the sole purpose of this study was to measure the reliability of the proposed grading system. A validation of the grading system is likely to take several years to complete because mild degrees of disease in asymptomatic patients do not warrant intervention and therefore there may not be any surgical or histopathological outcome data for this group. In

**Table 3** Table of proportion of absolute agreement in responses for experienced observers (1 and 2), and all three observers, demonstrating the limited agreement for grades B and C1

Grade	Frequency (%)	Complete agreement between observers	
		Observers 1 and 2 Proportion (%)	All 3 observers Proportion (%)
A	28	93	74
B	6	23	0
C1	12	33	11
C2	26	80	59
C3	28	77	63

these cases only stability in a longitudinal study will be a useful marker of the validity of mild disease grades. However, once validated, MR may then become a useful prognostic tool which may determine management in this group of patients.

A limitation of the study is that the database was constructed retrospectively and therefore the observers are likely to have seen some of the cases before. However the MR cases were selected from over 250 MR examinations performed over 5 years. They were anonymised and reported in isolation from any other clinical or radiological data. The reliability score for the least experienced observer, who had seen very few MoM MR cases, was equivalent to the more experienced observers. This suggests that previous exposure to the cases by the more experienced observers did not bias the scores in their favour and that the grading system described in this paper allows consistent evaluation of MoM prostheses with MR.

## Conclusion

MoM disease, demonstrated on MAR MR, can be reliably evaluated using the grading described in this study. Differentiation of normal hips and moderate or severe MoM disease is particularly robust, but it is more difficult to differentiate mild MoM disease from infection or some normal post-operative findings.

**Conflict of Interest** None

## References

1. Park Y, Moon Y, Lim S, Yang J, Ahn G, Choi Y. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am.* 2005;87(7):1515–21.
2. McKee GK, Watson-Farrar J. Replacement of arthritic hips by the McKee-Farrar prosthesis. *J Bone Joint Surg Br.* 1966;48(2):245–59.
3. August AC, Aldam CH, Pynsent PB. The McKee-Farrar hip arthroplasty. A long-term study. *J Bone Joint Surg Br.* 1986;68(4):520–7.
4. Charnley J. The long-term results of low-friction arthroplasty of the hip performed as a primary intervention. *J Bone Joint Surg Br.* 1972;54(1):61–76.
5. Schmalzried TP, Peters PC, Maurer BT, Bragdon CR, Harris WH. Long-duration metal-on-metal total hip arthroplasties with low wear of the articulating surfaces. *J Arthroplasty.* 1996;11(3):322–31.
6. Saito S, Ryu J, Watanabe M, Ishii T, Saigo K. Midterm results of Metasul metal-on-metal total hip arthroplasty. *J Arthroplasty.* 2006;21(8):1105–10.
7. Müller ME. The benefits of metal-on-metal total hip replacements. *Clin Orthop Relat Res.* 1995;(311):54–59.
8. Ladon D, Doherty A, Newson R, Turner J, Bhamra M, Case CP. Changes in metal levels and chromosome aberrations in the peripheral blood of patients after metal-on-metal hip arthroplasty. *J Arthroplasty.* 2004;19(8 Suppl 3):78–83.
9. Davies AP, Willert HG, Campbell PA, Learmonth ID, Case CP. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am.* 2005;87(1):18–27.
10. Willert H, Buchhorn GH, Fayyazi A, Flury R, Windler M, Köster G, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Jt Surg Am.* 2005;87(1):28–36.
11. Cahir JG, Toms AP, Marshall TJ, Wimhurst J, Nolan J. CT and MRI of hip arthroplasty. *Clin Radiol.* 2007;62(12):1163–71. discussion 1172–1173.
12. Lee MJ, Janzen DL, Munk PL, MacKay A, Xiang QS, McGowen A. Quantitative assessment of an MR technique for reducing metal artifact: application to spin-echo imaging in a phantom. *Skeletal Radiol.* 2001;30(7):398–401.
13. Chang SD, Lee MJ, Munk PL, Janzen DL, MacKay A, Xiang QS. MRI of spinal hardware: comparison of conventional T1-weighted sequence with a new metal artifact reduction sequence. *Skeletal Radiol.* 2001;30(4):213–8.
14. Gruber FW, Böck A, Trattnig S, Lintner F, Ritschl P. Cystic lesion of the groin due to metallosis: a rare long-term complication of metal-on-metal total hip arthroplasty. *J Arthroplasty.* 2007;22(6):923–7.
15. Toms AP, Nolan J, Barker T, Darrah C, Malcolm P. Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR. *Br J Radiol.* 2009;82(977):e87–91.
16. Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donell ST, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol.* 2008;63(1):49–58.
17. Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons CLM, et al. Pseudotumours associated with metal-on-metal hip resurfacings. *J Bone Joint Surg Br.* 2008;90(7):847–51.
18. Hart AJ, Sabah S, Henckel J, Lewis A, Cobb J, Sampson B, et al. The painful metal-on-metal hip resurfacing. *J Bone Joint Surg Br.* 2009;91(6):738–44.
19. Toms AP, Kasmai B, Williams S, Wilson P. Building an anonymized catalogued radiology museum in PACS: a feasibility study. *Br J Radiol.* 2006;79(944):666–71.
20. Haahr M. RANDOM.ORG - True Random Number Service. 2010. <http://www.random.org/>. Accessed 1 Jun 2010.
21. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging.* 2004;17(3):205–16.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.

# Building an anonymized catalogued radiology museum in PACS: a feasibility study

A P TOMS, FRCS, FRCR, B KASMAI, MSc, MIPeM, S WILLIAMS, MRCP, FRCR and P WILSON, MRCP, FRCR

*Department of Radiology, Norfolk & Norwich University Hospital NHS Trust, Colney Lane, Norwich, Norfolk NR4 7UY, UK*

**ABSTRACT.** The aim of this study was to test the feasibility of a software application that would allow the anonymization and cataloguing of whole DICOM datasets in order to build searchable radiology museums within PACS. The application was developed on a dedicated networked PC, using C# and HL7 coding. Whole DICOM datasets were pushed from PACS to a networked PC on which the application, Museum Builder, was developed. Museum Builder works by replacing the patient specific data (the forename, surname and hospital number) within each header of each DICOM file with terms from anatomical and surgical sieve menus. The date of birth is anonymized to 1 January of the same year. Whole DICOM datasets comprising hundreds of images can be anonymized and catalogued in a single episode. Museum Builder primes PACS with an HL7 script to receive a "new" patient. DICOM datasets are then pushed back to PACS where they are added to the database as "new" cases. The museum cases can then be searched for, on PACS, by any combination of terms that correspond to appropriate anatomical units, surgical sieve headings or radiological specialty. New radiology reports containing clinical histories, radiological descriptions, differential diagnoses and discussion can be added through the report window. Our institution has developed and used this tool to generate a PACS based radiology museum containing not only full DICOM datasets, but also relevant histological and clinical photographs. In conclusion, this technique offers a mechanism for generating anonymized catalogued radiology museums in PACS. Museum Builder represents a working prototype that demonstrates some of the archiving functions that are expected by teaching institutions from PACS.

Received 22 June 2005  
Revised 6 December 2005  
Accepted 9 December 2005

DOI: 10.1259/bjr/11861604

© 2006 The British Institute of  
Radiology

The radiology museum is an integral component of every radiological training scheme. Over the past 5 years, the practice of radiology has moved from film to PACS, but the ability to build radiology museums has not kept pace [1].

For many of us, the hardcopy ACR collection provided hours of study and exam practice material. Those cases we saw in our formative years of radiological training often become "index" cases against which those that followed were measured. These museums, which are becoming increasingly obsolete [2], often comprised cupboard-like rooms filled with shelf upon shelf of ageing radiographs in various states of disorganization. Radiological museums have now diversified into multiple digital formats. DICOM files can easily be converted and saved in a number of manageable formats [3]. Large institutional collections can be acquired on CD-ROM [4, 5]. Personal teaching collections can be created in any number of readily available image databases [6–10]. Some of these databases are specifically designed for archiving radiological teaching cases and sometimes for storage on servers [4, 11, 12] for sharing access across networks or the World Wide Web [13]. Online database applications allow free text searches across thousands of cases,

sometimes in multiple institutions [13]. These archives are used for research, teaching [2, 9, 10, 14] and for assessment of radiological expertise [15]. With this digital diversity has come a subtle change in the look and feel of the radiology museum. Hardcopy museum cases must be read in the same way that hardcopy radiology is practiced; with a light box. Digital museum cases are read from personal computers and not in the PACS environment in which many of us now work. One of the obstacles to replicating the PACS environment on a PC is the prohibitive size of the DICOM files. A solution to these problems is not to replicate PACS in order to build a work-like radiology museum, but to build a radiology museum within PACS. To the best of our knowledge, the major PACS manufacturers provide only limited tools for archiving radiology teaching cases, whereas most radiologists consider this sort of functionality important or even essential when considering the purchase of PACS [16]. Some provide a system of academic folders that require a system administrator to set up. These provide inadequate archiving and retrieval mechanisms for generating usable databases within PACS [1]. Neither does there appear to be any third party solutions that meet these criteria. PACS manufacturers prohibit access to their databases other than by their employees and, therefore, a novel approach is required for third parties to generate teaching cases on PACS. The aim of this study was to see if it was feasible to

This project was supported by a Kodak Radiology Fund Scholarship awarded by the Royal College of Radiologists in 2003.



develop a software application that could edit the DICOM headers to anonymize and catalogue teaching cases in order to build a radiology museum within PACS.

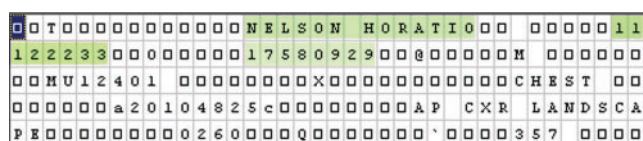
## Materials and methods

### Principle

Radiological studies stored on PACS can be identified and retrieved using a number of search fields common to all PACS – namely, the patient’s surname, forename, middle names and unique hospital number. This information is stored within a header in every DICOM file (Figure 1). This usually means that every image, in every series, in every radiological study contains this information embedded within it. Multiframe images, generated with ultrasound, contain the same information in a single header. Both of these types of DICOM file can be handled in the same way. After DICOM data has been generated by a radiological investigation, it is pushed to the PACS server to be archived. As PACS receives the DICOM data, it reads the DICOM header and stores the patient specific data in a database. When PACS is queried to search for a particular patient, it is this database that is searched and not the DICOM archive itself. This entry in the database, however, points to the DICOM dataset within the archive, which can then be retrieved and opened. Our application, called Museum Builder, exploits this process by replacing the patient specific data within the header of DICOM files that have been exported from PACS. When the DICOM files are returned to PACS, the new information within the header is added to the PACS database as a “new” patient. The DICOM files for the museum case are effectively duplicated, but with a new DICOM header. In effect, PACS sees Museum Builder as any other radiological modality contributing to the local PACS archive.

### Hardware and software

Our institution is a film-free hospital with a GE Centricity PACS (General Electric, Milwaukee, WI). Images can be viewed from a mixture of dedicated reporting workstations and PC-based web-browsers, which cover the whole hospital on a network with a 2 Gb s<sup>-1</sup> backbone and a 100 Mb s<sup>-1</sup> link to workstations. Museum Builder was developed using .NET technology and C# as the preferred language (Microsoft .Net Framework to run and Visual Studio to compile the C# source code). Museum Builder was installed on a networked PC (Pentium 4 CPU 2.80 Ghz with 1.0 Gb of



**Figure 1.** A screenshot of a readout from a hexadecimal editor of the header from a DICOM file demonstrating the ASCII values of the binary data. The patient specific binary data can be identified and edited with Museum Builder.

RAM running Microsoft Windows XP Professional, version 2002, with Service Pack 1). An academic licensed version of eFilm Workstation 1.9.4 [17] was installed as the helper application.

## Results

## Process

Museum Builder can work with most PC based DICOM browsers, which for the purpose of this article will be referred to as the helper application. After opening Museum Builder, the helper application's database can be browsed or searched using search fields for the patient's name or hospital number. Cases are then identified for museum archiving and highlighted. Once selected, the patient's surname, forename, hospital number and date of birth are displayed in a row of text fields. Below this a second row contains the text fields for the anonymized museum case. The patient's "new" surname and forename are selected from pre-defined menus. Optional middle names can also be entered. The new hospital number comprises a three letter code, again selected from a menu, followed by a unique four digit number generated by Museum Builder (Figure 2). The drop-down menu selections can be over-ridden by entering text into the fields manually. The patient's date of birth is automatically anonymized to 1 January of the year of their birth, thus preserving age-related information in the DICOM header. Once the minimum number of fields has been completed (surname, forename and hospital number), the case is ready for anonymization (Figure 2). Clicking the "anonymize" button will then replace the patient specific data in each header of each DICOM file within each directory of the chosen radiological study or studies and the helper application database is modified accordingly. All other identifiable labels, private or otherwise, are changed or erased. This one click anonymizes every radiological study in the eFilm database for that patient; there is no limit to the number of types of examination, *e.g.* conventional radiography, CT, ultrasound or MRI, that can be processed at once. Even for large datasets with multiple examinations this only takes a few seconds on the workstation described. The process is simple and is currently performed in our department by clerical staff. Every DICOM file generated by a medical imaging device contains a unique identifier (the study SOP instance UID). At this stage, Museum Builder assigns each new case with a new unique identifier, generated internally from a sub-delegate range offered by Medical Connections [18], so that PACS does not recognize it when it returns. PACS just sees another new case arriving from a medical imaging device. When coding and anonymization is complete, Museum Builder composes an HL7 Radiology order message, which is sent to the PACS HIS/RIS broker. On receipt of this acknowledgment the new museum case is pushed from the helper application back to PACS (Figure 3).

### Catalogue headings

The American College of Radiologists (ACR) has established a well-recognized classification system for

**Norwich Image Museum Builder**

File

Search:  Patient ID Surname Browse Search Next Exit Render Attributes Anonymize Study

**Current Patient Attribute Values**

Accession Number	Patient ID	DOB	Patient Name
000213809	805 100 112	19660909	TOMS*ANDONI

**New Patient Attribute Values**

Accession Number	Patient ID	DOB	Patient Name
MUSEUM1003	CNS1001	19660101	HEAD NORMAL BRAIN

**Study Coding**

Anatomy	Surgical Sieve	Organ	Organ	Other	Radiology System Code	Proposer
HEAD	NORMAL	BRAIN			CNS	
ABDOMEN	CONGENITAL	CHEST	ABDOMEN AND PELVIS	CORE	BREAST	ANONYMOUS
BREAST	DEGENERATIVE	HEAD	CHEST	HISTORICAL	CARDIOVASCULAR	CLINBK
CHEST	DEVELOPMENT	MANDIBLE	HEAD	JOINT	CENTRAL NERVOUS SYSTEM INC. OF	AT8
EXTREMITY	IATROGENIC	ORBIT	NECK	PROSTHESIS	GASTROINTESTINAL INC. BILIARY SY	
HEAD	INFECTION	PETROUS-TEMPOR	OTHER		HEAD & NECK INC. PETROS TEMPOR	
NECK	INFLAMMATORY	DENTAL	SPINE		LYMPHORETICULAR INC. SPLEEN	
PELVIS	MISCELLANEOUS	FACE	EXTREMITY		MUSCULOSKELETAL INC. PERIPHER	
SPINE	NEOPLASIA	NECK			PAEDIATRICS INC. ANTENATAL	
	NORMAL				REPIRATORY SYSTEM	
	TRAUMA				UROGENITAL SYSTEM INC. ADRENAI	
	VARIANT					
	VASCULAR					

**List of studies**

PatientID	Patient's Name	Study Description	Accession Number	Date of Birth	Se
805 100 112	TOMS*ANDONI	BRAIN	000213809	19660909	AX
805 100 112	TOMS*ANDONI	BRAIN	000213809	19660909	Axi
805 100 112	TOMS*ANDONI	BRAIN	000213809	19660909	DIF
805 100 112	TOMS*ANDONI	BRAIN	000213809	19660909	SA

**Image Directory:**

c:\Program Files\veFilm\DICOM

**Database File Name (Full Path)**

c:\Program Files\Database\

**MuseumBuilder Database (Full Path)**

c:\Program Files\MuseumBuilder\MuseumBuilder

Edit Save

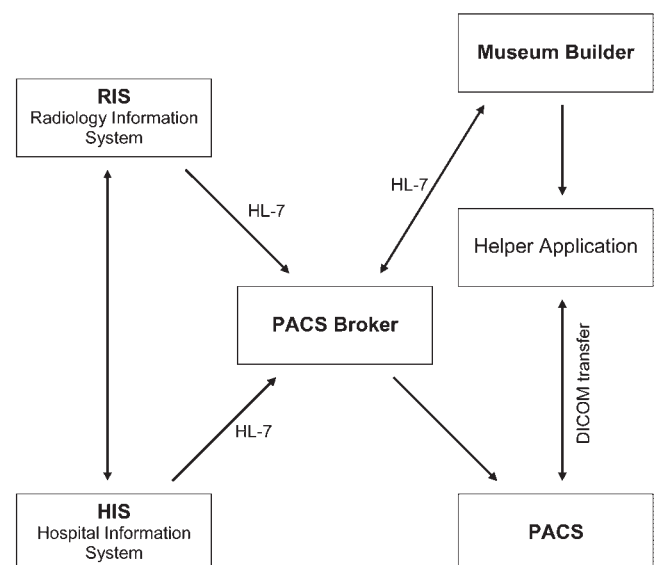
**Figure 2.** A screenshot from Museum Builder's anonymization window. The patient specific fields are replaced by catalogue headings that are selected from drop down menus.

radiology [19]. This system allows high-level discrimination of radiological diagnoses, and therefore accurate retrieval of data, particularly for research. However, the system is not entirely intuitive, and therefore a different classification system has been used in Museum Builder. The catalogue headings are almost universal within radiology. The patient's surname is changed to a radiological anatomical unit, which broadly defines the sections of the human body that radiological investigations cover. These consist of head, neck, chest, abdomen, pelvis, extremities and breast. The patient's forename is selected from a surgical sieve consisting of normal, developmental, trauma, infection, neoplasia, inflammation, vascular, metabolic. Thereafter there is an option to add one, two or more middle names from a selection of organ specific titles such as liver, lung, brain, adrenal and so on. The hospital number is replaced by a unique museum number that comprises a three-letter code followed by a four-digit number. The code reflects subspecialty interests within radiology and include MSK for musculoskeletal, GIT for gastrointestinal and H&N for head and neck (Figure 2).

### Non-radiological images

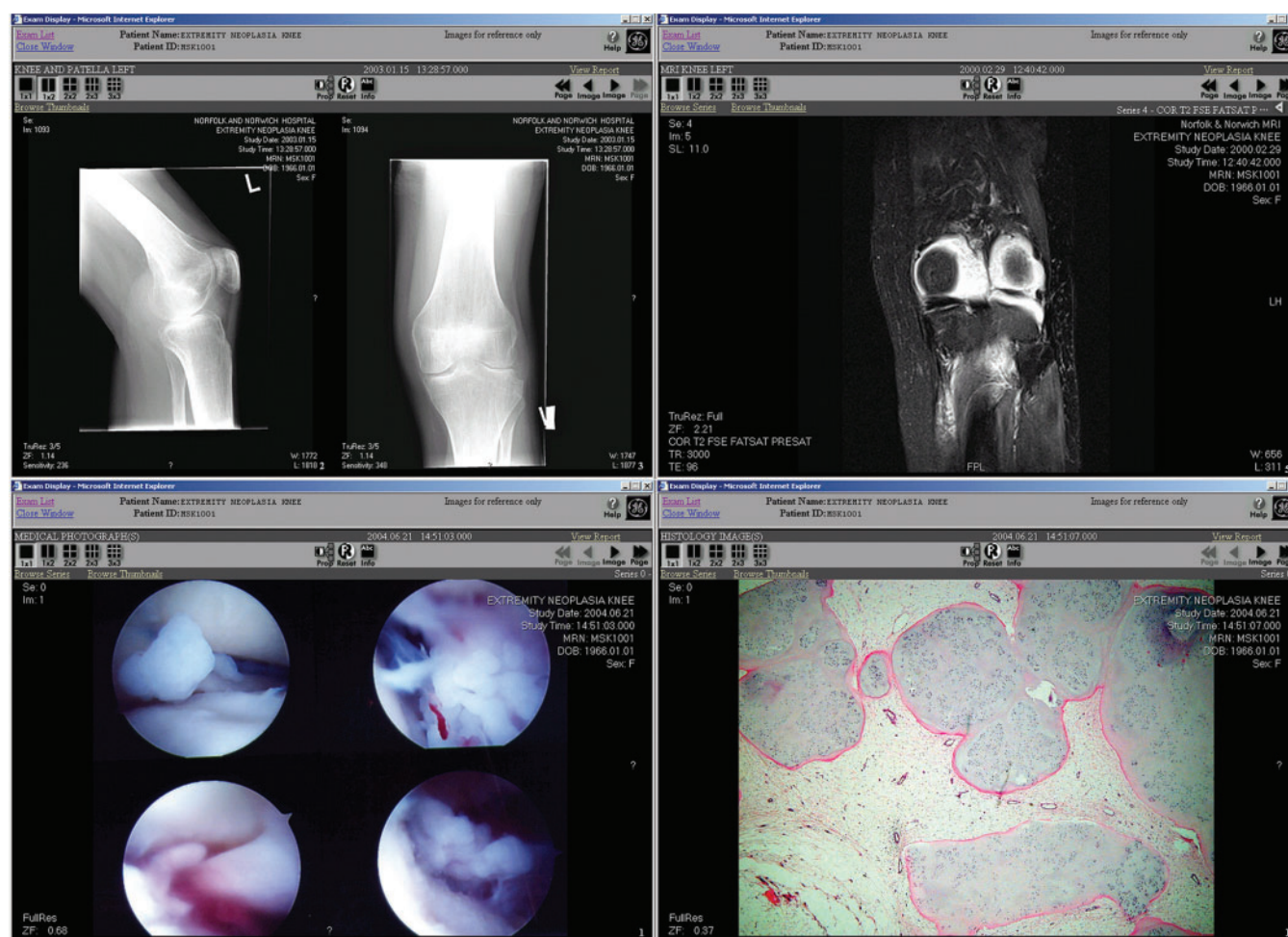
Many DICOM browsers will import non-DICOM image files such as JPEG and TIFF files. During the

import process, DICOM header fields are entered manually and new DICOM files are generated. This allows non-radiological images to be added to the



**Figure 3.** A schematic diagram demonstrating data flow between Museum Builder, the helper application, PACS and RIS/HIS.





**Figure 4.** A montage of screenshots demonstrating material from a single museum case including computed radiography, MRI, arthroscopy and histology.

museum case, including histology, arthroscopy, endoscopy and clinical photographs (Figure 4). The report field of these “new studies” can then contain pathology reports, operative notes and clinical findings.

## Reports

Reports for the museum cases can be added to the PACS RIS. These can be added by pasting text into a report window in Museum Builder (Figure 5). These reports can include copies of the original radiological report or can be entirely new and contain clinical histories, updated radiological reports, results of other special investigations, differential diagnoses and discussion.

## Discussion

Museum Builder is a fully working prototype that has some of the functionality required by teaching institutions to develop radiology museums within PACS. It uses a novel approach to generating anonymized searchable teaching cases without accessing the PACS database directly. Teaching cases can then be read in the PACS environment in which the radiology trainee and his or her trainers work. Instead of being presented with

single or selected digital images from a teaching archive or CD-ROM, the trainee has access to the whole DICOM dataset. When reading cross-sectional investigations, the trainee would have to interrogate the entire dataset, including localizers and sequences repeated because of technical problems, to gather the signs necessary to yield a diagnosis. It is this process that cannot be replicated by non-DICOM museums and teaching collections.

There are limitations to radiological museums created by Museum Builder. There is no free-text search function within the PACS browser window, which would allow the user to search for a specific diagnosis. However, the objective was to create a radiological museum that functioned in a similar manner to the ACR hard copy museum and therefore did not require the ability to immediately recall specific cases. Trainees can search through the database through catalogue headings based on anatomical site, disease process and radiology subspecialty.

Museum Builder can be used with any PACS and, in theory, integrated with any PACS broker, but this has only been tested with PACS in our institution. Whilst it should work with any PACS broker, the concept of Museum Builder does not allow for a “plug and play” solution. Museum Builder needs to be configured for each PACS broker in the same way that any CT or ultrasound machine must be configured to work with a

**Museum Builder Reporting System**

**Patient Data**

Patient ID: CNS1042 Patient Name: SPINE NEOPLASIA CERVICAL Step 1. Search For Study

DOB: 01/01/2019 Sex: F Address: Museum Builder

Accession Number: MU12031 Modality: MR Report Date/Time (dd/MM/yyyy HH:mm): 13/04/2005 15:58

Procedure Code: M.SPIP Procedure Description: MRI WHOLE SPINE PRE AND POST CONTRAST

**Report**

Diagnosis: Grade I pilocytic astrocytoma.

Comments: Rare tumour. Post operative radiotherapy planned because of a risk of local recurrence. Patient mobile with two sticks following surgery.

Validation: 1. Histology 2. Consensus

**HL7 Message**

Step 2. Compile Message(s)

```
MSH|^~\&|MUSEUMBUILDER|N||200504131605||ORU^R01|MU2005041316053
OIP|2.3||PID|1||CNS1042||SPINE NEOPLASIA CERVICAL||20190101|F||Museum
Builder||01603 999999||w|||||||||P
V1|1|NO|NO|P||||MUSEUM|C||||||20050413160530|O
RC|SC|MU12031|MU12031^MUSEUMBUILDER|CM|1|^N||20050413160530
```

**PACS Broker**

IP Address: 10.91.120.50 Port Number: 3328 Step 3. Reconnect to Server Disconnect from Server

**Exit** Step 4. Send Report

**Message Status**

**Activity Logbook**

```
Connected to server machine: 10.91.120.50 Port:
3328 Client Stopped Connection closed by machine
with IP address: 10.91.120.50 Close Handler
Connected to server machine: 10.91.120.50 Port:
3328
```

**Acknowledgment**

```
MSH|^~\&|HL7-RIS-INIMITRA-BROKER|MUSEUMB
UILDER|N|20050413161756|ACK|MU200504131605
30|P|2.3||MSA|A|MU20050413160530|
```

**Figure 5.** A screenshot demonstrating the "Report Window" where the text of the museum case has been pasted in prior to sending to RIS in an HL7 script.

particular PACS. However, configuring Museum Builder has been simplified by using variables in the HL7 code that can be defined from within the Access database according to the local PACS broker profiles.

The concept of Museum Builder is relatively simple and the coding is mainstream. It currently works as a bolt-on application to PC-based DICOM browsers, but there are a number of options for further development. Museum Builder could be coded to work as a DICOM client and, therefore, could stand alone in its integration with PACS and RIS. However, it does not make sense to repeat the work done by many affordable or free, readily available proprietary DICOM browsers. It would be easier to add Museum Builder's functionality to these

applications. The most elegant solution would be for the PACS manufacturers to add this functionality to their current systems. That way the radiologist could build his or her DICOM radiology museum without leaving PACS.

This sort of functionality within PACS has certain implications for governance of the educational material because other allied healthcare workers outside radiology, and IT personnel, also have access to the database. In our institution we developed a governance protocol that was approved by the Caldicott Guardian to ensure that the limitations of patients' consent to procedures and investigations were adhered to. Rather than being a risk, the radiology museum is considered a valuable



hospital-wide resource. Generating validated case material is always time consuming, and therefore the number of museum cases has, so far, made a negligible impact on the PACS archive capacity. In theory, however, duplicating large volumes of archive material could have serious implications for storage and therefore needs to be carefully controlled. In our institution, all museum material must be approved by a Radiology Museum Committee, which acts as a gatekeeper safeguarding the quality of the PACS museum and controlling its impact on the clinical archive.

## Conclusion

Museum Builder demonstrates that it is feasible to build anonymized catalogued radiology museums within PACS, by editing the patient specific headers within the DICOM files, and therefore without directly accessing the PACS database. Teaching cases generated with this tool allow the trainee to read the full DICOM datasets within the normal PACS working environment. By using PACS as the radiology museum repository, the problems of storing and transmitting large image files and directories can be overcome. Museum Builder provides a model for the some of functionality that many academic institutions would like to see added to PACS by the PACS manufacturers.

## References

1. Siegel E, Reiner B. Electronic teaching files: seven-year experience using a commercial picture archiving and communication system. *J Digit Imaging* 2001;14:125–7.
2. Wiggins RH, Davidson HC, Dilda P, Harnsberger HR, Katzman GL. The evolution of filmless radiology teaching. *J Digit Imaging* 2001;14:236–7.
3. Halsted MJ, Moskovitz J, Johnson N, Perry L. A simple method of capturing PACS and other radiographic images for digital teaching files or other image repositories. *AJR Am J Roentgenol* 2002;178:817–9.
4. Rosset A, Muller H, Martins M, Dfouni N, Vallee JP, Ratib O. Casimage project: a digital teaching files authoring environment. *J Thorac Imaging* 2004;19:103–8.
5. Bernard J, Gerber S, Oppenheim C, Marsault C, Dormont D. Creation of an educational CD-ROM using a PACS. *J Radiol* 2002;83:68–70.
6. Maldjian JA, Listerud J. Automated teaching file and slide database for digital images. *AJR Am J Roentgenol* 2000;175:1249–51.
7. Macura KJ, Macura RT, Morstad BD. Digital case library: a resource for teaching, learning, and diagnosis support in radiology. *Radiographics* 1995;15:155–64.
8. Weinberger E, Jakobovits R, Halsted M. MyPACS.net: a web-based teaching file authoring tool. *AJR Am J Roentgenol* 2002;179:579–82.
9. Strickland NH, Gishen P. From Silver to Silicon: film libraries of the future. *Imaging* 1994;6:143–6.
10. Strickland NH, Allison DJ, Gishen P. Technical note: a radiological educational system – organization of an image library. *Br J Radiol* 1995;68:524–7.
11. Rosset A, Ratib O, Geissbuhler A, Vallee JP. Integration of a multimedia teaching and reference database in a PACS environment. *Radiographics* 2002;22:1567–77.
12. Radfiler Website. Available at: [www.radfiler.com](http://www.radfiler.com) [Accessed 19 April 2005].
13. Lim CC, Yang GL, Nowinski WL, Hui F. Medical Image Resource Center—making electronic teaching files from PACS. *J Digit Imaging* 2003;16:331–6.
14. Dugas M, Trumm C, Stabler A, et al. Case-oriented computer-based-training in radiology: concept, implementation and evaluation. *BMC Med Educ* 2001;1:5.
15. Mullins ME, Will M, Mehta A, Novelline RA. Evaluating medical students on radiology clerkships in a filmless environment: use of an electronic test prepared from PACS and digital teaching collection images. *Acad Radiol* 2001;8:514–9.
16. Blunt D, O'Regan D. Using PACS as a teaching resource. *Br J Radiol* 2005;78:483–4.
17. Efilm Website. Available at [www.merge.com](http://www.merge.com) [Accessed 22 June 2005].
18. Medical Connections Website. Available at [www.medical-connections.co.uk](http://www.medical-connections.co.uk) [Accessed 22 June 2005].
19. The American College of Radiology Website. Available at [www.acr.org](http://www.acr.org) [Accessed 22 June 2005].

# Limitations of Single Slice Dynamic Contrast Enhanced MR in Pharmacokinetic Modeling of Bone Sarcomas

ANDONI P. TOMS<sup>1</sup>, LAWRENCE M. WHITE<sup>2</sup>, RITA KANDEL<sup>3</sup>, ROBERT R. BLEAKNEY<sup>2</sup>,  
MICHAEL NOSEWORTHY<sup>4</sup>, SHEPSTONE LEE<sup>5</sup>, MARTIN E. BLACKSTEIN<sup>6</sup> & JAY WUNDER<sup>7</sup>

<sup>1</sup>Department of Radiology, The Norfolk & Norwich University Hospital, Colney Lane, Norwich, Norfolk NR4 7UY, UK, <sup>2</sup>Department of Medical Imaging, Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G 1X5, Canada, <sup>3</sup>Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G 1X5, Canada, <sup>4</sup>Health Sciences Centre, Faculty of Health Sciences, McMaster University, 1200 Main Street West, Hamilton, ON L8N 3Z5, Canada, <sup>5</sup>Institute of Health, University of East Anglia, Norwich, Norfolk, UK, <sup>6</sup>Department of Oncology, Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G 1X5, Canada and <sup>7</sup>Musculoskeletal Oncology Unit, Mount Sinai Hospital, 600 University Avenue, Toronto, ON M5G 1X5, Canada

Andoni P. Toms, Lawrence M. White, Rita Kandel, Robert R. Bleakney, Michael Noseworthy, Shepstone Lee, Martin E. Blackstein, Jay Wunder. Limitations of Single Slice Dynamic Contrast Enhanced MR in Pharmacokinetic Modeling of Bone Sarcomas. *Acta Radiol* 2009;50:512–520.

**Background:** Single slice dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) appears to provide perfusion data about sarcomas in vivo that correlate with tumor necrosis on equivalent pathological sections. However, sarcomas are heterogeneous and therefore single slice DCE-MRI may not correlate with total tumor necrosis.

**Purpose:** To determine whether changes in pharmacokinetic modeling of DCE-MRI, during chemotherapy for primary bone sarcomas correlated with histological measures of total tumor necrosis.

**Material and Methods:** Twelve patients with appendicular primary bone sarcomas were included in the study. Each patient had DCE-MRI before, and after completion, of pre-operative chemotherapy. The mean arterial slope (A), endothelial permeability coefficient ( $K^{\text{trans}}$ ), and extravascular extracellular volume ( $V_e$ ) were derived from each data set using a modified two compartment pharmacokinetic model. Total tumor necrosis rates were compared with changes in A,  $K^{\text{trans}}$ , and  $V_e$ .

**Results:** Six patients had total tumor necrosis of  $\geq 90\%$  and six had a measure of  $< 90\%$ . The median percentage changes in A,  $K^{\text{trans}}$ , and  $V_e$  for the  $\geq 90\%$  necrosis group were  $-52.5\%$  ( $-83$  to  $6$ ),  $-66\%$  ( $-82$  to  $26$ ), and  $23.5\%$  ( $-26$  to  $40$ ), respectively. For the  $< 90\%$  necrosis group, A =  $-35\%$  ( $-75$  to  $132$ ),  $K^{\text{trans}}$  =  $-53$  ( $-66$  to  $149$ ) and  $V_e$  =  $-14.5\%$  ( $-42$  to  $40$ ). One patient with  $> 90\%$  necrosis had increases in all three measures. Comparison of the two groups generated *P*-values of 0.699 for A, 0.18 for  $K^{\text{trans}}$ , and 0.31 for  $V_e$ .

**Conclusion:** There was no statistically significant correlation between changes in pharmacokinetic perfusion parameters and total tumor necrosis. When using single slice DCE-MRI heterogeneous histology of primary bone sarcomas and repair mediated angiogenesis might both be confounding factors.

**Key words:** MRI; perfusion; sarcoma

*Dr. A.P. Toms, Department of Radiology, Norfolk & Norwich University Hospital, Colney Lane, Norwich, Norfolk NR4 7UY, UK (tel. +44 1603 286104, fax. +44 1603 286077, e-mail. andoni.toms@nnuh.nhs.uk)*

*Accepted for publication February 11, 2009*

It is recognized that conventional magnetic resonance imaging (MRI) sequences are limited in predicting the histological grade of tumors and

their response to chemotherapy (1). Dynamic contrast-enhanced MRI (DCE-MRI) is a robust imaging technique, which has been used for evaluating

angiogenesis in a number of tumor types (2, 3), and has become established as a research tool for use in primary bone sarcomas (4–8) but has so far proved to be unhelpful in differentiating benign from malignant bone tumors (9, 10).

The main focus of research to date has been in evaluating DCE-MRI as a technique for predicting patients' response to chemotherapy prior to surgical resection. Tumor necrosis of 90% or more is associated with greater than 80% 5-year survival rates (11, 12). Knowledge of the efficacy of an expensive therapy, with significant side effects, may contribute significantly to the decision to continue therapy in individual patients. DCE-MRI has most commonly been reported as a single examination performed after chemotherapy and prior to surgery (13–18). A number of measures have been derived from the time–intensity curves generated from the DCE-MRI data and correlated with histopathological measures (19, 20) of cell necrosis. The mean slope of the arterial phase of enhancement (A) has been reported as correlating with necrosis with accuracies ranging from 86 to 95% (16, 21–24). Other measures including Dynamic Vector Magnitudes (DVM) (17) and Factor Analysis of Medical Image Sequences (FAMIS) (25) can be calculated readily from the DCE-MRI data and correlate well with measures of necrosis, again when compared with single macro slices.

By applying compartmental modeling algorithms (26) to the data sets, further kinetic parameters can be derived including  $K^{trans}$  (endothelial cell permeability surface area product) and  $V_e$  (extravascular space volume) (27–29). Correlation of these kinetic parameters has also shown promise as predictors of necrosis in single slice histopathological comparisons (14, 15). Serial measures demonstrating changing perfusion parameters during chemotherapy have also been described with encouraging results (1, 17, 30).

A theoretical disadvantage of these techniques is that the histopathological macro slice may not be representative of the whole tumor either because the tumor is eccentrically shaped or because the differentiation of the cell types is heterogeneous. The null hypothesis of this study is that changes in pharmacokinetic endpoint measures, derived from single slice DCE-MRI, before and after chemotherapy do not correlate with total tumor necrosis.

## Material and Methods

### Patients

Approval for the project was obtained from the Hospital institutional review board in 2001. All

patients presenting to the musculoskeletal oncology unit with primary bone sarcomas between January 2002 and December 2005 were entered into the study ( $n=19$ ). Patients consented to a DCE-MRI perfusion study to be included with their routine pre and post-chemotherapy MRI examinations. Patients were subsequently excluded from the study if they failed to complete their course of chemotherapy ( $n=2$ ) or if either of the DCE-MRI perfusion studies were inadequate for the purposes of the analysis ( $n=3$ ). One patient died during treatment and one did not proceed to chemotherapy after the initial MRI. A total of 12 patients (eight men and four women, median aged 25, range 14–50) met the inclusion criteria and successfully underwent DCE-MRI before and after chemotherapy. The histological diagnoses included seven osteosarcomas (one telangiectatic), two leiomyosarcomas of bone, one malignant fibrous histiocytoma of bone, one chondroblastic osteosarcoma and, one sarcoma NOS (Not otherwise specified) (Table 1). The maximal long axis of the tumors ranged from 3.5 to 19 cm (median 9.5).

### Magnetic resonance imaging (MRI)

The DCE-MRI was performed on a 1.5T GE signa MRI system (GE Healthcare, Milwaukee, Wisc., USA). Following routine MR staging sequences (T1W FSE, TE 8.0, TR 700 and fat saturated T2W, TE 81.2 TR 4600 in axial, and sagittal or coronal planes) a single slice, oriented in an orthogonal plane along the longest axis of the tumor, was selected by the attending radiologist. About 20 ml of 0.5 mmol/ml gadolinium chelate (Omniscan, GE Healthcare, Wisc., USA) was injected intravenously, at 1.5 ml/s, into a 20 gauge cannula in the antecubital fossa, followed by 20 ml of normal

Table 1. demonstrating the histological type of the tumors in the twelve patients recruited in to the study

Patient	Histological diagnosis	Grade	Necrosis (%)
1	Osteosarcoma	3	80
2	Leiomyosarcoma	2	40–50
3	Malignant fibrous histiocytoma	3	15
4	Sarcoma NOS	NA*	50
5	Osteosarcoma	3	50
6	Chondroblastic osteosarcoma	3	50–75
7	Leiomyosarcoma	3	100
8	Osteosarcoma	3	>98
9	Osteosarcoma	3	95
10	Osteosarcoma	3	>90
11	Telangiectatic osteosarcoma	3	90
12	Osteosarcoma	3	>90

\*Not available as a result of limited sampling.

saline. 300 gradient echo T1W images (TE 2.7, TR 13.2, flip angle 20, 6.0 mm slice thickness, FOV  $28 \times 28$  cm, matrix  $256 \times 128$ ) were acquired over 5 min (temporal resolution of 2 s). The post-chemotherapy data was acquired, with the same parameters, using the pre-chemotherapy study as a guide for the slice localization. The median time between first and second MR examinations was 104 days (range 60–113). The second MRI examination was performed the day before surgery. Surgery was scheduled after the sixth cycle of chemotherapy when all blood indices had recovered.

### Histological evaluation

Patients underwent surgery the day after the second DCE-MRI examination. The resected tumor was fixed in formalin and serially sectioned, either transversely or sagittally, depending on the location of the tumor. Approximately 50% of the tumor was submitted for decalcification and processing for paraffin embedding and 5  $\mu$ m sections were cut, stained with haematoxylin and eosin, and examined by light microscopy. The percentage tumoral necrosis, or replacement by fibrous or fibro-osseous tissue, was determined following evaluation of all histological sections.

### Pharmacokinetic modeling

Pharmacokinetic modeling of the DCE-MRI perfusion data sets were performed using modified two compartment model (26) scripted into a bespoke software application (Matlab<sup>®</sup> version 6.0.0) running on a dedicated workstation (Sun Microsystems, Mountain View, Calif., USA). Regions of interest were drawn by one of two musculoskeletal radiologists using the conventional spin echo sequences as a template (Fig. 1). The following parameters for each MRI examination were obtained: mean arterial

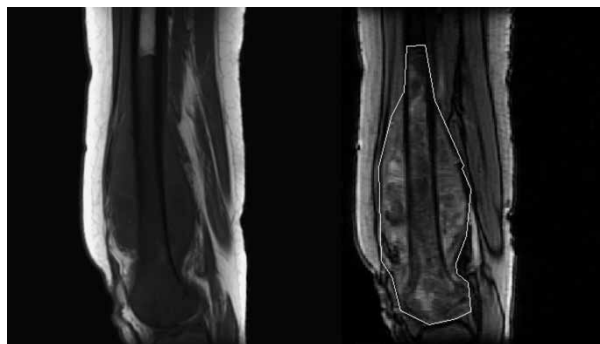


Fig. 1. Sagittal T1W MR (A) through the femur of a patient with an osteosarcoma used to identify the long axis of the tumor for selecting the ROI from the gradient echo T1W perfusion sequence (B).

slope (A), extravascular space volume ( $V_e$ ), and endothelial cell permeability surface area product ( $K^{trans}$ ). The  $K^{trans}$  and  $V_e$  parameters were determined based on the Tofts model (31). The Bera–Jarque test for normality was used on regions of interest (ROIs) and revealed a need for non-parametric based statistical analysis (below).

### Statistical analysis

The percentage change in mean arterial slope (A), endothelial permeability coefficient ( $K^{trans}$ ), and extravascular extracellular space ( $V_e$ ) was calculated (Table 2). These did not appear to follow a normal distribution and therefore a comparison of the spreads of the samples between the two groups (less than 90% necrosis and equal to, or greater than, 90% necrosis) was performed using the Mann–Whitney *U* test (SPSS for Windows 12.0.2) for which an assumption of a normal distribution is not required.

## Results

### Tumor necrosis

Sarcomas from six patients showed 90% or greater tumor necrosis and the remaining tumors ( $n=6$ ) showed less than 90% necrosis (Table 2).

### Pre-chemotherapy

Prior to chemotherapy the following median measures, with interquartile ranges, were obtained for patients with less than 90% total tumor necrosis:  $A = 2.822$  (1.897–3.817),  $K^{trans} = 0.941$  (0.570–1.896) and  $V_e = 0.392$  (0.275–0.533), compared to those patients with tumor necrosis rates  $\geq 90\%$ :  $A = 3.268$  (2.396–4.346),  $K^{trans} = 1.411$  (0.991–2.360), and  $V_e = 0.317$  (0.244–0.363). There was no evidence that the difference of the medians for the two groups was statistically significant: *P*-value for  $A = 0.6889$ ,  $K^{trans} = 0.2980$ , and  $V_e$  mean = 0.3785 (Mann–Whitney *U* Test) (Tables 2 and 3, Figs. 2–4).

### Post-chemotherapy

After to chemotherapy the following median measures, with interquartile ranges, were obtained for patients with less than 90% total tumor necrosis: mean  $A = 2.155$  (1.28–3.12), median  $A = 1.508$  (0.67–2.12),  $K^{trans} = 0.794$  (0.48–1.52), and  $V_e = 0.389$  (0.22–0.50). This compared to those patients with tumor necrosis rates  $\geq 90\%$ : mean  $A = 1.68$  (1.47–1.98), median  $A = 1.092$  (0.72–1.2),  $K^{trans} = 0.559$  (0.48–0.71) and  $V_e = 0.372$  (0.31–0.52). There was no evidence that the difference of the

Table 2. Table demonstrating the mean and median arterial slope (A), mean and median coefficient of endothelial permeability ( $K^{\text{trans}}$ ), and mean and median extracellular extravascular volume ( $V_e$ )

No	Necrosis (%)	Mean, median arterial slope $\pm$ / $\pm$ SD		Mean, median $K^{\text{trans}}$ $\pm$ / $\pm$ SD		Mean, median $V_e$ $\pm$ / $\pm$ SD	
		Pre	Post	Pre	Post	Pre	Post
1	80	3.52, 2.75 $\pm$ / $\pm$ 2.59	1.40, 0.70 $\pm$ / $\pm$ 1.97	1.60, 0.89 $\pm$ / $\pm$ 1.69	0.55, 0.19 $\pm$ / $\pm$ 1.13	0.24, 0.23 $\pm$ / $\pm$ 0.09	0.19, 0.15 $\pm$ / $\pm$ 0.24
2	40-50	2.86, 1.55 $\pm$ / $\pm$ 3.56	2.21, 1.32 $\pm$ / $\pm$ 3.01,	0.98, 0.42 $\pm$ / $\pm$ 1.54	0.71, 0.35 $\pm$ / $\pm$ 1.24	0.48, 0.34 $\pm$ / $\pm$ 0.61	0.67, 0.41 $\pm$ / $\pm$ 0.94
3	15	2.78, 1.80 $\pm$ / $\pm$ 3.10	4.78, 3.11 $\pm$ / $\pm$ 4.33	0.90, 0.49 $\pm$ / $\pm$ 1.28	2.03, 0.92 $\pm$ / $\pm$ 2.29	0.49, 0.36 $\pm$ / $\pm$ 0.50	0.45, 0.40 $\pm$ / $\pm$ 0.36
4	50	1.16, 0.77 $\pm$ / $\pm$ 1.69	2.56, 1.79 $\pm$ / $\pm$ 2.59	0.35, 0.21 $\pm$ / $\pm$ 0.73	0.87, 0.50 $\pm$ / $\pm$ 1.34	0.31, 0.25 $\pm$ / $\pm$ 0.40	0.40, 0.37 $\pm$ / $\pm$ 0.22
5	50	4.71, 4.29 $\pm$ / $\pm$ 3.30	2.10, 1.70 $\pm$ / $\pm$ 2.47	2.80, 1.85 $\pm$ / $\pm$ 2.70	1.34, 0.50 $\pm$ / $\pm$ 1.61	0.29, 0.26 $\pm$ / $\pm$ 0.17	0.23, 0.11 $\pm$ / $\pm$ 0.57
6	50-75	2.14, 1.30 $\pm$ / $\pm$ 3.46	0.92, 0.59 $\pm$ / $\pm$ 1.75	0.64, 0.35 $\pm$ / $\pm$ 1.25	0.26, 0.15 $\pm$ / $\pm$ 0.62	0.66, 0.42 $\pm$ / $\pm$ 0.97	0.38, 0.25 $\pm$ / $\pm$ 0.66
7	100	2.69, 1.24 $\pm$ / $\pm$ 3.43	1.55, 1.15 $\pm$ / $\pm$ 1.79	1.17, 0.34 $\pm$ / $\pm$ 2.10	0.44, 0.31 $\pm$ / $\pm$ 0.66	0.29, 0.26 $\pm$ / $\pm$ 0.16	0.36, 0.29 $\pm$ / $\pm$ 0.36
8	>98	3.28, 2.46 $\pm$ / $\pm$ 3.02	1.47, 0.77 $\pm$ / $\pm$ 2.35	1.63, 0.92 $\pm$ / $\pm$ 1.80	0.49, 0.20 $\pm$ / $\pm$ 0.97	0.26, 0.19 $\pm$ / $\pm$ 0.42	0.32, 0.23 $\pm$ / $\pm$ 0.45
9	95	4.63, 3.56 $\pm$ / $\pm$ 4.06	1.45, 0.59 $\pm$ / $\pm$ 2.94	1.90, 1.12 $\pm$ / $\pm$ 1.95	0.54, 0.15 $\pm$ / $\pm$ 1.31	0.38, 0.31 $\pm$ / $\pm$ 0.38	0.28, 0.22 $\pm$ / $\pm$ 0.36
10	>90	1.51, 1.05 $\pm$ / $\pm$ 1.59	1.81, 1.11 $\pm$ / $\pm$ 2.63	0.46, 0.28 $\pm$ / $\pm$ 0.61	0.58, 0.30 $\pm$ / $\pm$ 1.13	0.35, 0.30 $\pm$ / $\pm$ 0.41	0.39, 0.31 $\pm$ / $\pm$ 0.45
11	90	4.25, 2.97 $\pm$ / $\pm$ 3.83	1.84, 1.07 $\pm$ / $\pm$ 3.08	3.74, 1.19 $\pm$ / $\pm$ 4.13	0.69, 0.29 $\pm$ / $\pm$ 1.57	0.21, 0.18 $\pm$ / $\pm$ 0.20	0.39, 0.29 $\pm$ / $\pm$ 0.63
12	>90	3.25, 2.28 $\pm$ / $\pm$ 3.42	2.38, 1.34 $\pm$ / $\pm$ 3.72	1.19, 0.66 $\pm$ / $\pm$ 1.48	0.75, 0.36 $\pm$ / $\pm$ 1.42	0.36, 0.29 $\pm$ / $\pm$ 0.45	0.90, 0.39 $\pm$ / $\pm$ 1.36

medians for the two groups was statistically significant:  $P$ -value for mean A = 0.589, median A = 0.394,  $K^{\text{trans}}$  = 0.24, and  $V_e$  mean = 1.0 (Mann-Whitney  $U$  test) (Tables 2 and 4; Figs. 2-4).

#### Comparison of before and after chemotherapy

Following chemotherapy the median change in A,  $K^{\text{trans}}$ , and  $V_e$  for the <90% necrosis group were -0.939 (IQ range -2.244 to 1.549), -0.327 (-1.145 to 0.675), and -0.046 (-0.117 to 0.119), respectively (Figs. 2-4). This compared with median change in the  $\geq$ 90% necrosis group of: A = -1.477 (-1.247 to 3.215),  $K^{\text{trans}}$  = -0.934 (0.991 to 2.360), and  $V_e$  = 0.065 (0.01 to 0.270). Comparison of differences between the two groups (Tables 2 and 5; Figs. 2-4) using the Mann-Whitney  $U$  test revealed no significant difference in changes in any of the endpoint measures:  $P(A)$  = 0.4712,  $P(K^{\text{trans}})$  = 0.2298, and  $P(V_e \text{ mean})$  = 0.3785 (Tables 3-5).

A single patient, with a total tumor necrosis rating of >90%, had elevated readings of all three measures; A,  $K^{\text{trans}}$ , and  $V_e$  mean, following chemotherapy (Figs. 2-4 and 9, patient No. 4).

When the changes in A,  $K^{\text{trans}}$ , and  $V_e$  mean were considered as percentages of pre-chemotherapy results (Table 5) the median values for A,  $K^{\text{trans}}$ , and  $V_e$  mean fell for all groups except for the  $V_e$  mean in tumors with greater than or equal to 90% necrosis. In this group the median  $V_e$  mean rose very slightly (Figs. 5-8). Again there were no significant differences between the two categories of necrosis; the  $P$ -values were 0.699 for A, 0.180 for  $K^{\text{trans}}$ , and 0.310 for  $V_e$  mean.

#### Discussion

This study does not provide corroborative evidence that changes in pharmacokinetic endpoint measures, derived from single slice DCE-MRI, before and after chemotherapy correlate with total tumor necrosis. This is contrary to the interpretation of a number of recent studies examining DCE-MRI in bone sarcomas. There are a number of possible explanations for this. The first of which should be addressed are the limitations of the study.

There may have been selection bias introduced in to the study because of the number of patients who met initial entry criteria but subsequently failed to complete the two MRI examinations. It is possible that (along with the one death and one patient who did not have chemotherapy) this selected out patients with more aggressive tumors skewing the population. The number of patients in the study ( $n = 12$ ) is relatively small, and this is reflected in the

Table 3. Tabulated results for comparison of the mean arterial slope (A), mean coefficient of endothelial permeability ( $K^{trans}$ ), and mean extracellular extravascular volume ( $V_e$ ) between patients with tumors with 90% or greater necrosis and those with less, prior to chemotherapy

Variable	Necrosis <90% Median (Q1 to Q3*)	Necrosis ≥90% Median (Q1 to Q3*)	Median difference (95% CI) <sup>†</sup>	P-value <sup>‡</sup>
Mean arterial slope A	2.82 (1.90 to 3.81)	3.27 (2.40 to 4.35)	−0.45 (−2.09 to 1.35)	0.69
Median arterial slope A	1.67 (1.17 to 3.14)	2.37 (1.19 to 3.11)	−0.37 (−1.69 to 1.51)	0.81
Mean $K^{trans}$	0.94 (0.57 to 1.90)	1.41 (0.991 to 2.36)	−0.41 (−1.55 to 0.90)	0.30
Mean $V_e$	0.39 (0.28 to 0.53)	0.31 (0.24 to 0.36)	0.10 (−0.07 to 0.30)	0.38

\*Lower and upper quartiles.

<sup>†</sup>95% Confidence interval for difference in medians.

<sup>‡</sup>P-value from Mann-Whitney *U* test.

wide confidence intervals (Tables 3 and 4). However, encouraging or significant correlation between pharmacokinetic measures and histological outcomes have been reported with smaller sample sizes than our cohort: six (16), eight (15), ten (13, 25), and 12 subjects (30). In our study the mean arterial slope (A) was the endpoint with the greatest overall change in magnitude and numerical difference between the two groups. A sample size calculation suggests that to demonstrate a statistically significant difference of this magnitude ( $P=0.05$ , 90% power) would require a sample size of approximately  $n=40$ . This suggests that this current study is underpowered but also calls in to question the interpretation of studies with similarly small numbers.

The heterogeneity of the tumor type may be a cause for the disparity between these results and previously published data. Our cohort of 12 contained seven osteosarcomas and five primary bone sarcomas of other histological subtypes. Data from

similarly heterogeneous histological cohorts have also produced equivocal results (22). It seems that only data from series of osteosarcomas or Ewing's sarcomas produce encouraging results.

A further cause of error within the method may lie with the MR slice selection. Changes in body habitus and tumor size during treatment make reliable correlative interval slice selection difficult.

The lack of any statistically significant association between mean arterial slope, endothelial permeability coefficient, or the extravascular extracellular volume of the tumor in this study may reflect the fact that the comparison was made with necrosis rates derived from the complete resection specimens. Single slice DCE-MRI may correlate with a corresponding histological macro-slice, but may not represent overall necrosis where the distribution of necrosis is heterogeneous. Most previous studies examining the predictive properties of DCE-MRI have correlated measures of arterial perfusion or pharmacokinetic functions with one

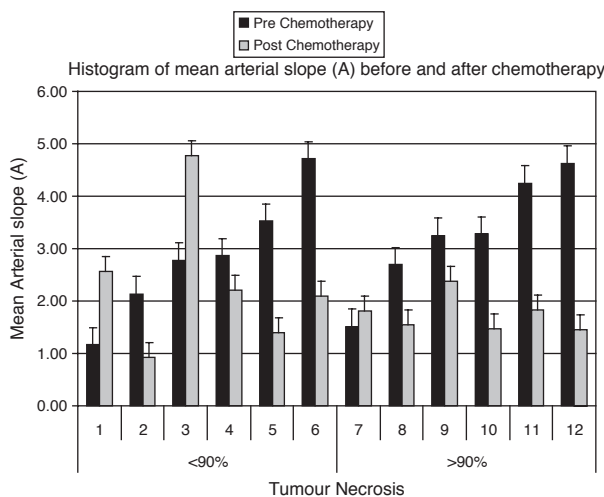


Fig. 2. Histogram demonstrating the mean arterial slope (A) for each patient, in each of the prognostic groups, before and after chemotherapy.

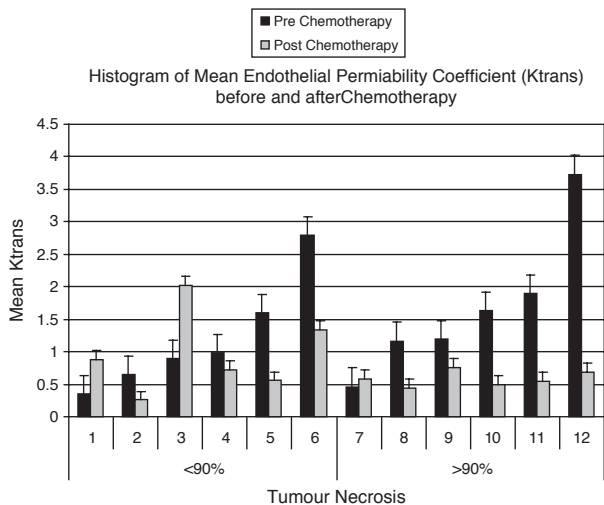


Fig. 3. Histogram demonstrating the mean endothelial permeability coefficient ( $K^{trans}$ ) for each patient, in each of the prognostic groups, before and after chemotherapy.

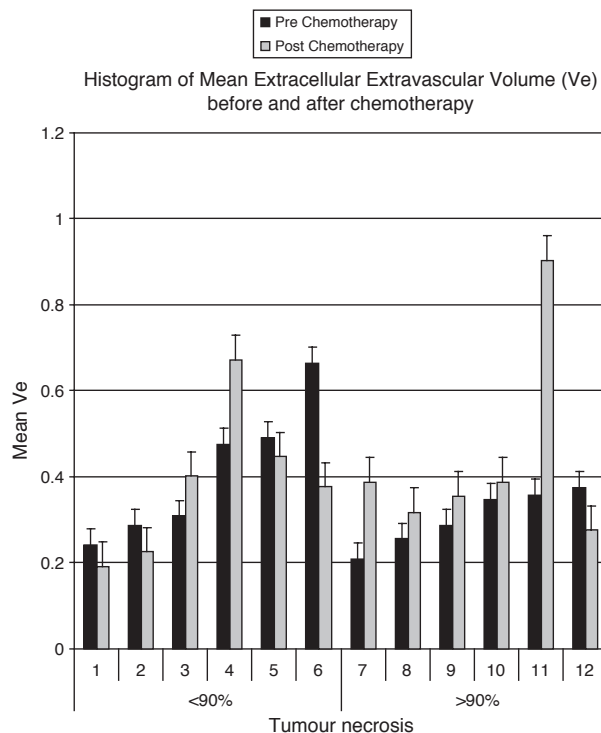


Fig. 4. Histogram demonstrating the mean extracellular extravascular volume ( $V_e$ ), for each patient, in each of the prognostic groups, before and after chemotherapy.

(14, 15, 17, 18, 22) or limited (15, 21) representative macroslices of the tumor. Recent work in animal models also suggests that DCE-MRI measures of tumor perfusion may vary by magnitudes of up to 46% in 1 hour. This instability in fluid dynamics within tumor may also be a cause of error (32, 33).

The final explanation is that mean arterial slope, endothelial permeability, extravascular extracellular space volume, and measures of any change with chemotherapy are independent of tumor necrosis. These are indirect measures of tumor necrosis that can be influenced by angiogenic factors, other than tumor viability.

Mean arterial slope and measures derived from A, have been the most commonly studied outcome measures to be compared with osseous tumor necrosis (14, 17, 21–23, 25, 30) and indicate that this tool might usefully be used to identify those patients who will fall in to the better prognostic group. Our study suggests that this may not be the case in more heterogeneous cohorts of primary bone tumors.

Unlike the results from this study, the association of endothelial permeability and tumor necrosis has been significantly correlated with good rates of tumor necrosis (14) and improved 4-year survival (34). The fact that the previously published studies included larger cohorts of 29 (14) and 31 (34)

Table 4. Tabulated results for comparison of the mean arterial slope (A), mean coefficient of endothelial permeability ( $K^{trans}$ ), and mean extracellular extravascular volume ( $V_e$ ) between patients with tumors with 90% or greater necrosis and those with less, after chemotherapy

Variable	Necrosis <90% Median (Q1 to Q3*)	Necrosis ≥90% Median (Q1 to Q3*)	Median difference (95% CI) <sup>†</sup>	P-value <sup>‡</sup>
Mean arterial slope A	2.16 (1.28 to 3.12)	1.68 (1.47 to 1.98)	−0.11 (−0.83 to 0.15)	0.59
Median arterial slope A	1.51 (0.67 to 2.12)	1.09 (0.72 to 1.2)	−0.59 (−0.67 to 0.21)	0.39
Mean $K^{trans}$	0.794 (0.48 to 1.52)	0.559 (0.48 to 0.71)	0.039 (−0.05 to 0.15)	0.24
Mean $V_e$	0.389 (0.22 to 0.50)	0.372 (0.31 to 0.52)	−0.25 (−0.68 to 0.01)	1.00

\*Lower and upper quartiles.

<sup>†</sup>95% Confidence interval for difference in medians.

<sup>‡</sup>P-value from Mann-Whitney *U* test.

Table 5. Tabulated results for comparison of the change in mean arterial slope (A), mean coefficient of endothelial permeability ( $K^{trans}$ ), and mean extracellular extravascular volume ( $V_e$ ), during chemotherapy, between patients with tumors with 90% or greater necrosis and those with less

Variable	Necrosis <90% Median (Q1 to Q3*)	Necrosis ≥90% Median (Q1 to Q3*)	Median difference (95% CI) <sup>†</sup>	P-value <sup>‡</sup>
Change in mean arterial slope (A)	−0.94 (−2.24 to 1.55)	−1.48 (−2.61 to −0.58)	0.83 (−1.24 to 3.22)	0.47
Change in median arterial slope (A)	−0.47 (−2.19 to 1.09)	−1.31 (−2.16 to −0.05)	0.94 (−1.12 to 2.74)	0.58
Change in mean $K^{trans}$	−0.33 (−1.15 to 0.68)	−0.93 (0.99 to 2.36)	0.82 (−0.50 to 2.27)	0.23
Change in mean $V_e$	−0.05 (−0.12 to 0.12)	0.07 (0.01 to 0.27)	−0.11 (−0.45 to 0.06)	0.38

\*Lower and upper quartiles.

<sup>†</sup>95% Confidence interval for difference in medians.

<sup>‡</sup>P-value from Mann-Whitney *U* test.

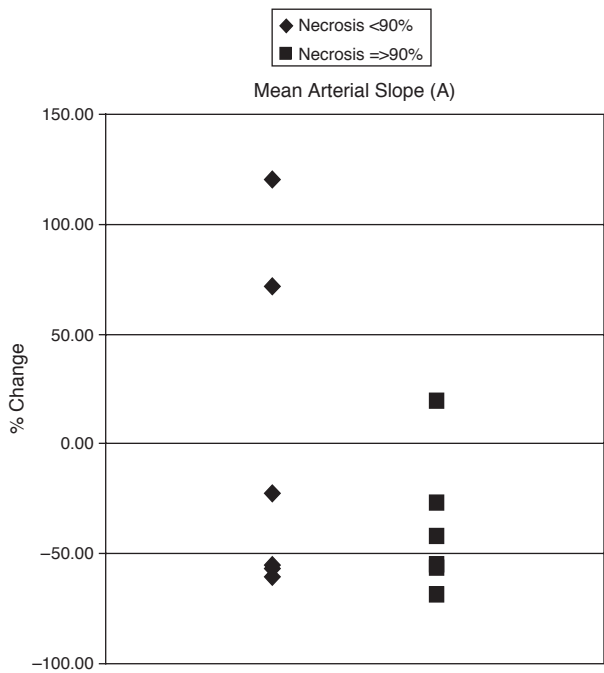


Fig. 5. Scatter plot of the percentage change in mean arterial slope (A) for the two prognostic groups demonstrating a trend toward a greater reduction in A in the tumors with 90% or greater necrosis.

patients, respectively, is likely to be an important factor in this discrepancy.

Increases in tumor volume during chemotherapy have been associated with poor responses to treatment (25, 34, 35) but this is not a universal finding

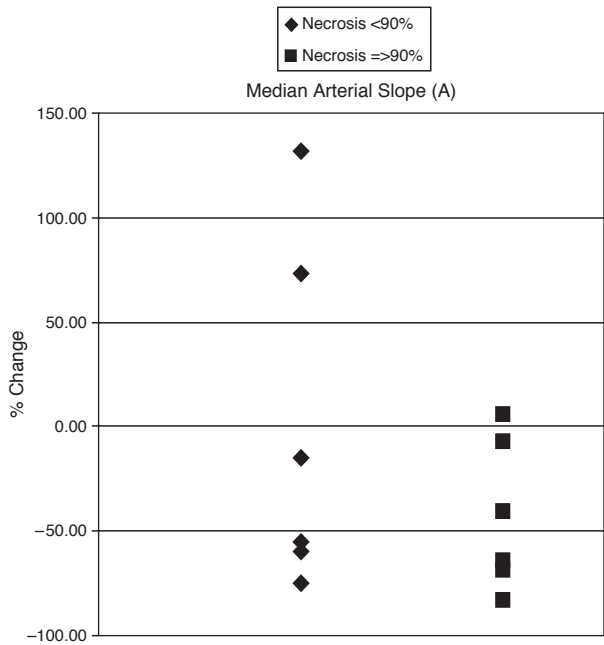


Fig. 6. Scatter plot of the percentage change in median arterial slope (A) for the two prognostic groups demonstrating a trend toward a greater reduction in A in the tumors with 90% or greater necrosis.

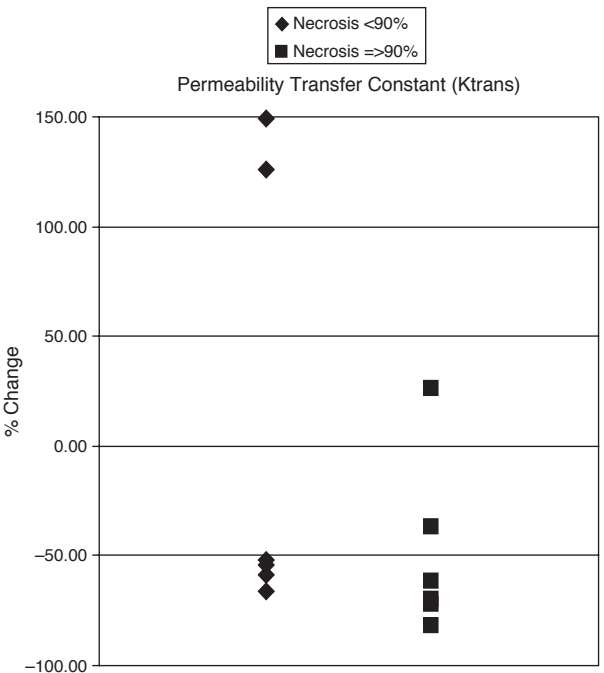


Fig. 7. Scatter plot of the percentage change in mean endothelial permeability coefficient ( $K^{trans}$ ) for the two prognostic groups again demonstrating a trend toward a greater reduction in  $K^{trans}$  in the tumors with 90% or greater necrosis.

(22). To our knowledge there has been no attempt to correlate extravascular extracellular space (EES) with necrosis rates. Any decrease in tumor

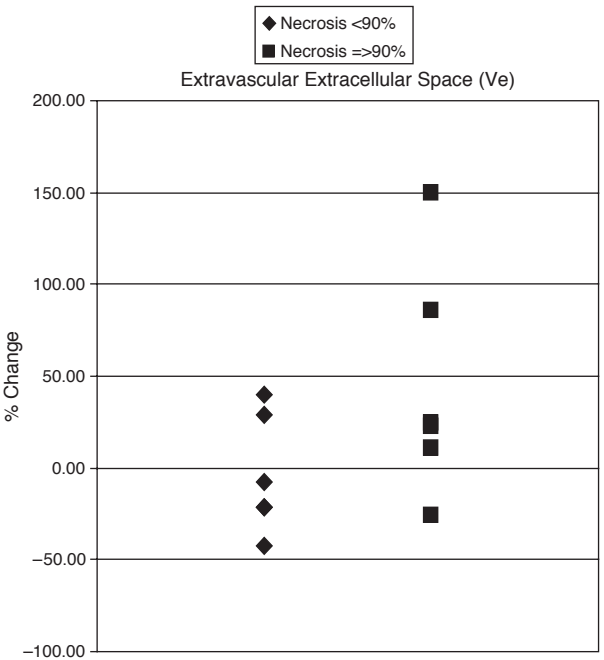


Fig. 8. Scatter plot of the percentage change in mean extravascular extracellular volume ( $V_e$ ), for the two prognostic groups demonstrating a trend toward an increase in  $V_e$  in the tumors with 90% or greater necrosis.



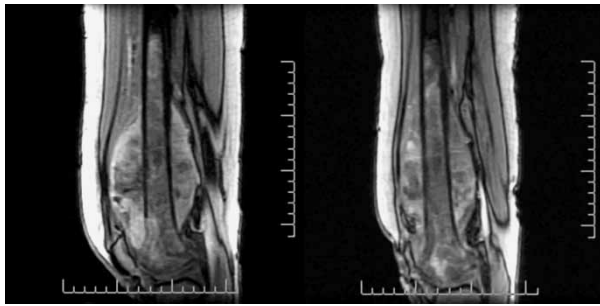


Fig. 9. Two sagittal gradient echo T1W MR images each positioned through the center of an osteosarcoma in the distal femur. The images have been obtained at peak enhancement before and after chemotherapy in a patient whose tumor demonstrated necrosis of over 90%. The images illustrate prominent perfusion both before and after chemotherapy despite there being little viable tumor after treatment.

size would be expected to have a positive correlation with a decrease in EES but this may be offset by necrosis, which may be inversely related to an increase in EES. This effect is hinted at in the results as the EES is the only parameter which increased (not statistically significant) in the group of patients with necrosis of 90% or (Fig. 8, Table 5).

Therefore, there may be more than one angiogenic mechanism influencing the results of this and other similar studies. Angiogenesis can be driven by both tumors and repair mediated pathway using similar proangiogenic factors and cytokines (36). As repair rates are variable between individuals it is easy to see how this could impact on results from DCE-MRI. This might explain some of the counter-intuitive results in the group of patients with good prognostic levels of necrosis. Angiogenesis accompanying a substantial repair process following >90% tumor necrosis might account for the solitary patient in whom all DCE-MRI markers were elevated following chemotherapy.

This possibility raises the question of the optimum timing of DCE-MRI. For those tumors that respond well with profuse early cell death (37) it might be more representative to perform the second DCE-MRI after the first or second cycle of chemotherapy, but other studies have not found this necessary to distinguish between prognostic groups (30). The relationship between the angiogenesis induced by tumors and that induced by normal repair processes is likely to be complex and until this relationship is untangled it is probably misleading to attribute all pharmacokinetic endpoint measures to changes in vascularity resulting from tumor necrosis.

In conclusion, this study indicates that pharmacokinetic modeling of single slice DCE-MRI of primary bone sarcomas does not predict total tumor necrosis in sample size of 12 patients with a

heterogeneous mix of primary bone sarcomas. Paradoxical changes in extravascular volumes and increases in perfusion in sarcomas with high necrotic fractions suggest that repair mediated angiogenesis might be a confounding factor when DCE-MRI is performed late in the chemotherapeutic course.

### Acknowledgements

This research was supported by a Marconi Medical Systems Travelling Bursary awarded by the Royal College of Radiologists.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

### References

1. Holscher HC, Bloem JL, Nooy MA, Taminiau AH, Eulderink F, Hermans J. The value of MR imaging in monitoring the effect of chemotherapy on bone sarcomas. *Am J Roentgenol* 1990;154:763–9.
2. Bhujwalla ZM, Artemov D, Glockner J. Tumor angiogenesis, vascularization, and contrast-enhanced magnetic resonance imaging. *Top Magn Reson Imaging* 1999;10:92–103.
3. Padhani AR, Husband JE. Dynamic contrast-enhanced MRI studies in oncology with an emphasis on quantification, validation and human studies. *Clin Radiol* 2001;56:607–20.
4. Hawighorst H, Libicher M, Knopp MV, Moehler T, Kauffmann GW, Kaick G. Evaluation of angiogenesis and perfusion of bone marrow lesions: role of semi-quantitative and quantitative dynamic MRI. *J Magn Reson Imaging* 1999;10:286–94.
5. Reddick WE, Taylor JS, Fletcher BD. Dynamic MR imaging (DEMRI) of microcirculation in bone sarcoma. *J Magn Reson Imaging* 1999;10:277–85.
6. Shapeero LG, Vanel D. Imaging evaluation of the response of high-grade osteosarcoma and Ewing sarcoma to chemotherapy with emphasis on dynamic contrast-enhanced magnetic resonance imaging. *Semin Musculoskelet Radiol* 2000;4:137–46.
7. van der Woude HJ, Bloem JL, Hogendoorn PC. Preoperative evaluation and monitoring chemotherapy in patients with high-grade osteogenic and Ewing's sarcoma: review of current imaging modalities. *Skeletal Radiol* 1998;27:57–71.
8. Lang P, Honda G, Roberts T, Vahlensieck M, Johnston JO, Rosenau W, et al. Musculoskeletal neoplasm: perineoplastic edema versus tumor on dynamic post-contrast MR images with spatial mapping of instantaneous enhancement rates. *Radiology* 1995;197:831–9.
9. van der Woude HJ, Verstraete KL, Hogendoorn PC, Taminiau AH, Hermans J, Bloem JL. Musculoskeletal tumors: does fast dynamic contrast-enhanced subtraction MR imaging contribute to the characterization? *Radiology* 1998;208:821–8.
10. Verstraete KL, De Deene Y, Roels H, Dierick A, Uyttendaele D, Kunnen M. Benign and malignant

- musculoskeletal lesions: dynamic contrast-enhanced MR imaging – parametric “first-pass” images depict tissue vascularization and perfusion. *Radiology* 1994; 192:835–43.
11. Meyers PA, Gorlick R, Heller G, Casper E, Lane J, Huvos AG, et al. Intensification of preoperative chemotherapy for osteogenic sarcoma: results of the Memorial Sloan-Kettering (T12) protocol. *J Clin Oncol* 1998;16:2452–8.
  12. Wunder JS, Paulian G, Huvos AG, Heller G, Meyers PA, Healey JH. The histological response to chemotherapy as a predictor of the oncological outcome of operative treatment of Ewing sarcoma. *J Bone Joint Surg Am* 1998;80:1020–33.
  13. de Baere T, Vanel D, Shapeero LG, Charpentier A, Terrier P, di PM. Osteosarcoma after chemotherapy: evaluation with contrast material-enhanced subtraction MR imaging. *Radiology* 1992;185:587–92.
  14. Dyke JP, Panicek DM, Healey JH, Meyers PA, Huvos AG, Schwartz LH, et al. Osteogenic and Ewing sarcomas: estimation of necrotic fraction during induction chemotherapy with dynamic contrast-enhanced MR imaging. *Radiology* 2003;228:271–8.
  15. Egmont-Petersen M, Hogendoorn PC, van der Geest RJ, Vrooman HA, van der Woude H, Janssen JP, et al. Detection of areas with viable remnant tumor in post-chemotherapy patients with Ewing's sarcoma by dynamic contrast-enhanced MRI using pharmacokinetic modeling. *Magn Reson Imaging* 2000;18:525–35.
  16. Hanna SL, Parham DM, Fairclough DL, Meyer WH, Le AH, Fletcher BD. Assessment of osteosarcoma response to preoperative chemotherapy using dynamic FLASH gadolinium-DTPA-enhanced magnetic resonance mapping. *Invest Radiol* 1992;27:367–73.
  17. Reddick WE, Bhargava R, Taylor JS, Meyer WH, Fletcher BD. Dynamic contrast-enhanced MR imaging evaluation of osteosarcoma response to neoadjuvant chemotherapy. *J Magn Reson Imaging* 1995;5:689–94.
  18. van der Woude HJ, Bloem JL, Verstraete KL, Taminiau AH, Nooy MA, Hogendoorn PC. Osteosarcoma and Ewing's sarcoma after neoadjuvant chemotherapy: value of dynamic MR imaging in detecting viable tumor before surgery. *AJR Am J Roentgenol* 1995;165:593–8.
  19. Huvos AG. Osteogenic sarcoma: pathological assessment of preoperative (neoadjuvant) chemotherapy. Philadelphia: WB Saunders; 1991. p. 122–8.
  20. Rosen G, Caparros B, Huvos AG, Kosloff C, Nirenberg A, Cacavio A, et al. Preoperative chemotherapy for osteogenic sarcoma: selection of postoperative adjuvant chemotherapy based on the response of the primary tumor to preoperative chemotherapy. *Cancer* 1982;49: 1221–30.
  21. Erlemann R, Sciuk J, Bosse A, Ritter J, Kusnierz-Glaz CR, Peters PE, et al. Response of osteosarcoma and Ewing sarcoma to preoperative chemotherapy: assessment with dynamic and static MR imaging and skeletal scintigraphy. *Radiology* 1990;175:791–6.
  22. Fletcher BD, Hanna SL, Fairclough DL, Gronemeyer SA. Pediatric musculoskeletal tumors: use of dynamic, contrast-enhanced MR imaging to monitor response to chemotherapy. *Radiology* 1992;184:243–8.
  23. Kawai A, Sugihara S, Kunisada T, Uchida Y, Inoue H. Imaging assessment of the response of bone tumors to preoperative chemotherapy. *Clin Orthop Relat Res* 1997;337:216–25.
  24. Uhl M, Saueressig U, van Buiren M, Kontny U, Niemeyer C, Köhler G, et al. Osteosarcoma: preliminary results of in vivo assessment of tumor necrosis after chemotherapy with diffusion- and perfusion-weighted magnetic resonance imaging. *Invest Radiol* 2006;41: 618–23.
  25. Bonnerot V, Charpentier A, Frouin F, Kalifa C, Vanel D, Di PR. Factor analysis of dynamic magnetic resonance imaging in predicting the response of osteosarcoma to chemotherapy. *Invest Radiol* 1992;27: 847–55.
  26. Tofts PS, Kermode AG. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. *Magn Reson Med* 1991;17:357–67.
  27. Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusible tracer: standardized quantities and symbols. *J Magn Reson Imaging* 1999;10:223–32.
  28. Brix G, Semmler W, Port R, Schad LR, Lauer G, Lorenz WJ. Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. *J Comput Assist Tomogr* 1991;15:621–8.
  29. Hoffmann U, Brix G, Knopp MV, Hess T, Lorenz WJ. Pharmacokinetic mapping of the breast: a new method for dynamic MR mammography. *Magn Reson Med* 1995;33:506–14.
  30. Ongolo-Zogo P, Thiesse P, Sau J, Desuzinges C, Blay JY, Bonmartin A, et al. Assessment of osteosarcoma response to neoadjuvant chemotherapy: comparative usefulness of dynamic gadolinium-enhanced spin-echo magnetic resonance imaging and technetium-99m skeletal angioscintigraphy. *Eur Radiol* 1999;9:907–14.
  31. Tofts PS. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J Magn Reson Imaging* 1997;7: 91–101.
  32. Baudelet C, Cron GO, Gallez B. Determination of the maturity and functionality of tumor vasculature by MRI: correlation between BOLD-MRI and DCE-MRI using P792 in experimental fibrosarcoma tumors. *Magn Reson Med* 2006;56:1041–9.
  33. Baudelet C, Cron GO, Ansiaux R, Crockart N, DeWever J, Feron O, et al. The role of vessel maturation and vessel functionality in spontaneous fluctuations of T2\*-weighted GRE signal within tumors. *NMR Biomed* 2006;19:69–76.
  34. Reddick WE, Wang S, Xiong X, Glass JO, Wu S, Kaste SC, et al. Dynamic magnetic resonance imaging of regional contrast access as an additional prognostic factor in pediatric osteosarcoma. *Cancer* 2001;91: 2230–7.
  35. Holscher HC, Bloem JL, Vanel D, Hermans J, Nooy MA, Taminiau AH, et al. Osteosarcoma: chemotherapy-induced changes at MR imaging. *Radiology* 1992;182: 839–44.
  36. Li WW. Tumor angiogenesis: molecular pathology, therapeutic targeting, and imaging. *Acad Radiol* 2000; 7:800–11.
  37. Preda A, Wielopolski PA, Ten Hagen TL, van Vliet M, Veenland JF, Ambagtsheer G, et al. Dynamic contrast-enhanced MRI using macromolecular contrast media for monitoring the response to isolated limb perfusion in experimental soft-tissue sarcomas. *MAGMA* 2004;17: 296–302.

## (ii) Diagnostic plain film radiology of the failing hip replacement

Andoni P Toms

Rajesh Botchu

John F Nolan

### Abstract

Conventional radiographs are the main imaging tool for investigating the failing total hip replacement (THR). THRs most commonly fail because of aseptic loosening. The aetiology of aseptic loosening lies in a number of mechanical and biological processes which can result in a variety of different radiological features. These may be used to support or confirm a diagnosis of a loose or failing THR. The reliability and accuracy of some, but not all, of these radiographic signs have been evaluated in the scientific literature. This paper reviews and illustrates these radiographic signs. The pathophysiology for each sign is explained and the evidence supporting the accuracy and reliability is discussed.

**Keywords** arthroplasty; failure; hip; loosening; radiograph

### Introduction

Over 66,000 total hip replacements (THRs) were performed in the UK in 2008.<sup>1</sup> After ten years approximately one third of components will have become loose<sup>2</sup> and 10% will have been revised,<sup>3,4</sup> usually for aseptic loosening. Therefore by 2018 approximately 20,000 loose THRs may require radiographic evaluation. Plain radiographs have been the mainstay of follow up for THR for 40 years. The aim of this paper is to review the diagnostic features on conventional radiography that indicate a failing THR. Each radiographic feature will be considered separately. The aetiology of each feature will be described and, where available, the diagnostic accuracy of each radiographic sign will be considered.

**Andoni P Toms FRCR** is a Consultant Radiologist within the Radiology Department, Norfolk and Norwich University Hospital, Norwich, UK.

**Rajesh Botchu MBBS MS(Orth) MRCS(Ed) MRCSI** is a Radiology StR within the Radiology Department, Norfolk and Norwich University Hospital, Norwich, UK.

**John F Nolan FRCS(Orth)** is a Consultant Orthopaedic Surgeon, Orthopaedic Department, Norfolk and Norwich University Hospital, Norwich, UK.

### Radiographic technique

Standard post-operative radiographs should include an AP of the pelvis and a lateral of the THR. The first post-operative radiograph acts as a baseline investigation with which all subsequent radiographs can be compared. The timing of these first post-operative radiographs varies between surgeons but it is probably best to perform them at the first outpatient visit following surgery. Radiographs that conform to accepted standards<sup>5</sup> can be most reliably produced using a radiographic table with patients who are relatively mobile and pain-free, as opposed to patients who are in hospital beds, with abduction wedges, in the first day or two after surgery. An iliac-oblique view (distally centred Judet) of the THR may be a useful adjunct to, or replacement of, the standard lateral projection.<sup>6</sup> The iliac-oblique view appears to be the optimal for assessing the cement bone interface of the acetabular component.<sup>7</sup> The timing of subsequent follow-up radiographs varies but there seems to be little to be gained by repeating the investigation in the early post-operative course.<sup>8</sup>

### Bone loss/osteolysis

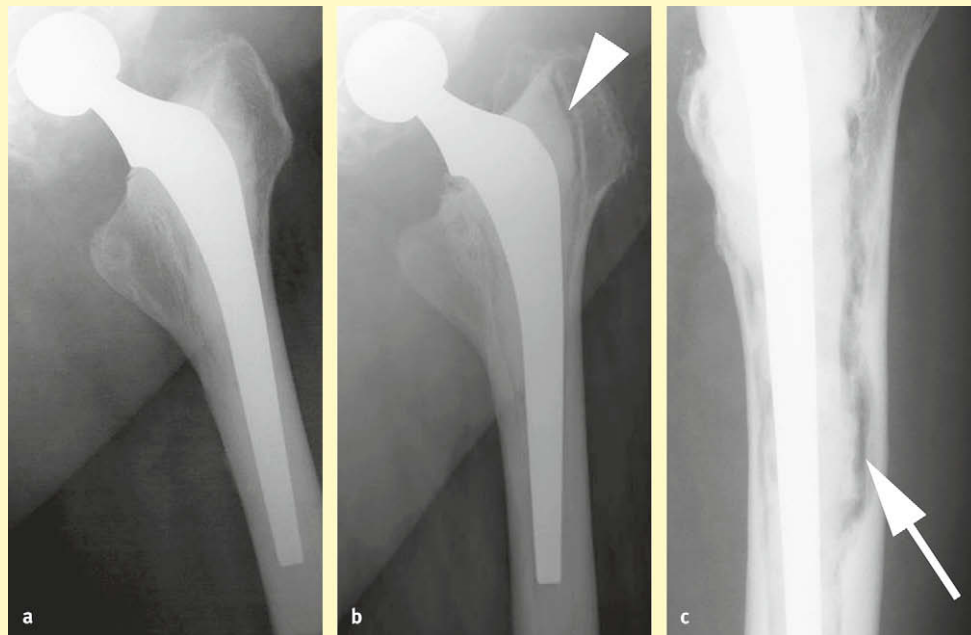
There are a number of different causes of loss of bone stock around a THR. These include infection, small particle disease and tumour but the most common causes are benign and are associated with aseptic loosening.

#### Normal

Preparation of the proximal femur and the exothermic process of setting methylmethacrylate cement causes thermal, chemical and traumatic necrosis of bone adjacent to the cement mantle. It is not uncommon for this layer of dead bone to be resorbed and replaced with an investing layer of fibrous tissue which creates a thin lucency (<2 mm thick) at the cement–bone interface on conventional radiographs. This lucency appears within 3 months of operation, is uniform in thickness and is demarcated by a thin sclerotic margin (Figures 1a and b).

#### Stress shielding

Stress shielding is the result of local osteopaenia, usually in the proximal femur, following THR particularly with long stemmed uncemented femoral prostheses (Figures 2a and b). Stress from loading through the joint is transmitted to the cortex of the femoral diaphysis through the femoral prosthesis and the cement mantle bypassing the trochanteric trabeculae, and therefore stress shielding is associated with cortical hypertrophy at the point of distal fixation. With a reduction in stress loading comes a net increase in resorption,<sup>9</sup> particularly in the greater trochanter,<sup>10</sup> which manifests as decreased mineralisation of the trabeculae and cortex on conventional radiographs. Bone mineralisation reaches a nadir at about two years after surgery and then increases to near normal levels by ten years after surgery. While initially considered a complication of THR,<sup>11</sup> particularly associated with wide and stiff femoral prostheses, there appears to be no association with adverse long term outcomes.<sup>12,13</sup> Inter-rater reliability for identifying stress shielding on conventional radiographs is excellent only if bone mineral loss is greater than 70%.<sup>14</sup> For less severe loss of bone stock CT osteodensitometry is required in order to achieve satisfactory measures of reliability.<sup>15</sup>



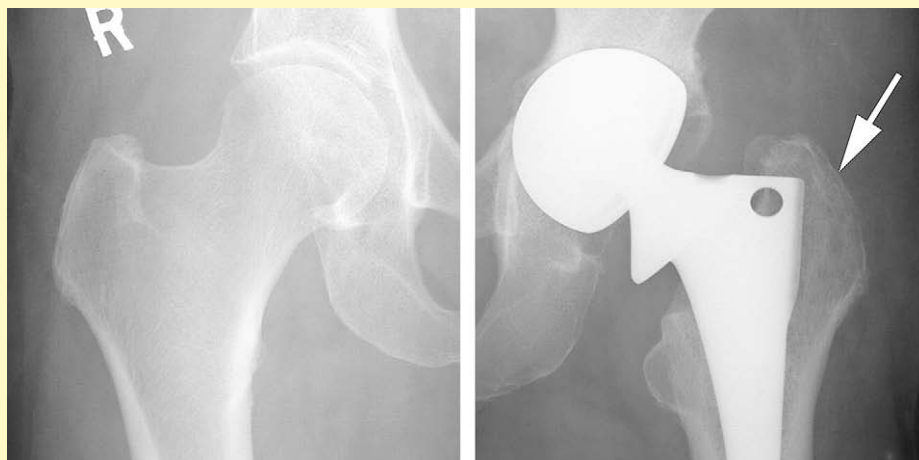
An immediate post-operative AP radiograph of a left THR **a** demonstrates a contiguous bone cement interface. **b** 6 months later there is a well defined linear lucency surrounding the cement mantle with a thin sclerotic margin (arrowhead) at the interface with bone. These appearances remained stable at follow up to 5 years consistent with a fibrous bone-cement interface. **c** In comparison thick irregular lucency (arrow) surrounding the cement mantle indicates osteolysis.

**Figure 1**

Narrowing of the femoral neck has been reported as a normal finding in hip resurfacing, both cemented and uncemented, which stabilises within 6 years of operation.<sup>16</sup> The authors suggest that this is also a result of osteolysis and cortical remodeling resulting from stress shielding. However this is not a universally held belief. Severe narrowing of the femoral neck is also a recognised feature of resurfacing prostheses which some argue is the result of metal ionosis. The resorption can be severe enough to compromise the mechanical integrity of

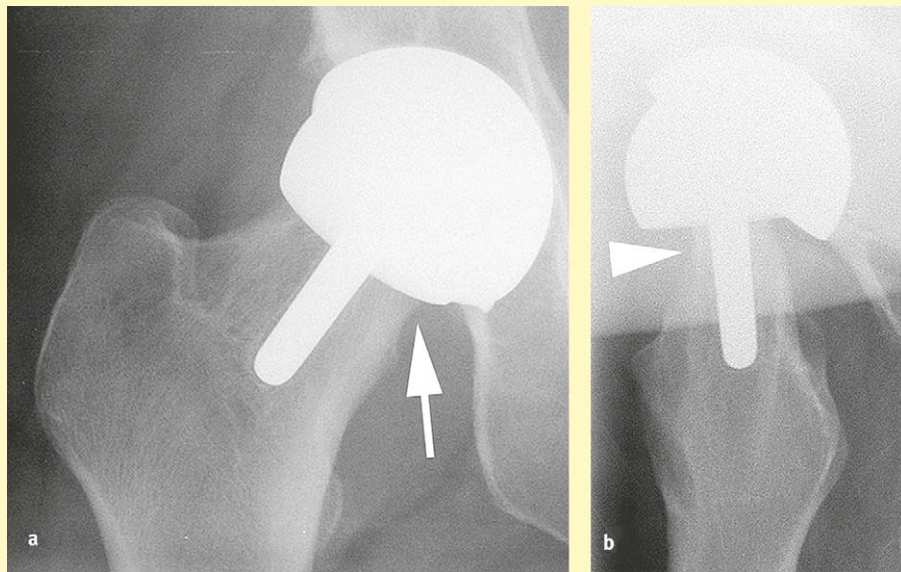
the femoral neck (Figure 3). Compromised vascularity has also been suggested as a cause for this resorption but has not been substantiated.

Proximal femoral osteoporosis can occur in the medulla interposed between the cement mantle and endosteum and should be differentiated from osteolysis. Osteoporosis presents as progressively reduced medullary attenuation surrounding the cement mantle but with preservation of secondary trabeculae and an absence of a demarcating line of sclerosis (Figure 4). Proximal



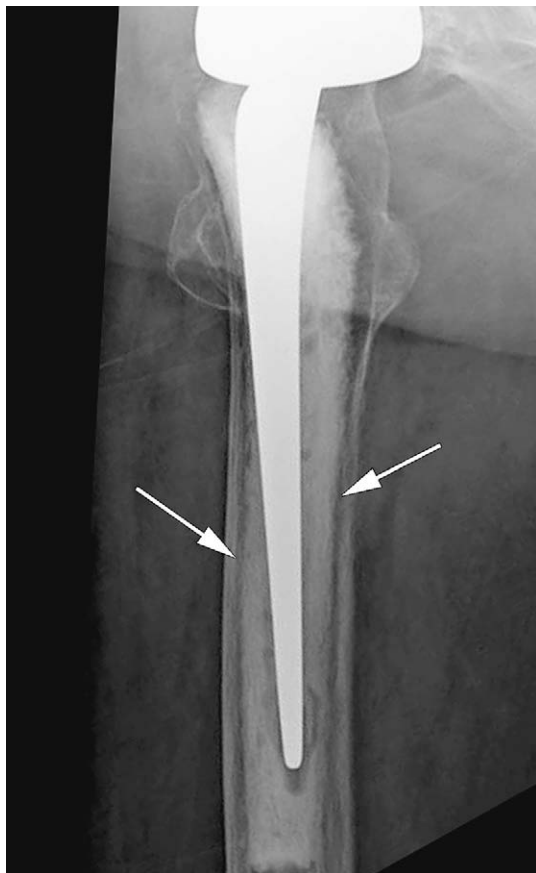
AP radiographs of the right and left hips of a 69 year old patient with a left distally fixed uncemented bipolar hemiarthroplasty (Bateman) demonstrating demineralisation of the trabecula and thinning of the cortex (arrow) of the left greater trochanter caused by stress shielding.

**Figure 2**



AP **a** and lateral **b** projections of the right hip in a patient with a Birmingham resurfacing total hip arthroplasty demonstrating narrowing of the femoral neck (arrow) which is much more apparent on the lateral view where the cortical resorption is so severe (arrowhead) that the femoral neck is at risk of fracture.

**Figure 3**



**Figure 4** Lateral radiograph of a right bipolar hemiarthroplasty demonstrating asymptomatic osteoporosis (arrows) adjacent to a normal cement mantle which should not be misinterpreted as osteolysis.

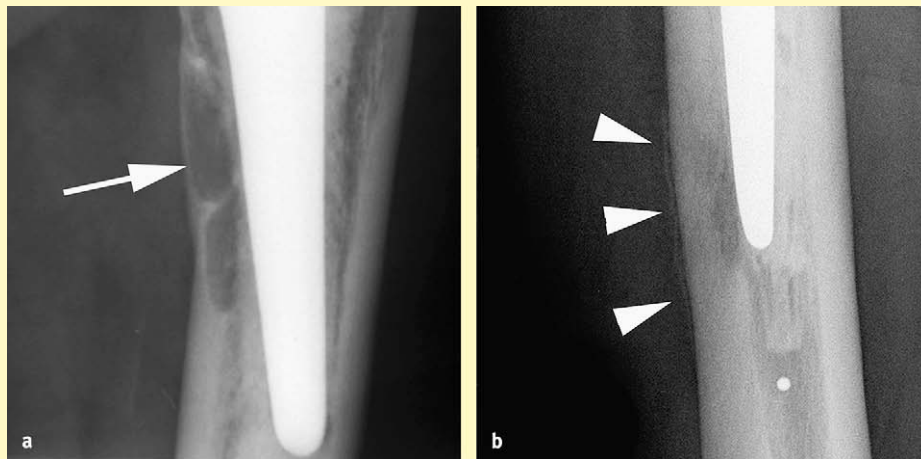
femoral peri-prosthetic osteoporosis is not known to be associated with symptoms or failure of the prosthesis.

### Aseptic loosening

Several overlapping processes that lead to aseptic loosening (principally mechanical and biological) are responsible for the most common radiographic appearances. Most commonly osteolysis is associated with polyethylene small particle disease, which typically forms foreign body granulomas,<sup>17</sup> in combination with elevated hydrodynamic pressures within the effective joint space that can dissect a plane along the prosthesis-cement interface.<sup>18</sup> This can result in linear, geographic or erosive patterns of osteolysis<sup>17</sup> which typically appear 7 to 8 years after surgery.<sup>19</sup> Linear osteolysis can be differentiated from a fibrous bone-cement interface if it is more than 2 mm thick, if its thickness is uneven or if it has progressed with time (Figure 1c). Geographic osteolysis, caused by small particle disease, is well defined with a thin sclerotic margin indicating the slow non-aggressive enlargement of the underlying granuloma. The description of geographic osteolysis suggests that the lesion is contained within bone, although this may often be expanded and thin. Contained osteolysis is amenable to treatment with morcelised bone graft at revision surgery (Figures 5a and b). Erosive osteolysis describes osteolysis that is not contained by bone and typically occurs at the margins of the joint such as the medial calcar and usually requires repair with a mesh to contain the bone graft (Figure 6).

The presence of osteolysis alone on conventional radiographs does not necessarily indicate aseptic loosening. However sensitivity for loosening is proportional to the extent of lysis around the cement mantle. If there is no lucency around the cement mantle of an acetabular component (including an iliac-oblique projection)





**a** AP radiograph of a femoral prosthesis demonstrating well defined lucencies with thin sclerotic margins scalloping the endosteum (arrow) caused by small particle granulomas. **b** A focal area of aggressive osteolysis indicated by ill defined margins and periosteal new bone (arrowheads). In this case the appearances are due to osteomyelitis but could be due to metastases.

**Figure 5**

then it is well fixed. If there is a lucency that surrounds the whole of cement mantle then there is a 94% incidence of the acetabular cup being loose.<sup>20</sup> In between these extremes the severity of lysis is proportionately associated with loosening. The rate of change of lysis is also important. Rapid progression of lysis, particularly early on in the post-operative course, appears to be a predictor for early aseptic loosening.<sup>21</sup>

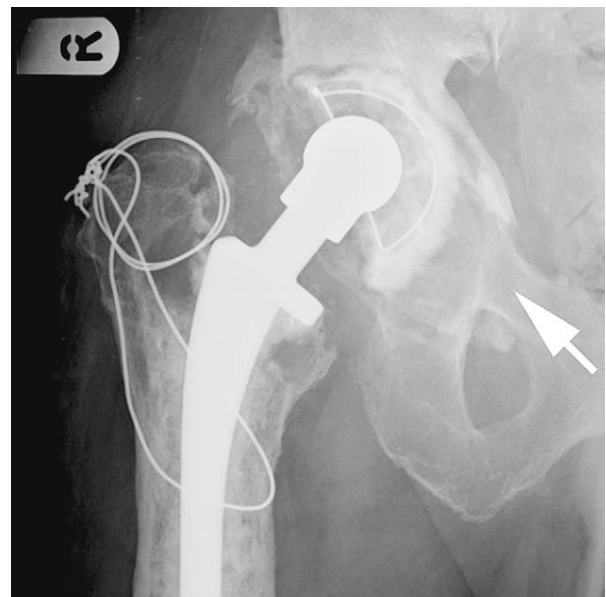
Specific osteolysis of the teardrop shadow<sup>22</sup> indicates disruption of the quadrilateral plate, or medial wall, of the acetabulum<sup>23</sup> (Figure 7). Similarly interruption of the iliopubic or ilioischial (Kohler's) lines, or a combination of the two, is a specific indicator of medial wall disruption with specificities of over 90% but

sensitivities of between 50% and 75%. Ballooning of both these lines in combination is also a reasonably useful sign of medial wall disruption with a specificity of over 80%<sup>24</sup> (Figure 8). Resorption of the medial femoral calcar may be due either to stress shielding or small particle disease but is usually not progressive and is not commonly associated with loosening.<sup>25</sup>

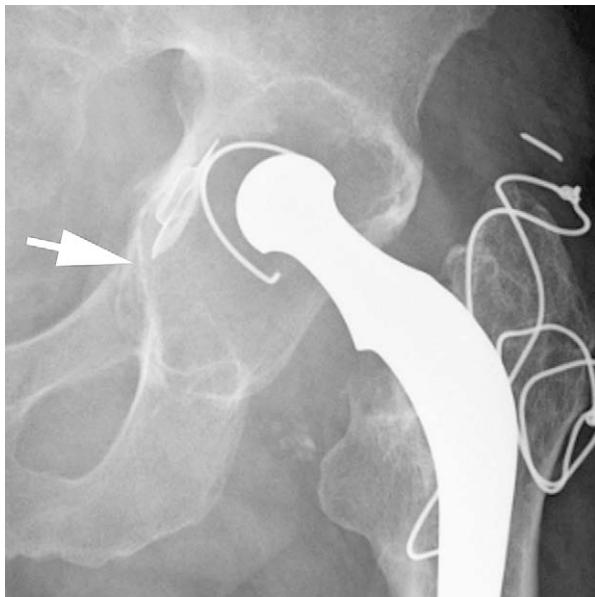
The detection of osteolysis with conventional radiographs has a high specificity, over 90%, but sensitivity is variable. The tendency is to overestimate the degree of osteolysis on the femoral side of the prosthesis<sup>26</sup> and underestimate the loss on the iliac side of the joint.<sup>27</sup> It is easier to detect a lytic lesion in the ilium



**Figure 6** A cemented THR demonstrating resorption of the medial calcar (arrow) by small particle disease.



**Figure 7** AP radiograph of a failing revision THR demonstrating uncontained osteolysis and an absent teardrop (arrow). At operation an osteolytic cavity extended through the medial wall and extended along the medullary canal of the superior pubic ramus.



**Figure 8** Extensive osteolysis around a cemented (barium free) THR demonstrating ballooning of the ilioischial and iliopubic lines (arrow) caused by extensive periacetabular small particle disease.

than in the ischium and acetabular rim on a single view particularly when lesions are smaller than 10 cm<sup>3</sup><sup>28</sup> but detection rates can be increased to 94% by using four different radiographic views (including a Judet iliac-oblique).<sup>27</sup> Inter-observer reliability for the detection of osteolysis from a single radiograph is poor, although intra-observer reliability can be good or excellent. However, inter-observer reliability can be improved by reviewing a series of radiographs rather than single examinations.<sup>29</sup>

Advanced acetabular osteolysis can, uncommonly, cause pelvic discontinuity where the ilium becomes separated from the ischium and pubis (Figure 9). This prevalence of pelvic discontinuity is associated with gender, being more common in women, patients with rheumatoid arthritis and significant complication rates following revision surgery.<sup>30</sup> Radiographic signs of pelvic discontinuity include fracture lines through the anterior and posterior columns and evidence of movement of the inferior hemipelvis, either medial translation or rotation.<sup>30</sup> There is no published evidence of the sensitivity or specificity of conventional radiographs in detecting pelvic discontinuity, probably because numbers of patients in any one centre are likely to be small. This is not important because CT is likely to replace conventional radiographs for the assessment of pelvic discontinuity.

Aggressive osteolysis is defined as focal bone loss with a poorly defined margin and may be accompanied by permeative changes (poorly defined small lucencies) in the adjacent marrow (Figure 5b). The most common causes for this relatively uncommon finding are osteomyelitis and metastases. When infection presents with permeative changes and periosteal new bone there is often accompanying infection in the extra-osseous soft tissues. Most patients with an infected THR will not have specific radiographic features of osteomyelitis and the appearances are often indistinguishable from aseptic loosening.<sup>31,32</sup> The periprosthetic infection in these cases is low grade and chronic being confined to a glycocalyx which occupies the areas



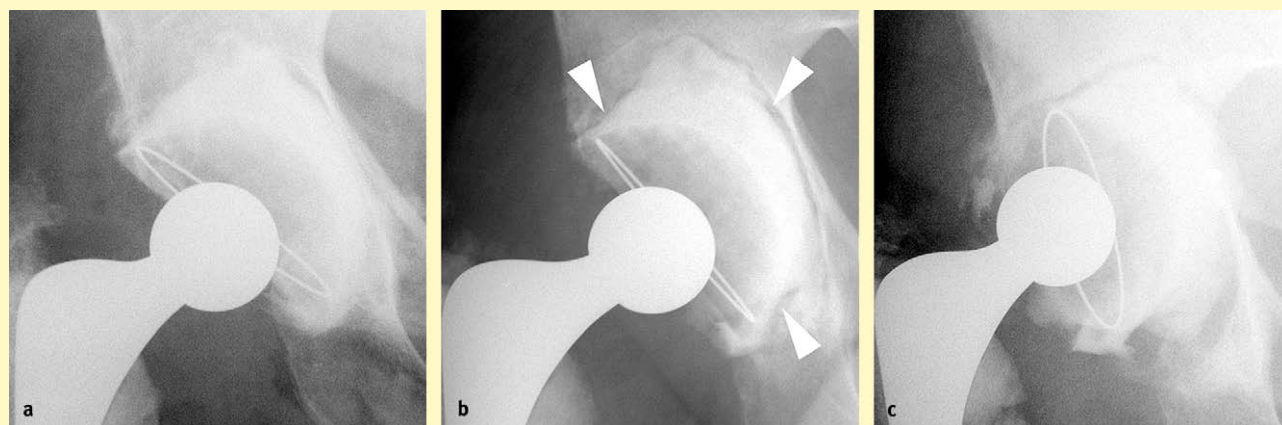
**Figure 9** Antero-posterior radiograph of a right THR demonstrating pelvic discontinuity with a sagittal fracture line through the roof of the acetabulum (arrowhead: required CT for confirmation). There is also an eccentrically placed femoral head indicating polyethylene wear and medial calcar resorption (arrow).

of osteolysis. Normal radiographs certainly do not exclude the presence of infection.

### Migration

The only definite radiographic sign of loosening is the change in position of a prosthesis from one radiograph to another.<sup>33–37</sup> For most patients this is assessed using standard interval radiography which can demonstrate movement but is probably inadequate for detecting early mechanical loosening.<sup>35,38</sup> Roentgen stereophotogrammetric analysis allows subtle differences in patient and radiographic positioning to be compensated for and is therefore more sensitive to early movement.<sup>39,40</sup> Other definitions of loosening, for instance radiolucency demarcating the whole of the margin of the acetabular component,<sup>41</sup> are highly effective statistical predictors but not direct evidence of loosening. Early osteolysis adjacent to the superolateral acetabulum (zone 1) is strongly associated with progression to acetabular loosening<sup>42,43</sup> (Figure 10).

The acetabular cup is the more common of the two components to migrate and become loose. Movement can be identified by measuring the acetabular inclination from a transverse



**a** There is a small focal lucency projected superolateral to the cement mantle (note poor cement interdigitation) 1 year after surgery. **b** 3 years after surgery the lucency has extended around the cement mantle (arrowheads), an indirect sign of loosening, and there has been a small rotational movement of the cup confirming that it is loose. **c** Within 1 year the acetabular cup rotates significantly.

**Figure 10**

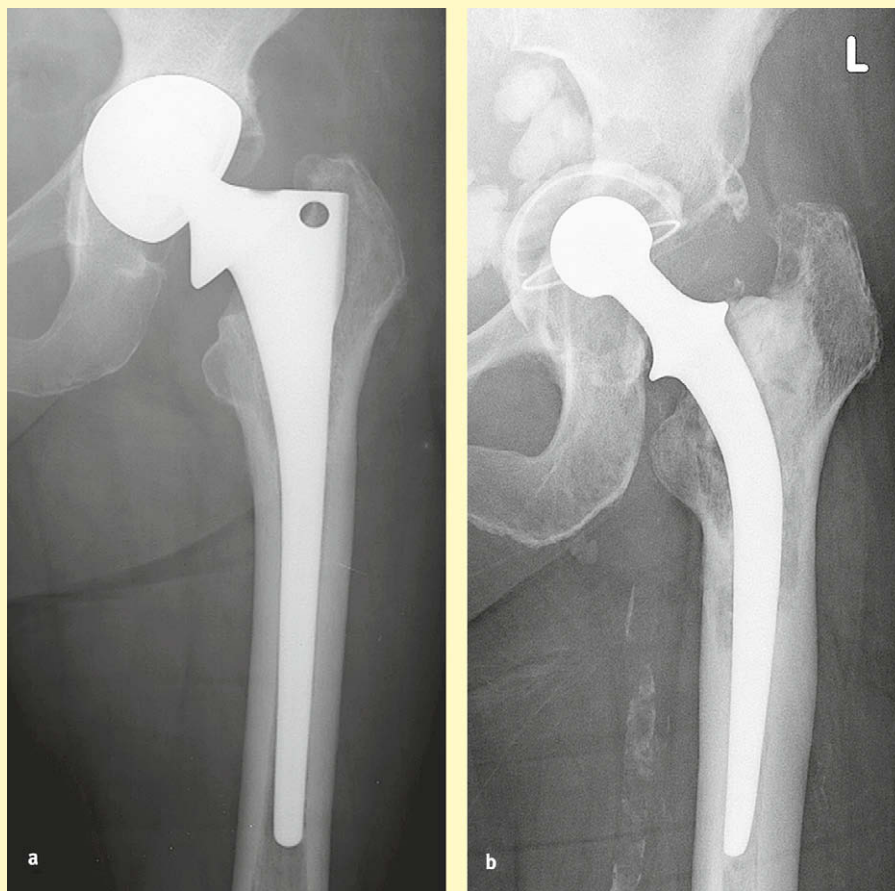
baseline,<sup>44</sup> which is usually a line drawn between the ischial tuberosities<sup>45</sup> or the inferior margins of the tear drop opacities<sup>46,47</sup> and demonstrating a difference in position between two radiographs. The absolute position of an acetabular cup does not indicate loosening nor, in isolation, is it associated with complications.<sup>48</sup> Although there are criteria for position and orientation of the pelvis for the optimal post-operative radiograph<sup>5</sup> in practice if pelvic tilt or rotation are less than  $10^\circ$  this does not have a significant effect on the projected angle of acetabular inclination.<sup>49,50</sup> Inter-observer reliability for measuring acetabular inclination has generally been reported as good<sup>33,51,52</sup> but the ability of observers to detect differences between radiographs is less encouraging. Inter-observer correlation coefficients for reporting loosening of the acetabular cup have been reported as only moderate (0.49–0.63)<sup>53</sup> and there is no clear evidence of what the minimal amount of detectable change in position on serial radiographs is required for reliable observations. What does appear to be clear is that the reliability of planar imaging to assess changes in anteversion is even less reliable and should probably only be attempted with CT or other three-dimensional imaging techniques.<sup>51</sup> Wear of the polyethylene liner of the acetabular cup by the femoral head will eventually lead to a loose articulation (Figure 11). The amount of polyethylene wear can be measured by choosing the shortest radius from the centre of the femoral head, on an AP radiograph, to the outer margin of the acetabular cup. This seems to be considered a reliable technique but it is not clear how much polyethylene wear is significantly associated with loosening or symptoms.<sup>54</sup>

As well as the acetabular cup rotating within the acetabulum the prosthesis can also migrate proximally or medially, or capsize (Figure 12). This occurs most frequently in patients with inflammatory arthritis compared to those with osteoarthritis.<sup>55</sup> Proximal migration, associated with a valgus stem position, appears to be the more common of the two directions with measurements taken from the inter-tear drop line. Medial migration (protrusio acetabuli prosthetica), associated with a varus stem position, is usually defined as migration medial to the ilio-ischial line.<sup>56,57</sup> It



**Figure 11** Marked superior migration of the head of a Charnley femoral prosthesis. The polyethylene erosion (arrow) has compromised the structural viability of the cup which has fractured (arrowhead).





**a** AP radiograph of a Bateman bipolar hemiarthroplasty demonstrating predominantly cranial migration of the prosthetic head which has eroded through the roof of the acetabulum. **b** Protrusio acetabuli prosthetic; a Stanmore THR has migrated medial to the ilioischial line. Cement in the pelvis indicates that there was a defect in the medial wall at the time of surgery. There is also loosening, osteolysis and cortical hypertrophy around the femoral stem.

**Figure 12**

is not certain how reliable this sign is but some have found defining the centre of rotation, using horizontal (x) and vertical (y) coordinates from the teardrop, to be useful for follow up studies of cup migration.<sup>58</sup>

The femoral stem angle can be measured on either the AP or the lateral views and requires a line to be drawn along the long axis of the stem and compared with the long axis of the proximal femoral diaphysis. Alternatively the position of the tip of the stem within the femoral medullary canal can be measured relative to either the cortex or endosteum. Although the identification of femoral loosening appears to be reliable with good correlation coefficients (0.74–0.8) between observers<sup>53</sup> the reliability of femoral stem angle measurements is not known. Whether or not these angles are particularly relevant is another question. Valgus stem migration is a recognised normal finding in certain prostheses. Mechanical loosening of the femoral prosthesis appears to be predominantly due to torque applied to the neck which results in rotational instability of the stem.<sup>59</sup> Measures of varus or valgus migration are unlikely to detect rotational loosening until it is quite advanced.

### Cement mantle

The cement mantle has failed when loss of integrity allows movement of the prosthesis. This is more common in the femoral than the acetabular component. The femoral cement mantle can be classified according to the number and extent of the defects in the mantle present after surgery. Mantles that lack cement over at least 50% of the margin or fail to cover the tip of the femoral prosthesis (Barrack C and D<sup>60</sup>) are strongly associated with early failure of the THR<sup>61</sup> but do not in themselves indicate a loose prosthesis. The THR can fail from either circumferential or longitudinal defects in the cement mantle. A circumferential defect leads to separation of the proximal and distal segments of the cement which allows the femoral prosthesis to subside. The distal segment of cement mantle is then displaced distally leaving a clear fracture line (Figure 13). Proximally the shoulder of the prosthesis may leave an empty impression in the adjacent cement as it migrates distally (Figure 14). This last sign, on its own, is not diagnostic of a loose prosthesis because some femoral components, such as those with collarless polished tapered stems, are designed to subside at

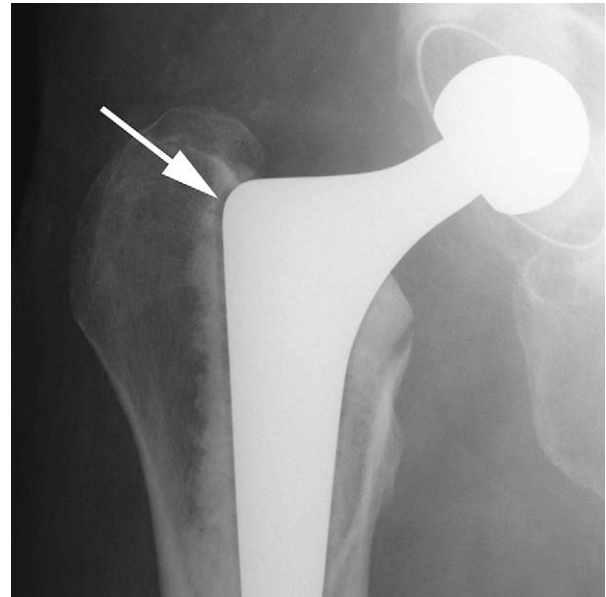


**Figure 13** Circumferential fracture of the femoral cement mantle (arrow) with distal displacement of the distal cement fragment indicating distal migration of the femoral component, resulting in apposition of the proximal femur and acetabulum.

predictable rates for several years following surgery.<sup>62</sup> Longitudinal defects in the mantle will allow the femoral component to rock in flexion and extension, or abduction and adduction, exacerbating osteolysis and eventually causing a fracture, typically of the medial calcar or around the tip of the prosthesis.

### Heterotopic ossification

Heterotopic ossification (HO) has a reported prevalence of between 26% and 34%.<sup>63–65</sup> While the prevalence does not appear to be related to the type of THR<sup>66</sup> there is some evidence to suggest that the prevalence may be higher in resurfacing prostheses.<sup>67</sup> There is also evidence that the position of HO is related to the surgical approach.<sup>65</sup> The diagnostic radiographic description of HO is of soft tissue opacities with a discrete cortex and a medulla that often contains slightly disorganised trabeculae. Without identifiable corticomedullary differentiation heterotopic bone cannot be differentiated from other causes of soft tissue calcification. The severity of HO varies; mild disease, with a few scattered small foci of bone, is the most common and is not clinically significant.<sup>64</sup> More extensive disease is less common but is associated with reduced functional outcomes<sup>68,67</sup> (Figure 15). The extent of HO can be reasonably reliably graded



**Figure 14** Migration of the femoral component distally or medially leaves a characteristic bare area of cement (arrow). Similar appearances in some polished collarless tapered stems are normal as 1–2 mm of settling of the prosthesis is allowed.



**Figure 15** Extensive (Brooker grade 4) heterotopic ossification that traverses the joint bridging the ilium and femur and was the cause of decreased range of movement in the patient.

using one of several systems, which have evolved from and since Brooker's original description.<sup>63</sup> Inter-observer reliability for the original Brooker grading system has been reported with kappa values as low as 0.43,<sup>69</sup> which is satisfactory, with more recent modifications producing good inter-observer reliability measures of between  $r = 0.69$  and  $0.8$ .<sup>70,71</sup>

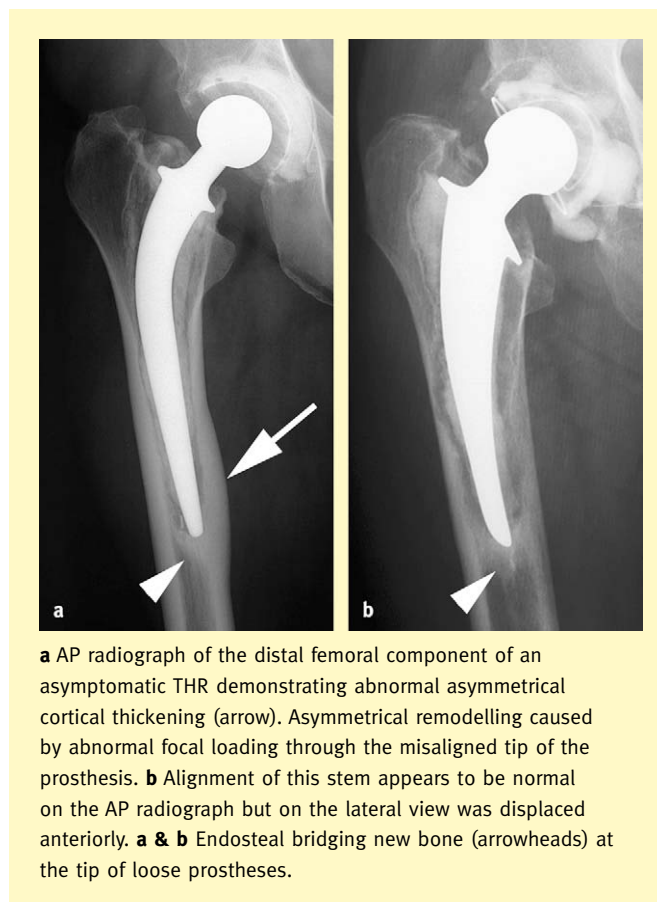
Another cause for periprosthetic soft tissue opacification is metallosis. This was a common complication of first generation metal-on-metal articulations but is uncommon with polyethylene-on-metal THR where it occurs after failure of the polyethylene cup so that the femoral head articulates with the metal liner of the acetabular component releasing large amounts of metal debris in to the joint. This can settle on the deep surface of the joint capsule producing cloud-like opacities which can be mistaken for HO (Figure 16).<sup>72</sup>

### Cortical remodelling and failure

Cortical remodelling is a normal response to THR. Within the first four years after surgery it is normal to see an increase in cortical thickness as periosteal new bone is laid down around the cement mantle in response to loading forces that are redistributed from the cement to the cortex.<sup>73</sup> On the other hand asymmetrical or focal cortical thickening is abnormal. It most commonly occurs around the tip of the femoral prosthesis (Figure 17). As a result of work in animal models it has been claimed that cortical remodeling is the result of loosening of the prosthesis particularly where the femoral component is fixed distally<sup>74</sup> although stem alignment may also influence abnormal focal cortical loading. Cortical remodeling may progress to form a pedestal that bridges the endosteum distal to the tip of the prosthesis.<sup>75</sup> Focal cortical thickening acts as a stress riser; a



**Figure 16** AP radiograph of a cemented THR in which the polyethylene cup has fractured allowing the head of the femoral component to articulate with the metal backing. Large amounts of metallic debris were shed in to the joint to lie deep to the joint capsule and produce characteristic cloud-like opacities (arrowhead).



**Figure 17**

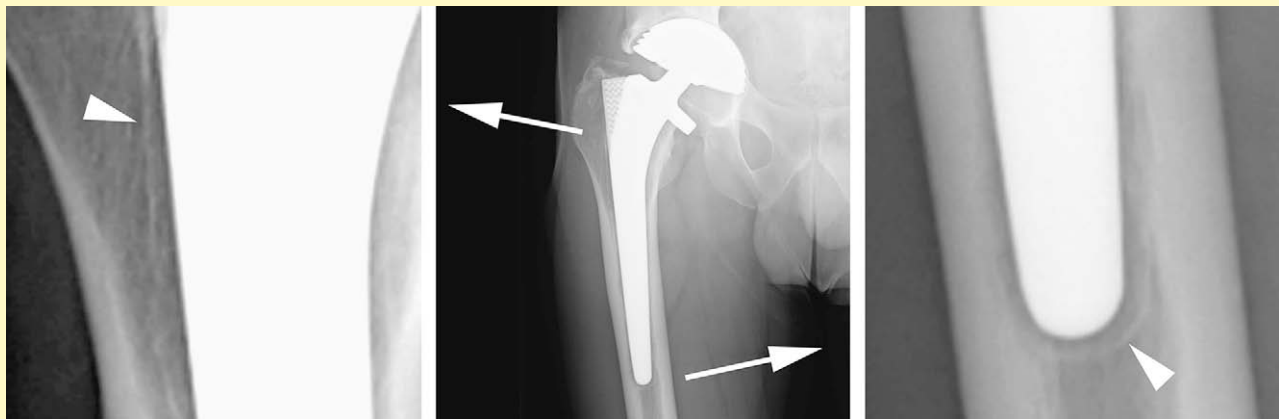
short transition between areas of cortex with different tensile strengths and therefore provides a point of weakness through which fractures propagate.

A neocortex is a feature of a loose uncemented stem. It is characterized by a thin sclerotic line that surrounds the femoral prosthesis from which it is separated by a lucency that is typically several millimeters thick<sup>76,77</sup> (Figure 18).

Delayed periprosthetic fractures in primary THR are uncommon but after revision THR the incidence rises to between 3.6% and 20.9%.<sup>78,79,80</sup> They occur most commonly at the tip of the femoral component<sup>81</sup> and appear to be associated with a patient age of over 70<sup>81</sup> with a loose, usually uncemented, femoral prosthesis.<sup>78</sup> Delayed fractures through the greater trochanter are rare but again are more common after revision surgery.<sup>82</sup> They can occur through areas of focal lysis caused by small particle disease.<sup>83</sup> They can also occur without obvious lysis when heat from a particularly thick cement mantle lateral to the shoulder of the prosthesis has been postulated as a cause for osteonecrosis of the greater trochanter (Figure 19). However the majority of patients with a periprosthetic fracture have a loose femoral stem and there are usually other radiographic features of loosening.<sup>79</sup>

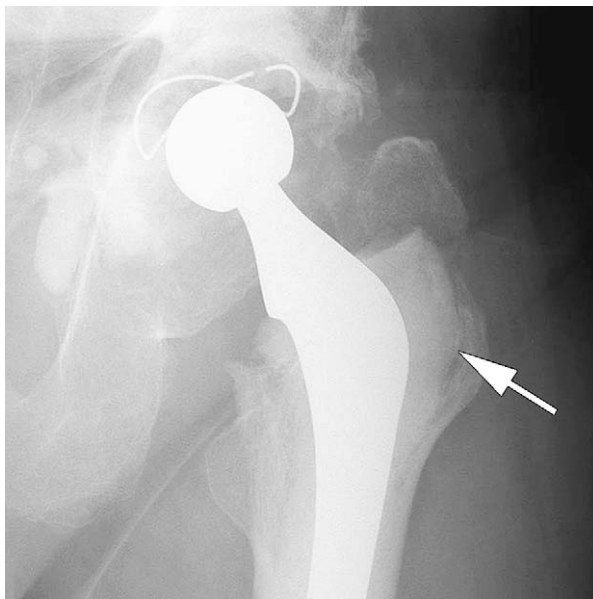
### Prosthetic fracture

The femoral prosthesis can fail because the stem suffers a fatigue fracture. This typically occurs in prostheses that are well fixed



Neocortex formation in a patient with a loose distally fixed uncemented THR. A thin line of sclerotic bone (arrowheads, in expanded images to left and right) surrounds the femoral prosthesis. Between this line and the prosthesis is a lucent area filled with a fibrous membrane.

**Figure 18**

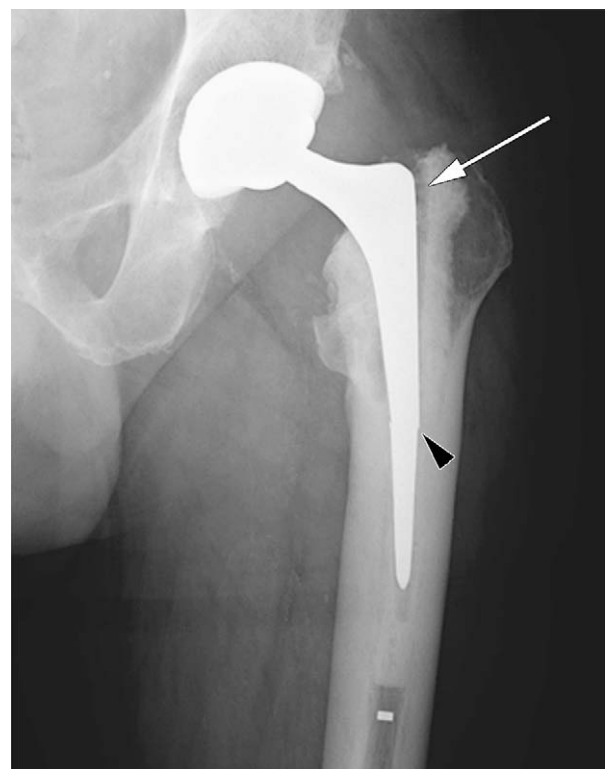


**Figure 19** AP radiograph of a cemented Charnley EliteTHR demonstrating a delayed fracture of the greater trochanter. Note the shell like cortex of the trochanter and the thick lateral cement mantle (arrow).

distally but are mobile proximally and result in fractures through the middle or proximal third of the stem<sup>84</sup> (Figure 20). Fracture of the acetabular cup usually occurs after severe focal wear of the polyethylene liner and can be demonstrated by discontinuity of the circumferential wire marker and medial migration of the head through the defect in the liner or cup (Figure 11).

## Conclusion

With a projected increasingly elderly population there will be a predictable continued increase in the number of hip arthroplasties



**Figure 20** Fractured femoral prosthesis (arrowhead) in a distally fixed stem with proximal loosening (arrow).

worldwide. As patients survive longer the proportion presenting with symptoms of a failing THR may also increase. Conventional radiographs are likely to remain the imaging investigation of choice for the near future. Many of the strengths and limitations of conventional radiography, for diagnosing the failing THR, have now been identified. These have been discussed in this review. ◆



## REFERENCES

- 1 National Joint Registry StatsOnline. Summary of annual statistics (England and Wales). [www.njrcentre.org.uk/](http://www.njrcentre.org.uk/) Last accessed 15th January 2009.
- 2 Harris WH, McCarthy JC, O'Neill DA. Femoral component loosening using contemporary techniques of femoral cement fixation. *J Bone Joint Surg Am* 1982; **64**: 1063–1067.
- 3 Temmerman OP, Raijmakers PG, David EF, et al. A comparison of radiographic and scintigraphic techniques to assess aseptic loosening of the acetabular component in a total hip replacement. *J Bone Joint Surg Am* 2004; **86-A**: 2456–2463.
- 4 Madey SM, Callaghan JJ, Olejniczak JP, Goetz DD, Johnston RC. Charnley total hip arthroplasty with use of improved techniques of cementing. The results after a minimum of fifteen years of follow-up. *J Bone Joint Surg Am* 1997; **79**: 53–64.
- 5 Massin P, Geais L, Astoin E, Simondi M, Lavaste F. The anatomic basis for the concept of lateralized femoral stems: a frontal plane radiographic study of the proximal femur. *J Arthroplasty* 2000; **15**: 93–101.
- 6 Cain TM, Fon GT, Brumby S, Howie DW. Plain film and arthrographic findings in painful total hip arthroplasties with surgical correlation. *Australas Radiol* 1990; **34**: 211–218.
- 7 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* 2007; **36**: 107–110.
- 8 Niskanen RO. Early repetitive radiography is unnecessary after an uncomplicated cemented hip or knee arthroplasty for osteoarthritis. *Acta Orthop Belg* 2005; **71**: 692–695.
- 9 Qureshi AA, Viridi AS, Didonna ML, et al. Implant design affects markers of bone resorption and formation in total hip replacement. *J Bone Miner Res* 2002; **17**: 800–807.
- 10 Schmidt R, Muller L, Kress A, Hirschfelder H, Aplas A, Pitto RP. A computed tomography assessment of femoral and acetabular bone changes after total hip arthroplasty. *Int Orthop* 2002; **26**: 299–302.
- 11 Tigges S, Stiles RG, Roberson JR. Complications of hip arthroplasty causing periprosthetic radiolucency on plain radiographs. *AJR Am J Roentgenol* 1994; **162**: 1387–1391.
- 12 Engh CAJ, Young AM, Engh CAS, Hopper RHJ. Clinical consequences of stress shielding after porous-coated total hip arthroplasty. *Clin Orthop Relat Res* 2003; **417**: 157–163.
- 13 Karachalios T, Tsatsaronis C, Efrimis G, Papadelis P, Lyritis G, Diakoumopoulos G. The long-term clinical relevance of calcar atrophy caused by stress shielding in total hip arthroplasty: a 10-year, prospective, randomized study. *J Arthroplasty* 2004; **19**: 469–475.
- 14 Engh CAJ, McAuley JP, Sychterz CJ, Sacco ME, Engh CAS. The accuracy and reproducibility of radiographic assessment of stress-shielding. A postmortem analysis. *J Bone Joint Surg Am* 2000; **82-A**: 1414–1420.
- 15 Schmidt R, Pitto RP, Kress A, et al. Inter- and intraobserver assessment of periacetabular osteodensitometry after cemented and uncemented total hip arthroplasty using computed tomography. *Arch Orthop Trauma Surg* 2005; **125**: 291–297.
- 16 Hing CB, Young DA, Dalziel RE, Bailey M, Back DL, Shimmin AJ. Narrowing of the neck in resurfacing arthroplasty of the hip: a radiological study. *J Bone Joint Surg Br* 2007; **89**: 1019–1024.
- 17 Bauer TW, Schils J. The pathology of total joint arthroplasty. II. Mechanisms of implant failure. *Skeletal Radiol* 1999; **28**: 483–497.
- 18 Aspenberg P, van der Vis H. Fluid pressure may cause periprosthetic osteolysis. Particles are not the only thing. *Acta Orthop Scand* 1998; **69**: 1–4.
- 19 Claus AM, Sychterz CJ, Hopper RHJ, Engh CA. Pattern of osteolysis around two different cementless metal-backed cups: retrospective, radiographic analysis at minimum 10-year follow-up. *J Arthroplasty* 2001; **16**: 177–182.
- 20 Hodgkinson JP, Shelley P, Wroblewski BM. The correlation between the roentgenographic appearance and operative findings at the bone-cement junction of the socket in Charnley low friction arthroplasties. *Clin Orthop Relat Res* 1988; **228**: 105–109.
- 21 Iwaki H, Scott G, Freeman MA. The natural history and significance of radiolucent lines at a cemented femoral interface. *J Bone Joint Surg Br* 2002; **84**: 550–555.
- 22 Bowerman JW, Sena JM, Chang R. The teardrop shadow of the pelvis; anatomy and clinical significance. *Radiology* 1982; **143**: 659–662.
- 23 Masri BA, Masterson EL, Duncan CP. The classification and radiographic evaluation of bone loss in revision hip arthroplasty. *Orthop Clin North Am* 1998; **29**: 219–227.
- 24 Walde TA, Mohan V, Leung S, Engh CAS. Sensitivity and specificity of plain radiographs for detection of medial-wall perforation secondary to osteolysis. *J Arthroplasty* 2005; **20**: 20–24.
- 25 Stauffer RN. Ten-year follow-up study of total hip replacement. *J Bone Joint Surg Am* 1982; **64**: 983–990.
- 26 Harris WH. Wear and periprosthetic osteolysis: the problem. *Clin Orthop Relat Res* 2001; 66–70.
- 27 Chiang PP, Burke DW, Freiberg AA, Rubash HE. Osteolysis of the pelvis: evaluation and treatment. *Clin Orthop Relat Res* 2003; 164–174.
- 28 Claus AM, Engh CAJ, Sychterz CJ, Xenos JS, Orishimo KF, Engh CAS. Radiographic definition of pelvic osteolysis following total hip arthroplasty. *J Bone Joint Surg Am* 2003; **85-A**: 1519–1526.
- 29 Engh CAJ, Sychterz CJ, Young AM, Pollock DC, SD T, Engh CAS. Interobserver. *J Arthroplasty* 2002; **17**: 752–759.
- 30 Berry DJ, Lewallen DG, Hanssen AD, Cabanela ME. Pelvic discontinuity in revision total hip arthroplasty. *J Bone Joint Surg Am* 1999; **81**: 1692–1702.
- 31 Cuckler JM, Star AM, Alavi A, Noto RB. Diagnosis and management of the infected total joint arthroplasty. *Orthop Clin North Am* 1991; **22**: 523–530.
- 32 Schmalzried TP. The infected hip: telltale signs and treatment options. *J Arthroplasty* 2006; **21**: 97–100.
- 33 Seelen JL, Bruijn JD, Hansen BE, Kingma LM, Bloem JL. Reproducible radiographs of acetabular prostheses. A method assessed in 35 patients. *Acta Orthop Scand* 1994; **65**: 258–262.
- 34 O'Neill DA, Harris WH. Failed total hip replacement. *J Bone Joint Surg Am* 1984; **66**: 540–546.
- 35 Mjoberg B, Selvik G, Hansson LI, Rosenqvist R, Onnerfalt R. Mechanical loosening of total hip prostheses. A radiographic and roentgen stereophotogrammetric study. *J Bone Joint Surg Br* 1986; **68**: 770–774.
- 36 Morscher EW. Current status of acetabular fixation in primary total hip arthroplasty. *Clin Orthop Relat Res* 1992; 172–193.
- 37 Gruen TA, McNeice GM, Amstutz HC. Modes of failure" of cemented stem-type femoral components: a radiographic analysis of loosening. *Clin Orthop Relat Res* 1979; 17–27.
- 38 Mjoberg B, Brismar J, Hansson LI, Pettersson H, Selvik G, Onnerfalt R. Definition of endoprosthetic loosening. Comparison of

- arthrography, scintigraphy and roentgen stereophotogrammetry in prosthetic hips. *Acta Orthop Scand* 1985; **56**: 469–473.
- 39 Alfaro-Adrian J, Gill HS, Murray DW. Cement migration after THR. A comparison of Charnley elite and Exeter femoral stems using RSA. *J Bone Joint Surg Br* 1999; **81**: 130–134.
- 40 Kiss J, Murray DW, Turner-Smith AR, Bithell J, Bulstrode CJ. Migration of cemented femoral components after THR. Roentgen stereophotogrammetric analysis. *J Bone Joint Surg Br* 1996; **78**: 796–801.
- 41 Mulroy RDJ, Harris WH. The effect of improved cementing techniques on component loosening in total hip replacement. An 11-year radiographic review. *J Bone Joint Surg Br* 1990; **72**: 757–760.
- 42 DeLee JG, Charnley J. Radiological demarcation of cemented sockets in total hip replacement. *Clin Orthop Relat Res* 1976; 20–32.
- 43 Eftekhari NS, Nercissian O. Incidence and mechanism of failure of cemented acetabular component in total hip arthroplasty. *Orthop Clin North Am* 1988; **19**: 557–566.
- 44 Murray DW. The definition and measurement of acetabular orientation. *J Bone Joint Surg Br* 1993; **75**: 228–32.
- 45 Herrlin K, Pettersson H, Selvik G. Comparison of two- and three-dimensional methods for assessment of orientation of the total hip prosthesis. *Acta Radiol* 1988; **29**: 357–361.
- 46 Hirakawa K, Mitsugi N, Koshino T, Saito T, Hirasawa Y, Kubo T. Effect of acetabular cup position and orientation in cemented total hip arthroplasty. *Clin Orthop Relat Res* 2001; 135–42.
- 47 Massin P, Schmidt L, Engh CA. Evaluation of cementless acetabular component migration. An experimental study. *J Arthroplasty* 1989; **4**: 245–251.
- 48 Rittmeister M, Callisjs C. Factors influencing cup orientation in 500 consecutive total hip replacements. *Clin Orthop Relat Res* 2006; **445**: 192–196.
- 49 Thoren B, Sahlstedt B. Influence of pelvic position on radiographic measurements of the prosthetic acetabular component. An experimental study on a pelvic model. *Acta Radiol* 1990; **31**: 133–136.
- 50 Anda S, Svenningsen S, Grontvedt T, Benum P. Pelvic inclination and spatial orientation of the acetabulum. A radiographic, computed tomographic and clinical investigation. *Acta Radiol* 1990; **31**: 389–394.
- 51 Kalteis T, Handel M, Herold T, Perlick L, Paetzel C, Grifka J. Position of the acetabular cup – accuracy of radiographic calculation compared to CT-based measurement. *Eur J Radiol* 2006; **58**: 294–300.
- 52 Olivecrona H, Weidenhielm L, Olivecrona L, et al. A new CT method for measuring cup orientation after total hip arthroplasty: a study of 10 patients. *Acta Orthop Scand* 2004; **75**: 252–260.
- 53 Kramhoft M, Gehrchen PM, Bodtker S, Wagner A, Jensen F. Inter- and intraobserver study of radiographic assessment of cemented total hip arthroplasties. *J Arthroplasty* 1996; **11**: 272–276.
- 54 Del Schutte HJ, Lipman AJ, Bannar SM, Livermore JT, Ilstrup D, Morrey BF. Effects of acetabular abduction on cup wear rates in total hip arthroplasty. *J Arthroplasty* 1998; **13**: 621–626.
- 55 Onsten I, Bengner U, Besjakov J. Socket migration after Charnley arthroplasty in rheumatoid arthritis and osteoarthritis. A roentgen stereophotogrammetric study. *J Bone Joint Surg Br* 1993; **75**: 677–680.
- 56 McBride MT, Muldoon MP, Santore RF, Trousdale RT, Wenger DR. Protrusion acetabuli: diagnosis and treatment. *J Am Acad Orthop Surg* 2001; **9**: 79–88.
- 57 Morley DCJ, Schmidt RH. Protrusion acetabuli prosthetica. *Orthop Rev* 1986; **15**: 135–141.
- 58 Gates HSr, Poletti SC, Callaghan JJ, McCollum DE. Radiographic measurements in protrusion acetabuli. *J Arthroplasty* 1989; **4**: 347–351.
- 59 Effenbger H, Heiland A, Ramsauer T, Plitz W, Dorn U. A model for assessing the rotational stability of uncemented femoral implants. *Arch Orthop Trauma Surg* 2001; **121**: 60–4.
- 60 Barrack RL, Mulroy RDJ, Harris WH. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. *J Bone Joint Surg Br* 1992; **74**: 385–389.
- 61 Chambers IR, Fender D, McCaskie AW, Reeves BC, Gregg PJ. Radiological features predictive of aseptic loosening in cemented Charnley femoral stems. *J Bone Joint Surg Br* 2001; **83**: 838–842.
- 62 Yates PJ, Burston BJ, Whitley E, Bannister GC. Collarless polished tapered stem: clinical and radiological results at a minimum of ten years' follow-up. *J Bone Joint Surg Br* 2008; **90**: 16–22.
- 63 Brooker AF, Bowerman JW, Robinson RA, Riley LHJ. Ectopic ossification following total hip replacement. Incidence and a method of classification. *J Bone Joint Surg Am* 1973; **55**: 1629–1632.
- 64 Nayak KN, Mulliken B, Rorabeck CH, Bourne RB, Woolfrey MR. Prevalence of heterotopic ossification in cemented versus noncemented total hip joint replacement in patients with osteoarthritis: a randomized clinical trial. *Can J Surg* 1997; **40**: 368–374.
- 65 Harwin SF. Trochanteric heterotopic ossification after total hip arthroplasty performed using a direct lateral approach. *J Arthroplasty* 2005; **20**: 467–472.
- 66 Purtill JJ, Eng K, Rothman RH, Hozack WJ. Heterotopic ossification. Incidence in cemented versus cementless total hip arthroplasty. *J Arthroplasty* 1996; **11**: 58–63.
- 67 Rama KR, Vendittoli PA, Ganapathi M, Borgmann R, Roy A, Lavigne M. Incidence and Severity of Heterotopic Ossification After Surface Replacement Arthroplasty and Total Hip Arthroplasty A Randomized Study. *J Arthroplasty* 2008.
- 68 Pohl F, Seufert J, Tauscher A, et al. The influence of heterotopic ossification on functional status of hip joint following total hip arthroplasty. *Strahlenther Onkol* 2005; **181**: 529–533.
- 69 Della VAG, Ruzo PS, Pavone V, Tolo E, Mintz DN, Salvati EA. Heterotopic ossification after total hip arthroplasty: a critical analysis of the Brooker classification and proposal of a simplified rating system. *J Arthroplasty* 2002; **17**: 870–875.
- 70 Wright JG, Moran E, Bogoch E. Reliability and validity of the grading of heterotopic ossification. *J Arthroplasty* 1994; **9**: 549–553.
- 71 Toom A, Fischer K, Martson A, Rips L, Haviko T. Inter-observer reliability in the assessment of heterotopic ossification: proposal of a combined classification. *Int Orthop* 2005; **29**: 156–159.
- 72 Khan RJ, Wimhurst J, Foroughi S, Toms A. The Natural History of Metallosis From Catastrophic Failure of a Polyethylene Liner in a Total Hip. *J Arthroplasty* 2008.
- 73 Adolphson P. Femoral cortical remodeling after uncemented total hip arthroplasty. A prospective radiologic study of 26 hips followed for 2 to 4 years. *J Arthroplasty* 1997; **12**: 99–105.
- 74 Bergh MS, Muir P, Markel MD, Manley PA. Femoral bone adaptation to unstable long-term cemented total hip arthroplasty in dogs. *Vet Surg* 2004; **33**: 238–245.
- 75 Manaster BJ. From the RSNA refresher courses. Total hip arthroplasty: radiographic evaluation. *Radiographics* 1996; **16**: 645–660.

- 76** Harris WH. Will stress shielding limit the longevity of cemented femoral components of total hip replacement? *Clin Orthop Relat Res* 1992; 120–123.
- 77** Nelissen RG, Bauer TW, Weidenhielm LR, LeGolvan DP, Mikhail WE. Revision hip arthroplasty with the use of cement and impaction grafting. Histological analysis of four cases. *J Bone Joint Surg Am* 1995; **77**: 412–422.
- 78** Lindahl H. Epidemiology of periprosthetic femur fracture around a total hip arthroplasty. *Injury* 2007; **38**: 651–654.
- 79** Lindahl H, Malchau H, Herberts P, Garellick G. Periprosthetic femoral fractures classification and demographics of 1049 periprosthetic femoral fractures from the Swedish National Hip Arthroplasty Register. *J Arthroplasty* 2005; **20**: 857–865.
- 80** Farfalli GL, Buttaro MA, Piccaluga F. Femoral fractures in revision hip surgeries with impacted bone allograft. *Clin Orthop Relat Res* 2007; **462**: 130–136.
- 81** Cook RE, Jenkins PJ, Walmsley PJ, Patton JT, Robinson CM. Risk factors for periprosthetic fractures of the hip: a survivorship analysis. *Clin Orthop Relat Res* 2008; **466**: 1652–1656.
- 82** Ornstein E, Atroshi I, Franzen H, Johnsson R, Sandquist P, Sundberg M. Early complications after one hundred and forty-four consecutive hip revisions with impacted morselized allograft bone and cement. *J Bone Joint Surg Am* 2002; **84-A**: 1323–1328.
- 83** Hsieh PH, Chang YH, Lee PC, Shih CH. Periprosthetic fractures of the greater trochanter through osteolytic cysts with uncemented MicroStructured Omnifit prosthesis: retrospective analyses of 23 fractures in 887 hips after 5–14 years. *Acta Orthop* 2005; **76**: 538–543.
- 84** Wilson LF, Nolan JF, Heywood-Waddington MB. Fracture of the femoral stem of the Ring TCH hip prosthesis. *J Bone Joint Surg Br* 1992; **74**: 725–728.

### Practice points

- Consider iliac-oblique Judet view to supplement or replace standard lateral view
- Standard radiography limited for early mechanical loosening
- Movement on interval radiographs is the only absolute feature of a loose prosthesis
- Other features have lesser sensitivities and specificities and may not be related to symptoms

PICTORIAL REVIEW

# CT and MRI of hip arthroplasty

J.G. Cahir<sup>a,\*</sup>, A.P. Toms<sup>a</sup>, T.J. Marshall<sup>a</sup>, J. Wimbhurst<sup>b</sup>, J. Nolan<sup>b</sup>

Departments of <sup>a</sup>Radiology, and <sup>b</sup>Orthopaedics, Norfolk and Norwich University Hospital, Colney Lane, Norwich, UK

Received 31 July 2006; received in revised form 29 March 2007; accepted 3 April 2007

---

Plain films are the initial imaging method of choice for evaluation of hip arthroplasty. Recent advances in technology and imaging techniques have largely overcome the problems of beam hardening in computed tomography (CT) and magnetic susceptibility artefact in magnetic resonance imaging (MRI). CT and MRI have now become useful imaging techniques in the assessment of hip arthroplasty.

© 2007 The Royal College of Radiologists. Published by Elsevier Ltd. All rights reserved.

---

## Introduction

Hip arthroplasty is a common treatment for patients with osteoarthritis and approximately 1.5 million procedures are performed worldwide each year.<sup>1</sup> Although complication rates are low, the large number of hip replacements means that the complications related to hip arthroplasty are common in clinical practice.

Problems that occur after arthroplasty include osteolysis, granulomatous disease, heterotopic new bone formation, dislocation, superficial and deep infection, mechanical aseptic loosening, prosthetic and periprosthetic fracture, and local nerve damage. These problems are a source of morbidity and may require surgical revision.<sup>2,3</sup> Plain films are the initial imaging method of choice for evaluation of hip arthroplasty but are limited in evaluation of complications due to their inability to delineate complex three-dimensional (3D) structures.<sup>4</sup> However, CT and MRI are now also useful imaging tools for assessing orthopaedic implants, as recent advances in technology and imaging techniques have largely overcome the problems of beam hardening in CT and magnetic

susceptibility artefact in MRI.<sup>3–12</sup> CT and MRI can detect periprosthetic collections, evaluate osteolysis due to small-particle disease, clearly define the periprosthetic soft tissues, and demonstrate loosening.

This article will describe the optimization of CT and MRI protocols for imaging hip arthroplasties. The role of CT and MRI in contributing to the management of the patient with complications related to hip arthroplasty will be highlighted.

## Technical considerations

### CT

Beam-hardening artefacts, which manifest as alternating high and low attenuation lines radiating from the prosthesis, are the major cause of image degradation in CT images of metallic implants. The degree of artefact is proportional to the proton density of the metallic implant with cobalt–chrome–steel and stainless steel alloys causing the most severe artefacts.<sup>10,13</sup> These artefacts can be reduced by increasing the signal to noise ratio by increasing the output of the tube (mA and kVp). The exact values will depend on the CT machine and the size of the patient but the mA should be in the region of 350–400 mA on modern

---

\* Guarantor and correspondent: J. Cahir, Department of Radiology, Norfolk and Norwich University Hospital, Colney Lane, Norwich, Norfolk NR4 7UY, UK. Tel.: +44 1603 286104; fax: +44 1603 286077.

E-mail address: [john.cahir@nnuh.nhs.uk](mailto:john.cahir@nnuh.nhs.uk) (J.G. Cahir).



machines.<sup>12</sup> Using a fast iterative algorithm<sup>14</sup> and an extended CT scale can reduce these artefacts.<sup>15</sup> Soft-tissue or smooth reconstruction filters reduce metal artefact with an inevitable, but acceptable, reduction of spatial resolution.

Viewing the CT images using wide window widths reduces the observer's perception of the artefact. Data acquired with multidetector CT, which is reconstructed with a soft-tissue algorithm and overlapping sections, can be reformatted in any plane, which may further reduce artefact.<sup>12</sup>

## MRI

MRI has, until recently, been limited in the imaging of postoperative orthopaedic patients due to magnetic susceptibility artefacts produced by metallic implants. Magnetization of the implant affects the local field gradient, proton dephasing, and spin frequency resulting in signal void, spatial distortion, and spurious high signal. As MRI machines have implemented higher magnetic field strengths, which induce greater magnetization of orthopaedic implants, so the size of the susceptibility artefacts has increased.<sup>2</sup>

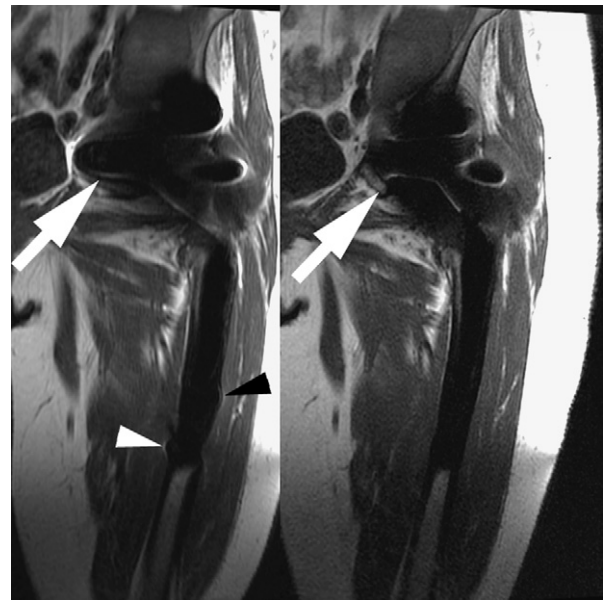
MRI parameters can be modified to minimize these artefacts. Increasing the frequency encoding gradient strength decreases the misregistration artefact proportionately. Fast spin-echo techniques refocus spins at a shorter interval than conventional spin-echo techniques and reduce diffusion-related signal intensity loss. Reducing the volume of the voxels (increasing the spatial resolution) reduces diffusion-related signal intensity loss. It also reduces the spatial definition of the signal void, and therefore, leads to reduction in apparent size of the void.<sup>2</sup> Spectral fat suppression is particularly susceptible to metallic artefact and should be avoided in favour of a short tau inversion recovery (STIR) sequence where some of the dephasing of proton spins, due to magnetic field inhomogeneity, is refocused by the 180° inversion pulse. The frequency encoding direction of the image is more susceptible to artefact, because of proton spin dephasing, than the phase encoding direction, and therefore, careful selection of the phase and frequency-encoding directions, in the three principal anatomical planes, will allow superior periprosthetic imaging in the frequency-encoding direction.<sup>2,10</sup> Combining all of these adjustments using metal artefact reduction sequences (MARS) can allow distinction of cortex, marrow, cement mantle, and disease in the region of the femoral stem of the implant (Fig. 1). Positioning the long axis of the prosthesis

parallel to the  $B_0$  magnetic field reduces susceptibility, which is why the stem of the femoral component responds better to these refinements than the obliquely oriented neck.

## CT versus MRI

MRI offers advantages over CT in assessing many aspects of hip arthroplasty because of its superior differentiation of soft tissues. However, it is still limited in the region of the acetabulum because susceptibility artefact has not yet been completely resolved by current techniques (Fig. 1). It is now possible to image periprosthetic fractures, osteolysis, marrow oedema, collections, extraosseous soft-tissue deposits, and adjacent musculature with routine MRI capabilities.<sup>2,10,16,17</sup>

CT has the advantage of speed. An axial volume acquisition through the pelvis and both femora takes a matter of minutes; multiplanar and surface-shaded reformats can be constructed after the patient has left the department. CT is preferable for imaging the roof of the acetabulum, and is probably superior to MRI for imaging cement, heterotopic ossification, and metallosis.



**Figure 1** Coronal T1-weighted MRI of a left metal-on-metal total hip arthroplasty. Part (a) was acquired using conventional fast spin-echo parameters and (b) was acquired using metal artefact reduction parameters described in the text. Mismatching of the signal from the femoral cortex (arrowheads) is nearly completely resolved by the MARS where the stem of the prosthesis lies parallel to  $B_0$ . The mismatching of the signal from the neck and acetabulum of the prosthesis is reduced so that the superior pubic ramus (arrow) is discernible.

## Complications of hip arthroplasty

### Osteolysis

Small-particle disease is a biological reaction to wear debris from hip arthroplasties, in particular those with polyethylene components.<sup>18</sup> Polyethylene debris is phagocytosed by macrophages, which then accumulate to form foreign-body granulomas. These masses of granulation tissue stimulate osteoclastic activity, and are a major cause of periprosthetic osteolysis.<sup>5,18–20</sup> However, repeated elevation of hydrostatic pressure within the effective postoperative joint space, which includes the bone cement interface of both components of the prosthesis, has also been implicated as a cause of osteolysis without the presence of granulomata.<sup>21</sup> This may have implications for the appearance of osteolysis on CT and MRI.

Conventional radiographs are routinely used to detect osteolysis around an implant but are limited by their planar representation of complex 3D pelvic anatomy.<sup>4,22</sup> CT has been reported to detect 87% of osteolytic lesions, compared with the 52% of lesions detected with a standard radiographic series, that are subsequently found at surgery.<sup>4</sup>

On CT, osteolysis manifests as well-defined lucencies devoid of osseous trabeculae.<sup>9</sup> The lucencies are typically in continuity with the prosthesis. In our experience granulomas are homogeneous soft-tissue masses with typical attenuation values of 30 HU (range –40 to 100 HU). Following the intravenous administration of iodinated contrast medium the attenuation is typically enhanced to about 90 HU (range 0 to 180 HU; the wide range of attenuation values reflects the noise from the adjacent prosthesis) and may predominate peripherally.

The typical signal characteristics of osteolysis on MRI are areas of low T1 signal and intermediate to slightly increased T2 signal (similar to skeletal muscle) with a well-defined additional line of low signal surrounding areas of marrow replacement. Both extra and intra-osseous granulomas have similar signal characteristics (Fig. 2). Peripheral enhancement and some internal irregular enhancement may be demonstrated with intravenous gadolinium.<sup>2,8</sup>

The primary indication for cross-sectional imaging in osteolysis is to provide a map of bone loss to aid planning of revision surgery. It is performed less commonly as an attempt to differentiate bland osteolysis from infection or tumour. The key areas of osteolysis in the pelvis that determine revision surgery are the medial wall and the roof of the acetabulum, and the anterior

and posterior columns. Critical loss of bone mass in the acetabular roof may require bone grafting and a mesh support for a new cemented cup or a large cementless cup with or without superior augmentation. Medial wall deficiencies may require a medial wall mesh or acetabular cage together with bone allograft, and are not reliably seen on plain films (Fig. 3). Defects that compromise either of the columns require internal fixation with reconstruction plates.

Pelvic discontinuity is a specific form of bone loss in which the ilium is separated from the pubis and ischium by osteolysis or a fracture through the acetabulum. It is more common in women and patients with rheumatoid arthritis, and pelvic radiotherapy is also thought to be a risk factor.<sup>23</sup> Pelvic discontinuity may require reconstruction with a plate and acetabular cage before insertion of allograft and a cemented cup or a custom-built cementless acetabular component.

Estimates of bone loss from conventional radiographs are typically low when compared with operative measurements.<sup>5</sup>

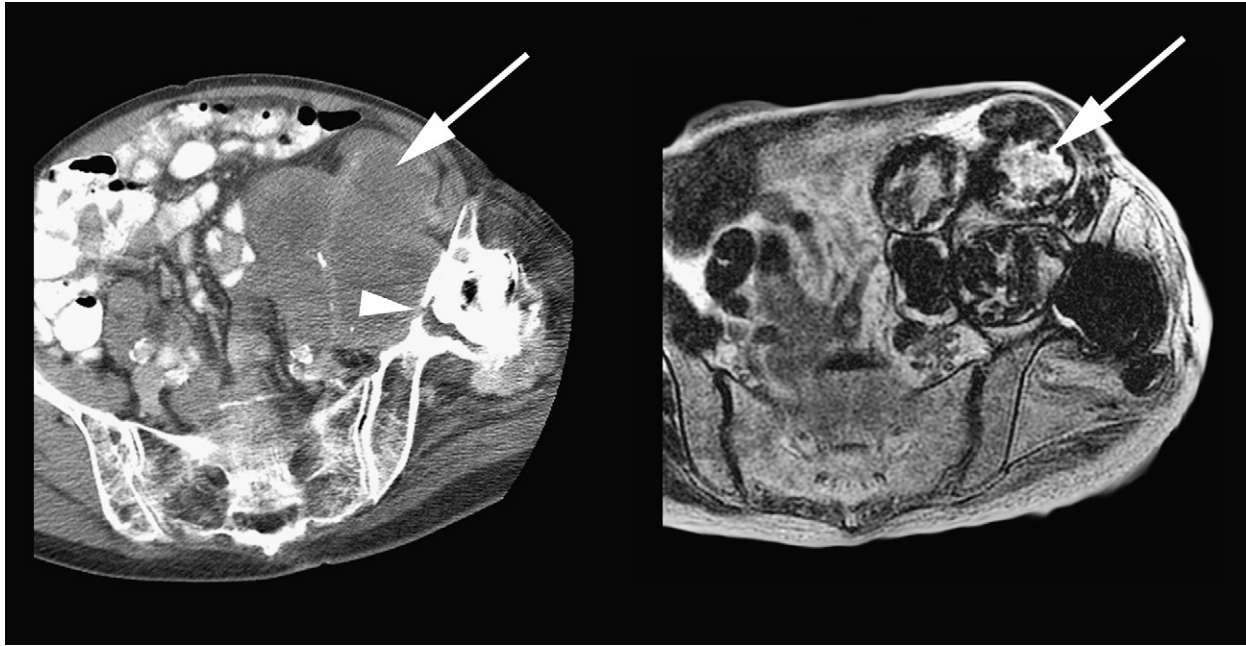
CT has been demonstrated to be more accurate than plain radiography in detecting the presence of osteolysis.<sup>1,4,9</sup> CT can also estimate the volume of osteolysis,<sup>1,4,24</sup> and thus, may help to estimate the quantity and type of bone graft that may be required at revision.

### Infection

Infection may be characterized on CT by an aggressive osteolysis with an ill-defined margin. By this stage there is frank osteomyelitis. More typically infected prostheses present with more chronic indolent changes that can be difficult to distinguish from other forms of osteolysis. Periostritis, indicated by periosteal new bone, on CT appears to be a very sensitive (100%), but not very specific (16%) marker of infection (Fig. 4). Other soft-tissue markers of infection such as joint distension, fluid-filled bursae, and soft-tissue fluid collections have been reported to increase the specificity of CT for infection to 87%. The absence of joint distension has a negative predictive value for infection of 96%.

Ultrasound has an important role in the evaluation of suspected infection as it can demonstrate a joint effusion or periprosthetic collection. Ultrasound is an invaluable technique to guide aspiration of an effusion or collection. Fluoroscopy also has a role in guiding hip joint aspiration.

The distinction between loosening and infection is important because loosening of a prosthesis is



**Figure 2** (a) Axial contrast medium-enhanced CT image through the left hip and pelvis demonstrates a peripherally enhancing lobulated mass of soft-tissue attenuation (arrow) abutting a defect in the medial wall of the left acetabulum (arrowhead) where there is osteolysis at the cement–bone interface of the acetabular cement mantle. (b) Axial T2-weighted MRI image at a similar level through the left hip demonstrates a lobulated mass that has irregular low signal margins and high signal centres (arrows) both on T1-weighted and T2-weighted sequences suggesting loculated cavities containing proteinaceous fluid. Ultrasound-guided biopsies of the wall of the lesion revealed non-specific inflammatory features only without evidence of polyethylene particles. Microbiology was negative.



**Figure 3** (a) A Plain film showing osteolysis involving both the acetabular and femoral components. (b) Coronal CT showing a large defect in the medial wall of the right acetabulum, which is not appreciated on the plain radiograph. This information is important in the planning of a revision arthroplasty.

treated with revision arthroplasty whereas infection may require a two stage revision.<sup>3</sup> MRI signal characteristics of osteolysis are different from those of infection. Infection has a signal intensity close to that of fluid and is poorly defined compared with small-particle disease (Figs. 5 and 6). Mechanical loosening may cause fluid collections that parallel the femoral stem; however, the specificity of periprosthetic fluid in the setting of pain after hip arthroplasty, and the features that distinguish early infective loosening from mechanical loosening are not known.<sup>2</sup>

### Tumour

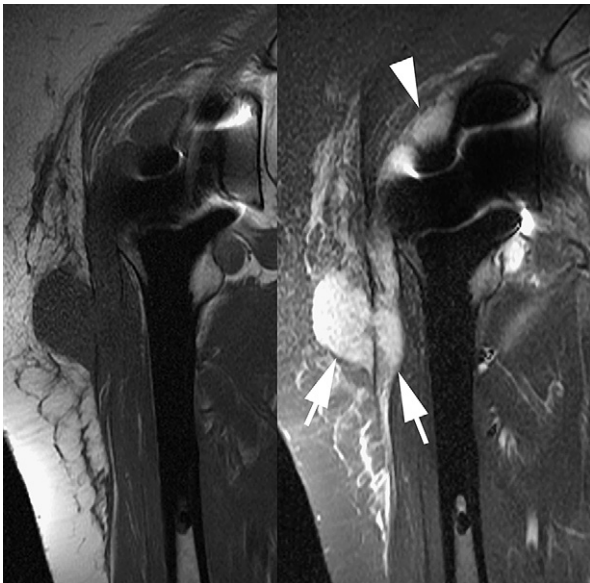
Bone destruction present on a conventional radiograph, due to aggressive osteolysis, may herald a bone metastasis or, rarely, a primary bone tumour. Differentiating a bone neoplasm in its early stages from granulomatous disease can be difficult. Small-particle disease is typically contiguous with the prosthesis but this can also occur with a tumour. Small-particle disease typically has a low signal interface with marrow and is not





**Figure 4** Coronal reformatted CT of the right hip demonstrating osteolysis at the bone–cement interface (arrowhead). The focal area of bone loss and adjacent periosteal reaction was due to infection with *Staphylococcus aureus* (arrow).

associated with bone marrow oedema,<sup>8</sup> whereas metastases are often aggressive with ill-defined margins and nearby reactive bone marrow and soft-tissue oedema. Disconcertingly large masses



**Figure 5** Coronal T1-weighted and T2-weighted images of a right THR with an collection arising from the joint (arrowhead) and tracking initially deep and then passing through a defect in the iliotibial tract (arrows).



**Figure 6** Sagittal T2-weighted MRI images of a revision THR with collections due to infection adjacent to the femur at multiple sites (arrows).

may result from small-particle disease (Fig. 2), and complications such as stress reactions, insufficiency, and pathological fractures may result in imaging findings that overlap with those of tumour. In the absence of a known primary tumour biopsy of the periprosthetic soft tissue may be required.

### Loosening

The most common cause of pain in patients after hip arthroplasty is mechanical loosening of the prosthesis. The criteria for diagnosing loosening of the prosthesis are the same as have been described on plain radiographs, namely migration of any component of the prosthesis with time, a fracture of the cement mantle and osteolysis surrounding 50–100% of either cement mantle.<sup>18</sup> Whether the cross-sectional capability, particularly of CT, offers any advantage over conventional radiography is unknown.

### Fractures

Periprosthetic fractures occur because of loosening, osteolysis, stress risers, osteoporosis, and unfavourable biomechanics. The acute periprosthetic fractures, such as at the tip of the femoral prosthesis and the greater trochanter are usually diagnosed with conventional radiographs, without the help of cross-sectional imaging. Some fractures though are occult causes of hip pain following THR and can be demonstrated using CT or MR.<sup>8</sup>

Pathological fractures through areas of osteolysis and insufficiency or stress fractures can be subtle and difficult to detect on conventional radiographs. The congruity of the medial wall of the acetabulum, the anterior and posterior columns, and the acetabular will all determine surgical technique and theatre time. Although MRI can now demonstrate fractures in close proximity to metal work (Fig. 7), CT is preferred because of the superior delineation of bone at the acetabulum (Fig. 8).

### Metal-on-metal arthroplasties

An unusual reaction, associated with particular second-generation metal-on-metal (cobalt–chromium alloy) hip prostheses, has recently been described.<sup>25</sup> Clinically a small proportion of patients present with pain, typically within the first 2–3 years after the operation. At surgery there is periprosthetic soft-tissue necrosis, sterile fluid collections, and bone necrosis. The histological picture is predominantly one of necrosis and occasionally a dense perivascular lymphocytic infiltrate, which may represent a unique disease process.<sup>26,27</sup> Conventional radiography is usually normal but MRI demonstrates striking changes that appear to correspond to macroscopic surgical findings (Fig. 9).

### Muscles and tendons

Greater trochanteric bursitis is a recognized complication of THR, particularly after a trans-gluteal



**Figure 7** A T2-weighted MARS MRI image demonstrating a fracture of the medial calcar with fluid/haemorrhage (arrow) tracking between the medial calcar and the adjacent intact cement mantle.

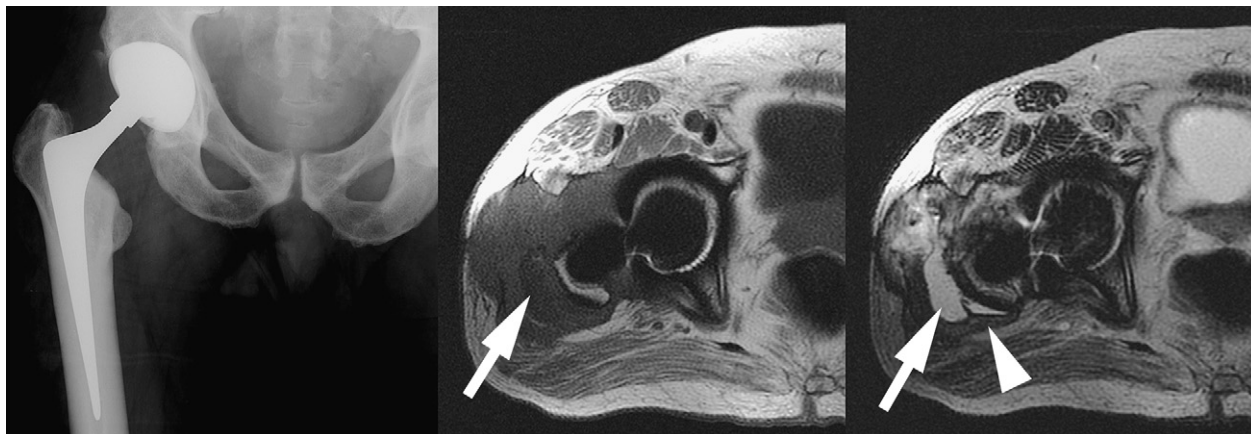


**Figure 8** Sagittal reformatted CT through the right hip demonstrating a stress fracture (arrow) of the posterior column that was not appreciated on conventional AP and coned lateral radiographs.

approach, and can be confirmed using MRI if there is any diagnostic uncertainty. It is demonstrated as an area of fluid-like signal, best visualized on fat-saturated T2 sequences, deep to the gluteus maximus tendon insertion into the greater trochanter, often extending posteriorly and medially.

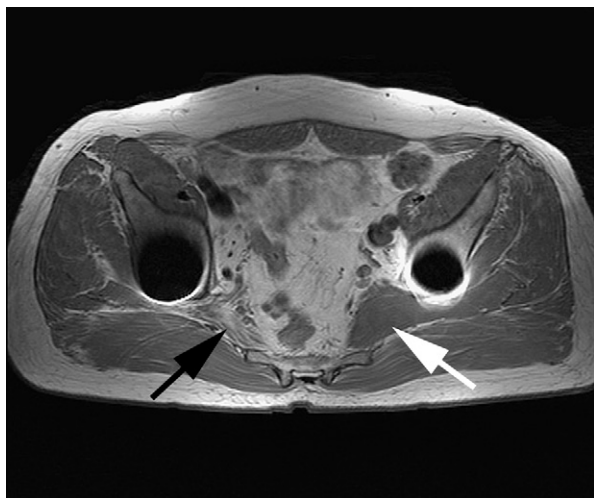
Spontaneous avulsion of the gluteal tendons is a recognized association of THR, which can present with an acute Trendelenberg gait. However, some of cases of gluteal avulsion are asymptomatic.<sup>16</sup> These may predate the surgery as a consequence of severe osteoarthritis but are phenomena that have yet to be fully explored.

Avulsion of the short external rotators is a recognized consequence of THR, usually after the posterior approach to the joint.<sup>16</sup> However, whether this is due to the prosthesis, or is associated with the surgical approach is not clear. The piriformis tendon is usually divided in the posterior approach to the hip and whether or not to repair is the subject some debate.<sup>28–30</sup> It may be that any repaired tendons commonly avulse early in the postoperative course and the consequence of this is unclear. The short external rotators are optimally imaged on axial T1-weighted or T2-weighted MRI. With a large matrix and thin sections the individual muscles can be identified (although the gemelli are difficult to separate). Small fields of view localized to the hip will often not cover all the muscles, particularly piriformis (Fig. 10).



**Figure 9** (a) Normal conventional radiograph of a patient with symptoms of recurrent hip pain 2 years after a right metal-on-metal hip arthroplasty. (b and c) Axial T1-weighted and T2-weighted MARS MRI images within 1 month of the radiograph demonstrates a large fluid collection (arrow) with a low signal T2-weighted rim (arrowhead) surrounding the neck of the prosthesis and the right greater trochanter. The gluteal tendons, which should be visible at this level, have been avulsed.

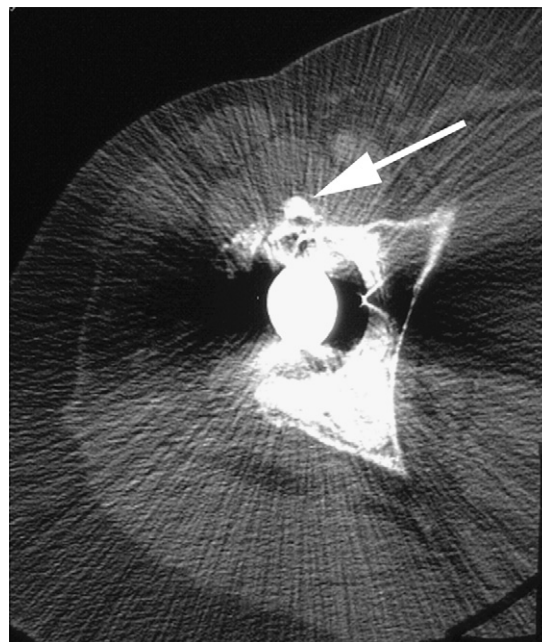
Psoas tendon impingement is an uncommon complication of THR that occurs because of friction against small anterior acetabular spurs or small extruded pieces of cement. These can be demonstrated on axial images and sagittal reformats. The spurs are usually small, and if there is diagnostic uncertainty, then a local anaesthetic injection targeted to the spur may confirm the diagnosis<sup>31</sup> (Fig. 11).



**Figure 10** Axial T1-weighted image through the pelvis in a patient with bilateral metal-on-metal hip arthroplasties demonstrating atrophy of the piriformis on the symptomatic diseased side (black arrow) whereas the piriformis muscle attached to the asymptomatic hip is preserved. Both hips were replaced using a posterolateral approach.

## Metallosis

Metallosis occurs when large volumes of metal debris enter the joint space. This occurs as



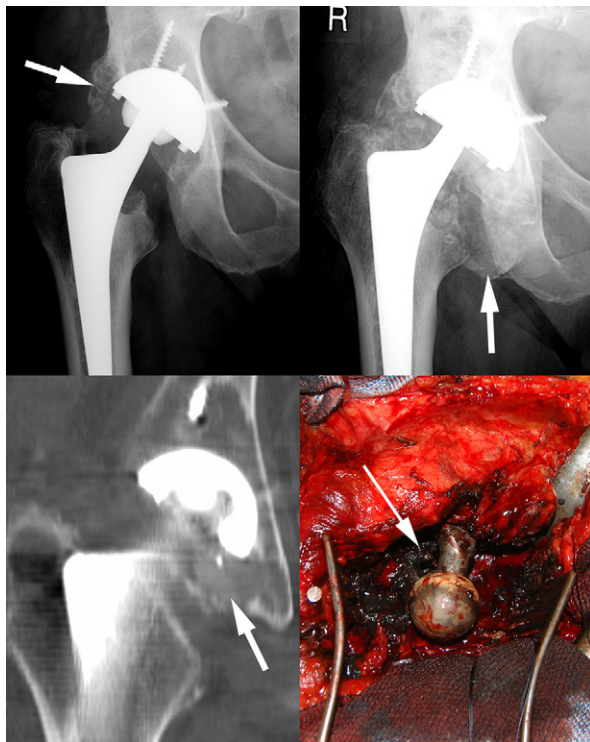
**Figure 11** Axial CT of the right hip in a patient with clinical features of psoas tendon irritation demonstrates a small piece of extruded cement (arrow) anterior to the acetabular component indenting the posterior psoas tendon. An ultrasound-guided local anaesthetic injection confirmed that this was the source of symptoms allowing successful treatment by surgical removal of the offending piece of cement. [Reproduced with permission and copyright of the British Editorial Society of Bone and Joint Surgery.<sup>31</sup>]



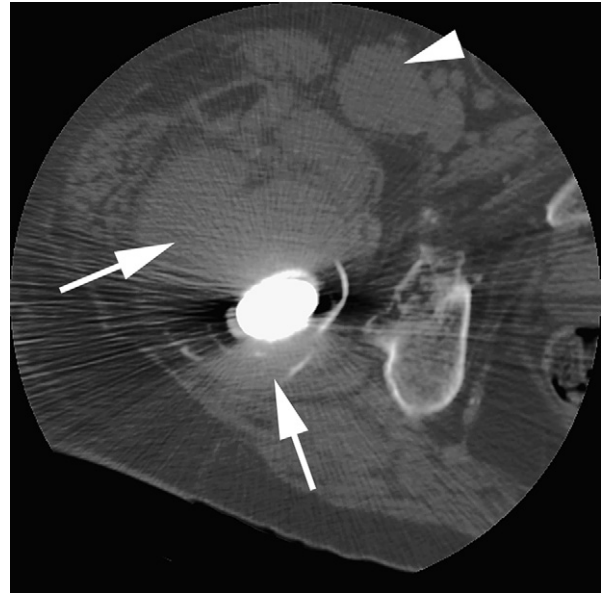
a consequence of full-thickness wear, or fracture, through the polyethylene liner of the acetabular component such that the femoral head articulates with the metal backing of the acetabular component. This presents as a homogeneous opacity, on plain radiography and CT, which conforms to the space defined by the pseudocapsule of the joint. It may be mistaken for heterotopic ossification but can be distinguished by its lack of cortico-medullary differentiation (Fig. 12).

### Lymphadenopathy

Dissemination of large amounts of metallic and polyethylene debris may result in regional, pelvic and para-aortic lymphadenopathy. However, this is



**Figure 12** (a) Conventional radiograph of a right total hip arthroplasty demonstrating heterotopic ossification adjacent to the supero-lateral acetabular margin (arrow). (b) Two years later there is more extensive opacification inferomedial to the neck of the prosthesis, which was considered to be heterotopic ossification (arrow). The femoral head has migrated into the acetabular cup indicating a full-thickness defect of the polyethylene liner. (c) Coronal reformatted CT of the hip prosthesis demonstrates this opacity to be homogeneous without the cortico-medullary differentiation that would be expected with ossification. (d) An intra-operative photograph demonstrates a black oily liquid (arrow), containing metal debris from the articulation of the head of the prosthesis with the titanium backing of the acetabulum, within the open joint.



**Figure 13** Axial CT through a right THR. CT was performed to delineate the extent of osteolysis before revision surgery. A large homogeneous mass of soft tissue surrounds the proximal femoral prosthesis and is associated with extensive osteolysis (arrows). Lymphadenopathy was noted in the right inguinal and external iliac territories. Biopsy of the periprosthetic mass showed non-specific inflammatory changes and no polyethylene debris. The patient died 3 weeks later and lymphoma was diagnosed at post-mortem examination.

a rarely reported phenomenon. Polyethylene particles have also been found in the liver and spleen in patients who have undergone hip revision surgery for mechanical failure.<sup>32</sup> Biopsy of affected lymph nodes reveals macrophages containing polyethylene particles. Lymphadenopathy associated with a THR is probably more frequently caused by infection or malignancy (Fig. 13).

### Conclusion

Improvements in metal artefact reduction protocols mean that diagnostic CT and MRI examinations can be produced on standard equipment available in most radiology departments.

CT probably remains the technique of choice for evaluating the volume and extent of osteolysis while planning revision surgery.

MRI promises to be an increasingly useful diagnostic technique, particularly for soft-tissue disease associated with THR.

### References

1. Looney RJ, Boyd A, Totterman S, et al. Volumetric computerized tomography as a measurement of periprosthetic

- acetabular osteolysis and its correlation with wear. *Arthritis Res* 2002;4:59–63.
2. White LM, Kim JK, Mehta M, et al. Complications of total hip arthroplasty: MR imaging-initial experience. *Radiology* 2000;215:254–62.
  3. Cyteval C, Hamm V, Sarabere MP, et al. Painful infection at the site of hip prosthesis: CT imaging. *Radiology* 2002;224:477–83.
  4. Leung S, Naudie D, Kitamura N, et al. Computed tomography in the assessment of periacetabular osteolysis. *J Bone Joint Surg Am* 2005;87:592–7.
  5. Chiang PP, Burke DW, Freiberg AA, et al. Osteolysis of the pelvis: evaluation and treatment. *Clin Orthop Relat Res* 2003;164–74.
  6. Czerny C, Krestan C, Imhof H, et al. Magnetic resonance imaging of the postoperative hip. *Top Magn Reson Imaging* 1999;10:214–20.
  7. Olsen RV, Munk PL, Lee MJ, et al. Metal artifact reduction sequence: early clinical applications. *RadioGraphics* 2000;20:699–712.
  8. Potter HG, Nestor BJ, Sofka CM, et al. Magnetic resonance imaging after total hip arthroplasty: evaluation of periprosthetic soft tissue. *J Bone Joint Surg Am* 2004;86-A:1947–54.
  9. Puri L, Wixson RL, Stern SH, et al. Use of helical computed tomography for the assessment of acetabular osteolysis after total hip arthroplasty. *J Bone Joint Surg Am* 2002;84-A(4):609–14.
  10. Sofka CM, Potter HG. MR imaging of joint arthroplasty. *Semin Musculoskelet Radiol* 2002;6:79–85.
  11. Stulberg SD, Wixson RL, Adams AD, et al. Monitoring pelvic osteolysis following total hip replacement surgery: an algorithm for surveillance. *J Bone Joint Surg Am* 2002;84-A(Suppl. 2):116–22.
  12. White LM, Buckwalter KA. Technical considerations: CT and MR imaging in the postoperative orthopedic patient. *Semin Musculoskelet Radiol* 2002;6:5–17.
  13. Wang JC, Yu WD, Sandhu HS, et al. A comparison of magnetic resonance and computed tomographic image quality after the implantation of tantalum and titanium spinal instrumentation. *Spine* 1998;23:1684–8.
  14. Wang G, Frei T, Vannier MW. Fast iterative algorithm for metal artifact reduction in x-ray CT. *Acad Radiol* 2000;7:607–14.
  15. Link TM, Berning W, Scherf S, et al. CT of metal implants: reduction of artifacts using an extended CT scale technique. *J Comput Assist Tomogr* 2000;24:165–72.
  16. Pfirrmann CW, Notzli HP, Dora C, et al. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology* 2005;235:969–76.
  17. Claus AM, Walde TA, Leung SB, et al. Management of patients with acetabular socket wear and pelvic osteolysis. *J Arthroplasty* 2003;18(Suppl. 1):112–7.
  18. Harris WH, McCarthy Jr JC, O'Neill DA. Femoral component loosening using contemporary techniques of femoral cement fixation. *J Bone Joint Surg Am* 1982;64:1063–7.
  19. Masri BA, Masterson EL, Duncan CP. The classification and radiographic evaluation of bone loss in revision hip arthroplasty. *Orthop Clin North Am* 1998;29:219–27.
  20. Wimhurst JA. The pathogenesis of aseptic loosening. *Curr Orthop* 2002;16:407–10.
  21. Aspenberg P. Wear and osteolysis in total joint replacements. *Acta Orthop Scand* 1998;69:435–6.
  22. Southwell DG, Bechtold JE, Lew WD, et al. Improving the detection of acetabular osteolysis using oblique radiographs. *J Bone Joint Surg Br* 1999;81:289–95.
  23. Berry DJ, Lewallen DG, Hanssen AD, et al. Pelvic discontinuity in revision total hip arthroplasty. *J Bone Joint Surg Am* 1999;81:1692–702.
  24. Stamenkov R, Howie D, Taylor J, et al. Measurement of bone defects adjacent to acetabular components of hip replacement. *Clin Orthop Relat Res* 2003;117–24.
  25. Toms AP, Marshall TJ, Cahir J, et al. MR Imaging of early failing metal-on-metal total hip arthroplasty with surgical and histological correlation. *Eur Radiol* 2006;16(Suppl. 1):164.
  26. Davies AP, Willert HG, Campbell PA, et al. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am* 2005;87:18–27.
  27. Willert HG, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am* 2005;87:28–36.
  28. Weeden SH, Paprosky WG, Bowling JW. The early dislocation rate in primary total hip arthroplasty following the posterior approach with posterior soft-tissue repair. *J Arthroplasty* 2003;18:709–13.
  29. Dixon MC, Scott RD, Schai PA, et al. A simple capsulorrhaphy in a posterior approach for total hip arthroplasty. *J Arthroplasty* 2004;19:373–6.
  30. Stahelin T, Vienne P, Hersche O. Failure of reinserted short external rotator muscles after total hip arthroplasty. *J Arthroplasty* 2002;17:604–7.
  31. Jasani V, Richards P, Wynn-Jones C. Pain related to the psoas muscle after total hip replacement. *J Bone Joint Surg Br* 2002;84:991–3.
  32. Urban RM, Jacobs JJ, Tomlinson MJ, et al. Dissemination of wear particles to the liver, spleen, and abdominal lymph nodes of patients with hip or knee replacement. *J Bone Joint Surg Am* 2000;82:457–76.



## (iii) CT and MRI of hip replacements

John G Cahir

Andoni P Toms

### Abstract

Computed tomography (CT) and magnetic resonance imaging (MRI) are now useful imaging techniques in the evaluation of hip arthroplasty. The recognised problems of beam hardening in CT and magnetic susceptibility artefact in MRI have been significantly reduced.

MRI is useful for assessing the peri-prosthetic soft tissues and in evaluation of the painful replacement with normal plain films.

CT is better than plain films in evaluating bone stock around a hip replacement.

**Keywords** arthroplasty; CT; hip; MR

### Introduction

There are approximately 1.5 million hip arthroplasties performed world wide each year.<sup>1</sup> In 2000 it was reported that the NHS in England was performing over 30 000 hip replacements at a cost of £140 million.<sup>2</sup> Even though complication rates following arthroplasty are low the large number of arthroplasties performed means that complications are encountered frequently. All prostheses are likely to fail eventually given enough time as a result of infection, mechanical failure or flaws in surgical technique or prosthesis manufacture. Most patients have a satisfactory long-term outcome however an annual revision burden (derived by calculating the percentage of revision arthroplasties relative to the total number of primary and revision arthroplasties) of up to 17% has been described.<sup>3</sup>

Complications following hip arthroplasty include mechanical aseptic loosening, osteolysis secondary to granulomatous disease, heterotopic new bone formation, infection both superficial and deep, peri-prosthetic and prosthetic fracture and infrequently local nerve damage. Some of these complications are a source of morbidity and may require revision surgery.<sup>4,5</sup>

When revision surgery is performed it includes both the acetabular and femoral component in 37% of cases, the acetabular component only in 40% of cases and the femoral stem in approximately 22% of cases.<sup>6</sup>

Plain films are the first line investigation in evaluation of the painful total hip replacement. While plain radiographs are the

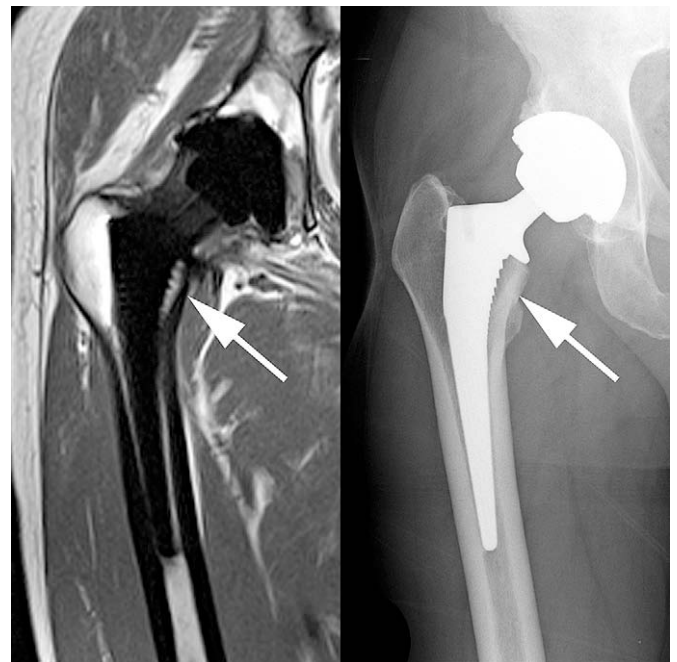
initial modality of evaluation they do have limitations due to the complex three dimensional geometry of the pelvis.<sup>7</sup>

Additional useful information can be obtained with both CT and MRI as the previous limiting factors of beam hardening in CT and susceptibility artefact in MRI can now be reduced considerably. CT is particularly useful in the evaluation of bone stock and the integrity of the medial acetabular wall and the anterior and posterior columns. MRI can now accurately evaluate the peri-prosthetic soft tissues, which is useful in the evaluation of infection or any other associated soft tissue abnormality. MRI has been shown to be a useful problem solving tool in unexplained failed total hip replacement. In one study it resulted in an unsuspected diagnosis in over 50% of cases.<sup>8</sup>

### Technical considerations

#### MRI

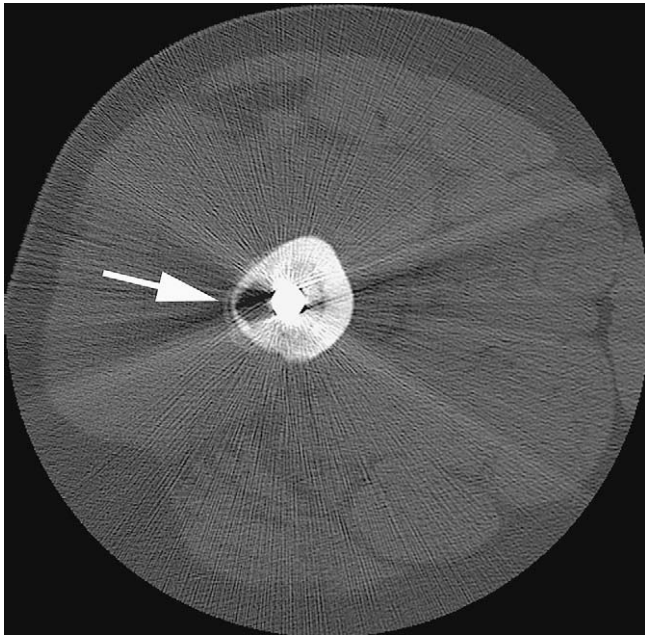
MRI when first introduced had limited scope in the imaging of metallic implants. This was due to magnetic susceptibility artefact produced by orthopaedic hardware. Magnetization of implants distorts the local field gradients causing proton dephasing which gives rise to signal voids and spatial distortion. The introduction of higher field strengths unfortunately results in even greater distortion. Technical alterations however can now largely overcome these problems. Artefact secondary to metallic implants can be reduced by increasing frequency encoding gradients, using fast spin echo techniques, reducing the volume of voxels and using short tau inversion recovery (STIR) sequences rather than using spectral fat suppression. The frequency encoding direction of the image is more susceptible to artefact. Selective orientation of the frequency and phase encoding directions of the acquisition can result in reduction of artefact in the regions of specific interest.<sup>4,9</sup>



**Figure 1** Coronal T1 MRI (left) and plain radiograph (right) of a titanium stem with a ceramic bearing. Metal artefact reduction allows differentiation of the cortex, marrow and even the ridges on the edge of the prosthesis (arrows).

**John G Cahir MRCP FRCS** is a Consultant Musculoskeletal Radiologist at the Department of Radiology, Norfolk and Norwich University Hospital, Colney Lane, Norwich, UK.

**Andoni P Toms FRCS FRCS** is a Consultant Musculoskeletal Radiologist at the Department of Radiology, Norfolk and Norwich University Hospital, Colney Lane, Norwich, UK.

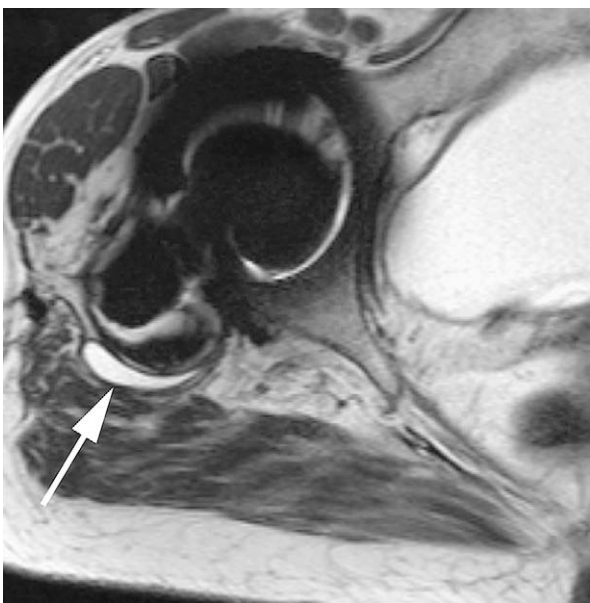


**Figure 2** Infection at the tip of a femoral prosthesis. There is cortical destruction and also a periosteal reaction (arrow).

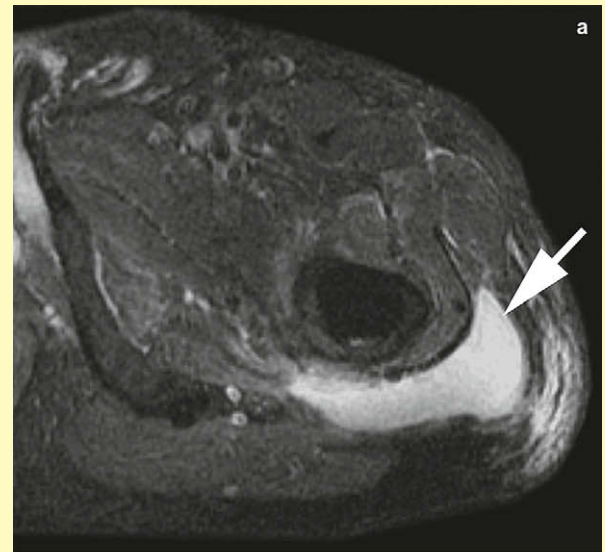
The exact imaging parameters necessary to optimize sequences vary depending on manufacturer and the available software and may require some trial and error in order to optimize imaging. The manipulation of these technical factors using metal artefact reduction sequences (MARS) enables visualisation of the cortex, cement and marrow around the femoral stem (Figure 1). It is more difficult to reduce artefact arising from the acetabular cup due to the oblique orientation in relation to the direction of the magnetic field.

#### CT

The most significant problem encountered when imaging orthopaedic hardware with CT is beam hardening. Beam hardening is



**Figure 3** A normal post-operative finding of a small seroma in the line of the post surgical scar.



**a** Axial STIR image of an infected collection (arrow) of the posterior aspect of the hip. There is oedema of the overlying soft tissues.  
**b** Axial T1 post gadolinium imaging demonstrates rim enhancement of the collection (arrowheads).

**Figure 4**

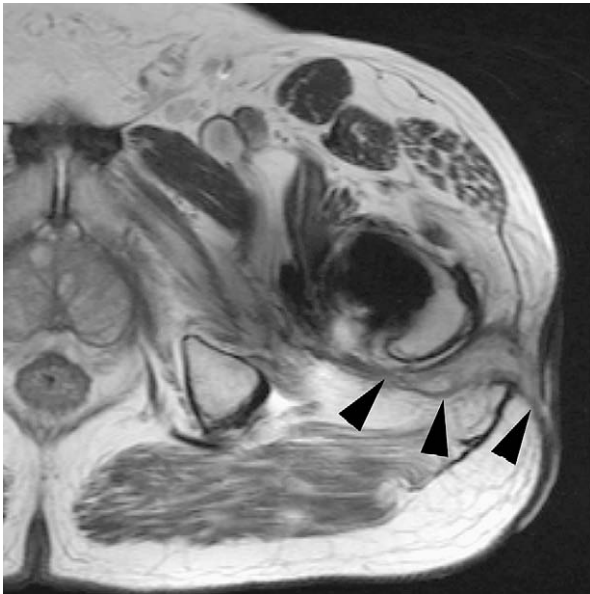
seen as alternating high and low attenuation lines which appear to radiate from implants. This problem can be overcome by increasing the output from the tube (mA and kVp). Technical factors mean that modern multislice CT scanners have a lower baseline artefact than older equipment. The data set obtained with multi slice CT can be reformatted in any plane with a soft tissue and smooth reconstruction filters with overlapping sections which also reduces artefact.<sup>10</sup>

#### Complications

##### Infection

In the early stages of infection plain radiographs will be normal in appearance. When infection is suspected ultrasound has an





**Figure 5** Axial T2 weighted image of a sinus tract (arrowheads) communicating with the posterior aspect of the joint.

important role in the assessment of the hip prosthesis as it can demonstrate the presence of periprosthetic collections and joint effusions. Ultrasound or fluoroscopy can guide aspiration of the joint. CT has been shown to be a useful diagnostic tool in the assessment of infection. Soft tissue changes are a more significant finding than periprosthetic bone abnormalities. Periprosthetic bone abnormalities do not allow differentiation of infection from complications not related to sepsis apart from periostitis which has a 100% specificity but only a 16% sensitivity (Figure 2). In cases of proven infection the bone may appear normal on CT in 25% of cases. Soft tissue abnormalities such as joint distension, fluid filled bursae and fluid collections in



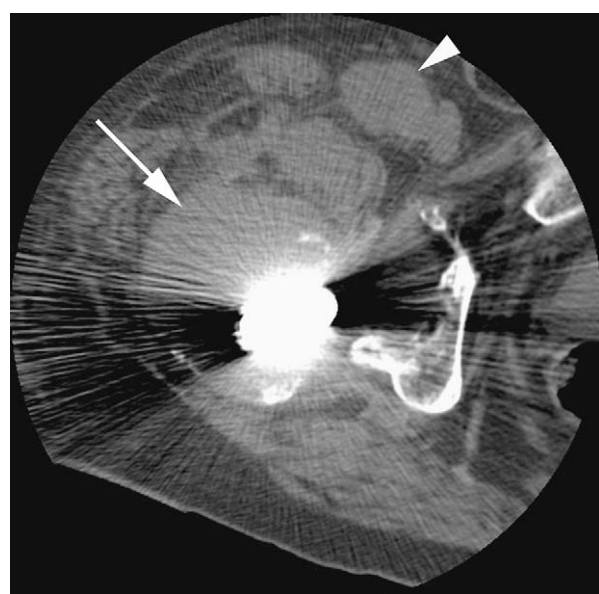
**Figure 7** Coronal CT showing extensive cortical destruction of the proximal femur secondary to osteolysis. The arrow heads depict the original outline of the lateral cortex. The lesser trochanter is fractured (arrow).

muscles and perimascular fat result in a 83% sensitivity, 96% specificity and 94% accuracy for infection when evaluating a painful prosthesis.<sup>5</sup>

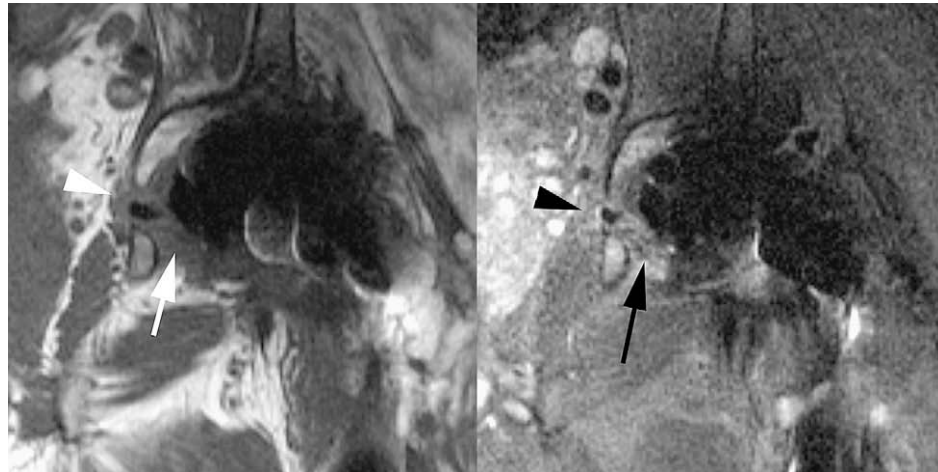
MRI is a better imaging modality than CT for demonstrating periprosthetic soft tissues. MRI is better at depicting soft tissues around the femoral stem than in the region of the acetabulum. It is not uncommon to see a small seroma in the line of



**Figure 6** Coronal T1 weighted image of an infected prosthesis which shows osteomyelitis of the medial calcar (arrowhead) and also infection tracking through a defect in the iliotibial tract (arrow).



**Figure 8** Axial CT image of a soft tissue mass and cortical destruction due to small particle disease (arrow). The osteolysis is of similar density to muscle (arrowhead).



**Figure 9** Coronal T1 image (left) of osteolysis (arrow) of the acetabulum. The small particle disease is of low signal. There is a breach of the medial wall (arrowheads) Coronal STIR imaging (right), where the osteolysis is of intermediate signal.

surgical approach (Figure 3). Infection on MRI has a signal intensity similar to that of fluid and will usually have an ill-defined margin (Figures 4a and 4b). MRI can depict bony involvement and assess the extent and route of sinus tracts. (Figures 5 and 6)

#### Osteolysis

Polyethylene wear particles are phagocytosed by macrophages which accumulate to create foreign body granulomas. The subsequent inducement of osteoclastic activity is a significant cause of periprosthetic osteolysis. Patients may remain asymptomatic until extensive bone loss has occurred.<sup>11–13</sup>

Plain radiographs are the first line investigation in evaluation of osteolysis but are limited in their ability to assess the complex anatomy of the hip particularly around the acetabulum. CT has been shown to be more accurate at identifying periacetabular osteolysis than plain radiographs and will show medial wall perforations not detectable on a plain film.<sup>7</sup> The presence of well demarcated lucencies adjacent to the socket or screws with absence of osseous trabeculae with CT are characteristic of osteolysis.<sup>14</sup>

Osteolysis may cause extensive cortical destruction (Figure 7). Soft tissue masses due to granulomatous disease have a similar density on CT to that of muscle (Figure 8). The typical signal characteristics of osteolysis on MRI are low signal on T1 imaging and low to intermediate high signal on T2 weighted imaging (Figure 9) there may be a surrounding low signal rim. Peripheral enhancement may be demonstrated on post gadolinium studies.<sup>4</sup>

The main indication for imaging of patients with osteolysis is to provide an accurate assessment of bone loss which is then helpful to the surgeon in planning revision surgery. Important areas which may be affected by osteolysis are the medial wall, the roof of the acetabulum and the anterior and posterior columns. Extensive osteolysis may result in pelvic discontinuity.<sup>15</sup>

#### Loosening

Mechanical loosening of the prosthesis is the most common cause of pain following hip arthroplasty. The plain radiograph findings of prosthesis movement, progressive lucency around

components over time and cement fracture are indicative. It is not known if CT or MRI offer any benefit over plain films in the assessment of mechanical loosening (Figure 10).

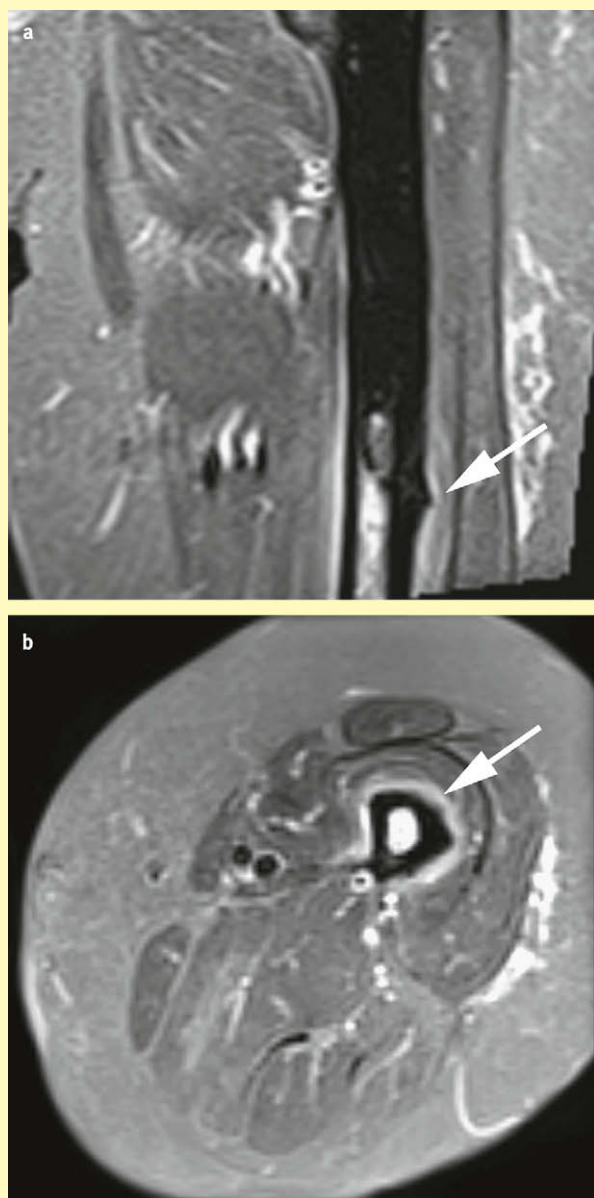
#### Fractures and pelvic discontinuity

Periprosthetic fractures occur for a variety of reasons such as osteolysis, loosening, stress risers and unfavourable biomechanics.<sup>15</sup> Periprosthetic fractures involving the bone adjacent to the tip of the femoral stem or the greater trochanter are usually diagnosed without difficulty with plain radiographs. Cross sectional imaging is useful in the evaluation of hip pain which may be due to occult fracture.<sup>16,8</sup>

Some fractures may be difficult to detect on plain radiographs. Subtle insufficiency fractures, stress fractures and fractures through areas of osteolysis may not be appreciated with plain radiographs. Both MRI and CT can demonstrate fractures immediately adjacent to metalwork however CT is better at evaluating the bone near the acetabulum (Figure 11). Pelvic discontinuity is a separation of the cephalad aspect of the pelvis from the caudal aspect as a result of fracture or bone loss, wherein the anterior and posterior columns are non-supportive. Bone loss may be due to infection, osteolysis or mechanical abrasion (Figure 12).<sup>17</sup> Pelvic discontinuity is more common in women and patients with rheumatoid arthritis. Pelvis radiotherapy is also a risk factor.<sup>18</sup> The integrity of the medial wall of the acetabulum and the anterior and posterior columns are important factors which determine surgical planning, technique and theatre time. Pelvic discontinuity may require reconstruction with a plate and anti-protrusion acetabular cage before insertion of bulk allograft and a cemented cup or a custom built acetabular component.<sup>17,19</sup>

#### Metal-on-metal soft tissue reaction

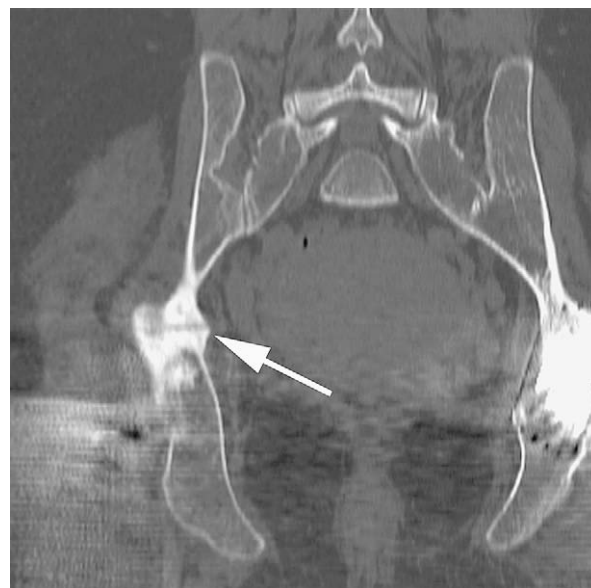
Technical advances with improvements in machine tolerances have result in the development of second and third generation metal on metal prostheses. These prostheses have the potential advantage of avoiding the problems of small particle disease and osteolysis which are recognized complications associated with conventional metal on polyethylene prostheses.<sup>20</sup> An unusual



**a** STIR image with cortical thickening and periosteal reaction (arrow) several cms distal to the tip of the femoral stem. Oedema is also present in the bone marrow. **b** Axial STIR image of the same patient with a fluid periosteal stress reaction (arrow) surrounding the femur.

**Figure 10**

reaction associated with metal on metal (cobalt chromium alloy) prostheses, both total and resurfacing has been described. A perivascular infiltration of lymphocytes and an accumulation of plasma cells in association with macrophages containing metallic wear debris particles have been demonstrated.<sup>21–23</sup> These histological appearances have been termed aseptic lymphocytic vasculitis-associated lesions. The clinical presentation is typically early recurrence of pain or discomfort after the initial surgery which may be initially induced with exercise.<sup>24</sup>



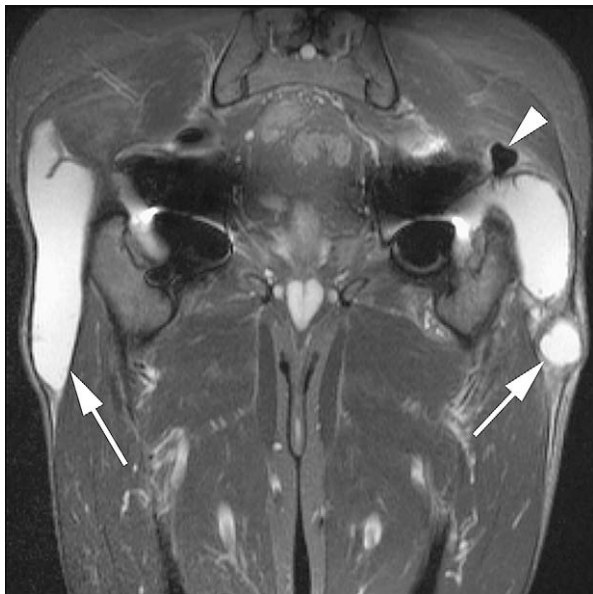
**Figure 11** Coronal CT in a patient with bilateral THRs, there is an insufficiency fracture (arrow) of the posterior column on the right side.

The plain radiographs are typically normal and may therefore be falsely reassuring. The imaging findings suggestive of a metal on metal soft tissue reaction are a fluid collection which has a thick irregular wall surrounding the neck of the prosthesis, bone marrow oedema in the proximal femoral diaphysis, and oedema of the surrounding musculature. The associated collection may sometimes extend a considerable distance from the joint (Figure 13). The associated soft tissue reaction and necrosis may damage important adjacent structures (Figure 14).<sup>20,23</sup> The exact cause of this specific soft tissue reaction, which has not been described with other implants, has not been established.



**Figure 12** Coronal CT image showing pelvic discontinuity (arrow). There is also extensive wear of the acetabular component of this right THR with an eccentric position of the head of the femoral component in relation to the cup.





**Figure 13** Coronal STIR image which showing large non-infective collections (arrows) associated with bilateral resurfacing prostheses. There is a low signal lymph node (arrowhead) which contained metallic particles. The patient required revision surgery of both hips.

Theories include a hypersensitivity reaction with a type IV immune response to the metal alloy components.<sup>24</sup>

#### Tendon attachments

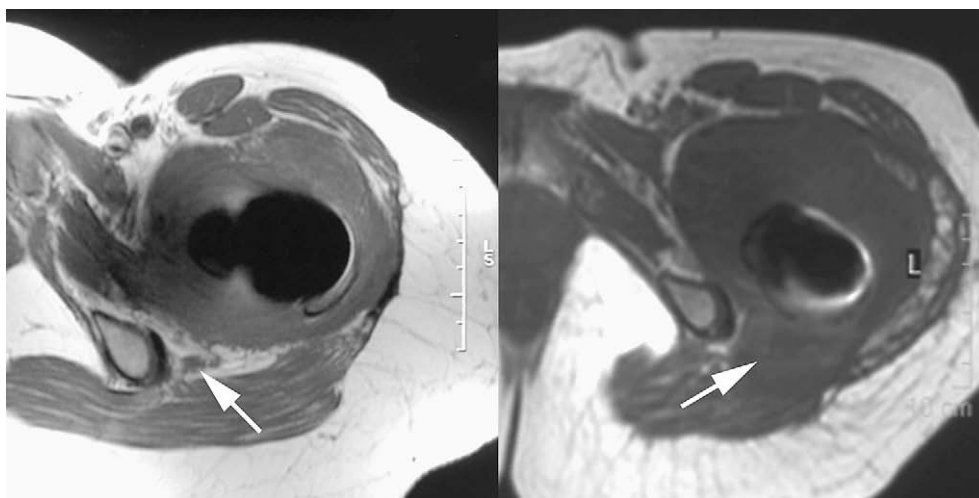
Avulsion of the gluteal tendons is a recognized complication following hip arthroplasty. Patients may have weakness of hip abduction however many cases appear to be asymptomatic. Short external rotator avulsion is also a recognized complication following THR typically when a posterior approach has been used.<sup>25</sup> Muscle atrophy and tendon avulsion may also be related to severe osteoarthritis and may be findings which are present prior to arthroplasty.

When the posterior approach to the joint is used the piriformis tendon is usually divided and is not necessarily repaired.<sup>26,27</sup>



**Figure 15** The psoas bursa (arrow) is visible even in close proximity to the acetabular cup. The acetabular cup usually produces more artefact than the femoral stem on MRI.

The short external rotators are best imaged with axial T1W imaging. Coronal imaging is useful in evaluation of the gluteus minimus and medius attachments.<sup>25</sup> A transgluteal approach may predispose to greater trochanteric bursitis. In the presence of clinical uncertainty of the diagnosis then MRI will demonstrate fluid like signal deep to the gluteus maximus tendon. There may be also associated gluteus minimus or minimus tendinopathy. It is however common to visualise some fluid signal on fat saturated T2 weighted imaging in the region of the greater trochanter in asymptomatic patients.<sup>28</sup> Ilio-psoas bursitis may be a cause of hip pain following arthroplasty. This can be seen on MRI even though it is immediately adjacent to the acetabular cup. (Figure 15)



**Figure 14** THR with a metal-on-metal reaction. The left image shows the sciatic nerve (arrow) with surrounding fat with an adjacent metal-on-metal soft tissue reaction. The image on the right shows enlargement of the collection with has now engulfed the sciatic nerve (arrow) resulting in a sciatic nerve palsy.

## Metallosis

Metallosis is an uncommon condition that occurs when excessive wear produces shedding of a large number of metallic particles into the joint. It typically occurs when there has been failure or fracture through the polyethylene liner of the acetabular component. The resultant abnormal articulation of the femoral head with the metal backing of the acetabular component causes generation of the particles. Metallosis can also occur in excessive wear in metal on metal arthroplasty.<sup>29,30</sup>

Titanium components are more commonly associated with metallosis than chromium cobalt prosthesis.<sup>31</sup> The metallic debris causes a chronic inflammatory reaction and stimulates periprosthetic osteolysis.<sup>32</sup> The metallosis may result in a linear radiodensity which outlines the periphery of the joint capsule. This appearance has also been termed the bubble sign.<sup>33</sup> It may also appear as homogeneous increased density of the pseudocapsule of the joint on plain radiography. The lack of corticomedullary differentiation differentiates it from heterotopic ossification. Aspiration of affected joints typically reveals thick black oily liquid. The periprosthetic tissues may be stained black and be infiltrated with histiocytes containing metal debris.<sup>15,34</sup>

## Conclusion

Relatively simple modifications in imaging protocols can produce adequate metal artefact reduction on standard equipment. CT and MRI are now effective tools in the evaluation of the symptomatic hip arthroplasty. ◆

## REFERENCES

- Looney RJ, Boyd A, Totterman S, et al. Volumetric computerized tomography as a measurement of periprosthetic acetabular osteolysis and its correlation with wear. *Arthritis Res* 2002; **4**: 59–63.
- Public Accounts Committee. Hip replacements: getting it right first time. House of Commons; 2000.
- Kurtz S, Mowat F, Ong K, Chan N, Lau E, Halpern M. Prevalence of primary and revision total hip and knee arthroplasty in the United States from 1990 through 2002. *J Bone Joint Surg Am* 2005; **87**: 1487–97.
- White LM, Kim JK, Mehta M, et al. Complications of total hip arthroplasty: MR imaging-initial experience. *Radiology* 2000; **215**: 254–62.
- Cyteval C, Hamm V, Sarabere MP, Lopez FM, Maury P, Taourel P. Painful infection at the site of hip prosthesis: CT imaging. *Radiology* 2002; **224**: 477–83.
- Bozic KJ, Durhakula S, Berry DJ, et al. Differences in patient and procedure characteristics and hospital resource use in primary and revision total joint arthroplasty: a multicenter study. *J Arthroplasty* 2005; **20**: 17–25.
- Leung S, Naudie D, Kitamura N, Walde T, Engh CA. Computed tomography in the assessment of periacetabular osteolysis. *J Bone Joint Surg Am* 2005; **87**: 592–7.
- Johnston C, Kerr J, Ford S, O'byrne J, Eustace S. MRI as a problem-solving tool in unexplained failed total hip replacement following conventional assessment. *Skeletal Radiol* 2007; **36**: 955–61.
- Sofka CM, Potter HG. MR imaging of joint arthroplasty. *Semin Musculoskelet Radiol* 2002; **6**: 79–85.
- White LM, Buckwalter KA. Technical considerations: CT and MR imaging in the postoperative orthopedic patient. *Semin Musculoskelet Radiol* 2002; **6**: 5–17.
- Wimhurst JA. The pathogenesis of aseptic loosening. *Curr Orthop* 2002; **16**: 407–10.
- Harris WH, McCarthy JC, O'Neill DA. Femoral component loosening using contemporary techniques of femoral cement fixation. *J Bone Joint Surg Am* 1982; **64**: 1063–7.
- Chiang PP, Burke DW, Freiberg AA, Rubash HE. Osteolysis of the pelvis: evaluation and treatment. *Clin Orthop Relat Res* 2003; **164**–74.
- Puri L, Lapinski B, Wixson RL, Lynch J, Hendrix R, Stulberg SD. Computed tomographic follow-up evaluation of operative intervention for periacetabular lysis. *J Arthroplasty* 2006; **21**: 78–82.
- Cahir JG, Toms AP, Marshall TJ, Wimhurst J, Nolan J. CT and MRI of hip arthroplasty. *Clin Radiol* 2007; **62**: 1163–71 (discussion 1172–3).
- Potter HG, Nestor BJ, Sofka CM, Ho ST, Peters LE, Salvati EA. Magnetic resonance imaging after total hip arthroplasty: evaluation of periprosthetic soft tissue. *J Bone Joint Surg Am* 2004; **86-A**: 1947–54.
- DeBoer DK, Christie MJ, Brinson MF, Morrison JC. Revision total hip arthroplasty for pelvic discontinuity. *J Bone Joint Surg Am* 2007; **89**: 835–40.
- Berry DJ, Lewallen DG, Hanssen AD, Cabanela ME. Pelvic discontinuity in revision total hip arthroplasty. *J Bone Joint Surg Am* 1999; **81**: 1692–702.
- Choplin RH, Henley CN, Edds EM, Capello W, Rankin JL, Buckwalter KA. Total hip arthroplasty in patients with bone deficiency of the acetabulum. *Radiographics* 2008; **28**: 771–86.
- Toms AP, Marshall TJ, Cahir J, et al. MR Imaging of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008; **63**: 49–58.
- Park YS, Moon YW, Lim SJ, Yang JM, Ahn G, Choi YL. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am* 2005; **87**: 1515–21.
- Davies AP, Willert HG, Campbell PA, Learmonth ID, Case CP. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am* 2005; **87**: 18–27.
- Fang CS, Harvie P, Gibbons CL, Whitwell D, Athanasou NA, Ostlere S. The imaging spectrum of peri-articular inflammatory masses following metal-on-metal hip resurfacing. *Skeletal Radiol* 2008; **37**: 715–22.
- Willert HG, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am* 2005; **87**: 28–36.
- Pfrrmann CW, Notzli HP, Dora C, Hodler J, Zanetti M. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology* 2005; **235**: 969–76.
- Weeden SH, Paprosky WG, Bowling JW. The early dislocation rate in primary total hip arthroplasty following the posterior approach with posterior soft-tissue repair. *J Arthroplasty* 2003; **18**: 709–13.
- Stahelin T, Vienne P, Hersche O. Failure of reinserted short external rotator muscles after total hip arthroplasty. *J Arthroplasty* 2002; **17**: 604–7.
- Blankenbaker DG, Ullrick SR, Davis KW, De Smet AA, Haaland B, Fine JP. Correlation of MRI findings with clinical findings of trochanteric pain syndrome. *Skeletal Radiol* 2008; **37**: 903–9.
- Gambera D, Carta S, Crainz E, Fortina M, Maniscalco P, Ferrata P. Metallosis due to impingement between the socket and the femoral head in a total hip prosthesis. A case report. *Acta Biomed* 2002; **73**: 85–91.
- Cipriano CA, Issack PS, Beksac B, Della Valle AG, Sculco TP, Salvati EA. Metallosis after metal-on-polyethylene total hip arthroplasty. *Am J Orthop* 2008; **37**: E18–25.

- 31** Weissman BN, Scott RD, Brick GW, Corson JM. Radiographic detection of metal-induced synovitis as a complication of arthroplasty of the knee. *J Bone Joint Surg Am* 1991; **73**: 1002–7.
- 32** Heffernan EJ, Alkubaidan FO, Nielsen TO, Munk PL. The imaging appearances of metallosis. *Skeletal Radiol* 2008; **37**: 59–62.
- 33** Su EP, Callander PW, Salvati EA. The bubble sign: a new radiographic sign in total hip arthroplasty. *J Arthroplasty* 2003; **18**: 110–2.
- 34** Ottaviani G, Catagni MA, Matturri L. Massive metallosis due to metal-on-metal impingement in substitutive long-stemmed knee prosthesis. *Histopathology* 2005; **46**: 237–8.

### Learning Points

- CT is better than plain radiographs for assessing bone stock around a THR.
- Metal artefact reduction sequences can successfully be implemented on most clinical MR machines.
- MR is useful for assessing peri-prosthetic soft tissues.
- In a painful THR with normal radiographs MR is a useful diagnostic tool.



# Early failure of the Ultima metal-on-metal total hip replacement in the presence of normal plain radiographs

S. T. Donell,  
C. Darrah,  
J. F. Nolan,  
J. Wimhurst,  
A. Toms,  
T. H. W. Barker,  
C. P. Case,  
J. K. Tucker,  
Norwich Metal-on-Metal Study Group

*From the Norfolk & Norwich University Hospital, Norwich, United Kingdom*

■ S. T. Donell, FRCS(Orth), MD, Honorary Professor  
■ C. Darrah, RGN, MSc, Research Nurse Manager  
■ J. F. Nolan, FRCS(Orth), Consultant Orthopaedic Surgeon  
■ J. Wimhurst, MChir, FRCS(Trauma & Orth), Consultant Orthopaedic Surgeon  
■ J. K. Tucker, FRCS, Consultant Orthopaedic Surgeon  
Department of Trauma & Orthopaedics  
■ A. Toms, FRCR, Consultant Radiologist  
Department of Radiology  
■ T. H. W. Barker, FRCPath, Consultant in Histopathology  
Department of Histopathology  
Norfolk & Norwich University Hospital, Colney Lane, Colney, Norwich NR4 7UZ, UK.  
■ C. P. Case, FRCPath, MRCPath, Senior Lecturer in Histopathology  
Bristol Implant Research Centre, Avon Orthopaedic Centre, Southmead Hospital, Bristol BS10 5NB, UK.

Correspondence should be sent to Professor S. T. Donell; e-mail: simon.donell@nnuh.nhs.uk

©2010 British Editorial Society of Bone and Joint Surgery  
doi:10.1302/0301-620X.92B11.24504 \$2.00

*J Bone Joint Surg [Br]*  
2010;92-B:1501-8.

Received 21 January 2010;  
Accepted after revision 18 June 2010

**Metal-on-metal total hip replacement has been targeted at younger patients with anticipated long-term survival, but the effect of the production of metal ions is a concern because of their possible toxicity to cells. We have reviewed the results of the use of the Ultima hybrid metal-on-metal total hip replacement, with a cemented polished tapered femoral component with a 28 mm diameter and a cobalt-chrome (CoCr) modular head, articulating with a 28 mm CoCr acetabular bearing surface secured in a titanium alloy uncemented shell.**

Between 1997 and 2004, 545 patients with 652 affected hips underwent replacement using this system. Up to 31 January 2008, 90 (13.8%) hips in 82 patients had been revised. Pain was the sole reason for revision in 44 hips (48.9%) of which 35 had normal plain radiographs. Peri-prosthetic fractures occurred in 17 hips (18.9%) with early dislocation in three (3.3%) and late dislocation in 16 (17.8%). Infection was found in nine hips (10.0%).

At operation, a range of changes was noted including cavities containing cloudy fluid under pressure, necrotic soft tissues with avulsed tendons and denuded osteonecrotic upper femora. Corrosion was frequently observed on the retrieved cemented part of the femoral component. Typically, the peri-operative findings confirmed those found on pre-operative metal artefact reduction sequence MRI and histological examination showed severe necrosis.

**Metal artefact reduction sequence MRI proved to be useful when investigating these patients with pain in the absence of adverse plain radiological features.**

Some early designs of total hip replacement (THR) included metal-on-metal (MoM) bearings<sup>1</sup> but the variability in the tolerance at that time during manufacture caused high rates of failure. However, when early failure did not occur the implants outlasted the contemporary Charnley metal-on-polyethylene design<sup>2-4</sup> because the wear rates were considerably less.<sup>5</sup>

Low wear rates and minimal soft-tissue necrosis have been observed after 30 years for the McKee-Farrar MoM THR<sup>6</sup> and a custom-made prosthesis,<sup>7</sup> demonstrating that a MoM bearing can achieve long survival. Subsequently they have been re-introduced in THR<sup>8</sup> as a result of improved understanding of the bearing tribology,<sup>9</sup> manufacturing processes and the quest for increased longevity of implants. The use of this type of implant has been targeted at younger patients. However, the production of metal debris and corrosion has been a concern since metal ions may have a toxic effect on cells. There is evidence that an immunological reaction may ensue with the theoretical risk of an oncogenic response.<sup>10</sup>

The production of metal ions may depend on a number of factors. Wear between the bearing surfaces has been reduced by improved design and manufacturing.<sup>11</sup> However, increased production of metal ions is reported to be associated with excessive inclination of the acetabular component,<sup>12-16</sup> the use of incompatible metals,<sup>17</sup> excess movement at the cement-stem interface in the presence of poor positioning of the stem<sup>18</sup> and impingement.<sup>19</sup> Whether production of metal ions is directly proportional to use of the implant is open to debate.<sup>20-22</sup> Their production may provoke a hypersensitivity reaction, resulting in an immune response.<sup>23-29</sup>

When a metal-on-polyethylene THR fails there is, typically, visible evidence of loosening on plain radiographs, most notably radiolucent lines.<sup>30</sup> These are due, in part, to the biological response to wear debris which leads to an increase in osteoclastic activity and resultant re-absorption of bone.<sup>31,32</sup> A recent study has shown failure in a MoM THR when the plain radiographs appeared to be normal, but

metal artefact suppression MR scans detected gross abnormalities.<sup>33</sup>

We report a series of patients with a MoM bearing and a conventional design of femoral stem who underwent revision for a variety of reasons including pain, but who had normal radiographs.

### Patients and Methods

Between February 1997 and August 2004, 545 patients with 652 affected hips underwent hybrid THR with a MoM cobalt-chromium (Co-Cr) alloy bearing. Of these, 434 hips were treated in a public hospital (Norfolk and Norwich University Hospital) and 218 in a private hospital (Spire Hospital, Norwich).

There were 302 males and 243 females with a mean age of 57 years (15 to 81). Of these, 366 (67.1%) had a unilateral MoM THR with an unaffected contralateral hip and 107 (19.6%) had bilateral MoM THRs. A further 75 (13.8%) had a MoM THR on one side, and a non-MoM implant on the other. Additionally, seven patients (1.3%) had the MoM implant as a revision from a failed non-MoM primary THR.

The diagnosis was primary osteoarthritis (OA) in 487 patients (89.3%) and in 58 (10.6%) secondary OA from developmental dysplasia of the hip in 18, avascular necrosis in ten, post-traumatic OA in ten, slipped upper femoral epiphysis or Perthes' disease in seven and Paget's disease in one. A total of 12 patients (1%) had an inflammatory arthropathy, ankylosing spondylitis in eight and rheumatoid arthritis in four.

The first 20 patients were recruited as part of an ethically-approved multicentre clinical investigation into the safety and efficacy of a new MoM THR which was managed and funded by Johnson and Johnson Professional, New Milton, United Kingdom, now DePuy International Ltd, Leeds, United Kingdom. The purpose was to demonstrate conformance of the device with the essential requirements of the Council of European Communities directive on medical devices 93/42/EEC,<sup>34</sup> and was approved by the United Kingdom regulatory agency. The use of different femoral components was allowed in the various centres, but in our study all patients had the same femoral component. The inclusion criteria were the same as for standard primary THR using a cementless acetabular component and a cemented femoral component, namely pain, deformity and loss of function which was not responsive to medical treatment. The exclusion criteria were revision THR, rheumatoid arthritis, previous sepsis of the hip, previous inclusion of the contralateral hip in the study, recent use of high-dosage corticosteroids, recent therapeutic radiation, metabolic disorders of calcified tissue, Charcot arthropathy, unsuitability for cementless fixation of the acetabular component, the requirement of a segmental acetabular bone graft including protrusion, previous evidence of sensitivity to Co-Cr or titanium alloy (Ti-6Al-4V), inability to attend post-operative follow-up visits and psychosocial



Fig. 1

Photograph of the components of the implant.

factors which might limit rehabilitation. The clinical and radiological data were gathered prospectively pre-operatively and at review at six weeks, six months and then annually.

On completion of the initial investigation on our 20 patients (and a further 30 recruited from elsewhere), the relevant *Conformité Européenne* (CE) marking was obtained by the manufacturer and the patients were recruited as part of a post-market clinical follow-up study. The same inclusion criteria were followed, but the presence of inflammatory arthropathy, contralateral hip replacement and inability to attend follow-up were no longer applied as exclusion criteria. Outpatient review followed the routine department protocol in the public hospital. This consisted of a clinical and radiological review at six weeks after operation and then at yearly intervals. Patients treated in private hospitals were followed up under the supervision and protocol of the individual surgeon (JFN, JKT).

**The implant.** The Ultima hip replacement system (DePuy International Ltd) was used in all the patients. The acetabular component was uncemented of wrought Ti-6Al-4V porous-coated outer shell with three holes for supplementary Ti-6Al-4V screw fixation. There were 11 sizes of shell from 48 mm to 68 mm in increments of 2 mm. An acetabular insert manufactured from high-carbon Co-Cr molybdenum alloy (CoCrMo) with a 28 mm diameter hemispherical articular surface was secured in the shell by a taper arrangement. A 10° augmented insert was available if required. The modular femoral head was 28 mm in diameter and made of low carbon CoCrMo alloy, with a diametral clearance of 0.030 mm to 0.075 mm to allow polar bearing. The stem was a standard cemented Ultima TPS collarless, double-tapered, polished stem made from CoCrMo alloy (Fig. 1).

The femoral component was cemented with either plain Simplex, Simplex P (Howmedica International, Limerick,

Ireland) low-viscosity cement or Palacos R (Biomet Inc, Warsaw, Indiana) high-viscosity cement with gentamicin.

**Operative technique.** In all, 647 operations (99.2%) were carried out under the care of three specialist hip surgeons, including two authors (JFN, JKT) one of whom performed 407 (62.4%). Of the operations undertaken, 605 (92.8%) were performed by consultants as the primary surgeon and the rest by experienced trainees.

The approach chosen was according to the preference of the surgeon, either anterolateral detaching the anterior two-thirds of gluteus medius (most surgeons) or posterolateral. The acetabulum was prepared using hemispherical reamers and the acetabular component selected was 2 mm larger than the last reamer. Any defects were packed with autograft. This was used in 38 (5.8%) procedures. The femoral component was inserted using third-generation cementing techniques.<sup>35</sup> The surgeons who favoured the anterolateral approach used predominantly Palacos R cement and the surgeon using the posterolateral approach Simplex P cement.

**Renal function.** Since chronic renal failure is recognised to influence the levels of metal ions,<sup>36</sup> the pre- and post-operative renal function was determined from the serum levels of urea and creatinine extracted from the public hospital software program (ICE, Anglia Healthcare Systems, version 1, Norwich, United Kingdom). The initial value was obtained at the nearest time to the primary operation, but not more than one month before, and similarly any pre-revision value was obtained not more than one month before the revision procedure. The latest value available was taken as the level after the revision. Renal impairment was considered to be present if the recorded value exceeded the normal laboratory range (urea 2.5 mmol/l to 10.5 mmol/l and creatinine 44 µmol/l to 124 µmol/l).

**Radiological analysis.** Plain radiographs of the pelvis were taken with standard standing anteroposterior (AP) and lateral views. The following measures were acquired from each radiograph and measured independently by consensus opinion: acetabular inclination, the height of the acetabular component, leg length, lateral offset and alignment of the femoral component according to a published protocol.<sup>37</sup> They were reviewed and independently analysed. The position of the component, the adequacy of cementing using the Barrack system<sup>38</sup> and the presence of radiolucent lines, bone loss, and soft-tissue swelling were noted. Peri-prosthetic fractures were classified according to the Vancouver system.<sup>39</sup>

As the study advanced it became apparent that the plain radiographs were not showing any adverse features when some patients with normal radio-isotopic bone scans and normal blood parameters reported pain. MRI was undertaken using metal artefact reduction sequences with one of two 1.5T MR machines (Siemens Symphony; Siemens, Erlangen, Germany or GE Sigma; GE Healthcare, Slough, United Kingdom). MR acquisitions included axial T1- and T2-weighted fast spin echo, coronal T1-weighted and Short

T1 Inversion Recovery and a sagittal T2-weighted fast spin echo with matrix sizes up to 320 × 384 voxels and extended receiver bandwidths (e.g. 640 Hz/voxel using the Siemens apparatus). Phase and frequency-encoding directions were arranged so that the peri-prosthetic soft tissues could be visualised optimally on at least two of the sequences.<sup>33</sup>

All the pre-operative MR examinations were reviewed on diagnostic Picture Archive Communication System workstations (GE Centricity, GE Healthcare) using 2K high-resolution monitors (Barco, Kortrijk, Belgium) by two trained musculoskeletal radiologists (including AT). The findings were recorded after agreement of the two observers. Atrophy or avulsion of the short external rotators was recorded as positive if there was evidence of disease in any one of the muscles.

**Histological analysis.** Tissues retrieved at revision surgery were processed by standard haematoxylin and eosin stains, as well as a connective tissue stain, Elastin Ponceau S.

**Statistical analysis.** Survival of the implant was estimated according to the Kaplan-Meier method using the SPSS version 10.0 software (SPSS Inc., Chicago, Illinois) with 95% confidence intervals (CI). Failure was defined as revision for any reason. The time to revision was calculated as the interval between the date of implantation and that of revision. Patients without revision were censored on the date of death. The chi-squared test was used with significance set at a p-value ≤ 0.05.

## Results

**Revision cohort.** Up to 31 January 2008, 82 patients had undergone revision of their primary MoM THR, of whom four were revised to another MoM THR before the nature of the problem was appreciated. They have since undergone a further revision. Another four patients have had bilateral revisions. Therefore the revision series comprised 90 hips. There were 44 males and 38 females with a mean age of 51 years (15 to 75). The mean body mass index (BMI) was 29.5 g/cm<sup>2</sup> (21 to 42.5).

Revision of all the components was performed in all except five hips. Of the latter, two had exchange of the acetabular component and one an isolated revision of the femoral component, which subsequently required total revision. In the remaining two patients only the head and liner were exchanged, one of which had a proven infection.

When it became clear that there was an unacceptably high rate of failure, the Research and Ethics Committees were informed and at their request an investigation was undertaken by a group which included clinicians, radiologists, histopathologists, research staff, a statistician and representatives of DePuy International Ltd. The United Kingdom national competent authority (the Medicines and Healthcare products Regulatory Agency) was informed by the manufacturer.

**Clinical findings.** Generally patients were pleased with the outcome, but if a problem emerged they would describe the

**Table I.** The causes of revision in the 90 failed total hip replacements correlated with the plain radiological findings

Cause	Normal (%)	Loosening (%)	Total (%)
Pain	35 (39.0)	9 (10.0)	44 (48.9)
Peri-prosthetic fracture	14 (15.6)	3 (3.3)	17 (18.9)
Dislocation			
Early	3 (3.3)	0 (0.0)	3 (3.3)
Late	12 (13.3)	4 (4.0)	16 (17.8)
Infection	6 (6.7)	3 (3.3)	9 (10.0)
Implant fracture	1 (1.1)	0 (0.0)	1 (1.1)
Total	71 (78.9)	19 (21.1)	90 (100.0)

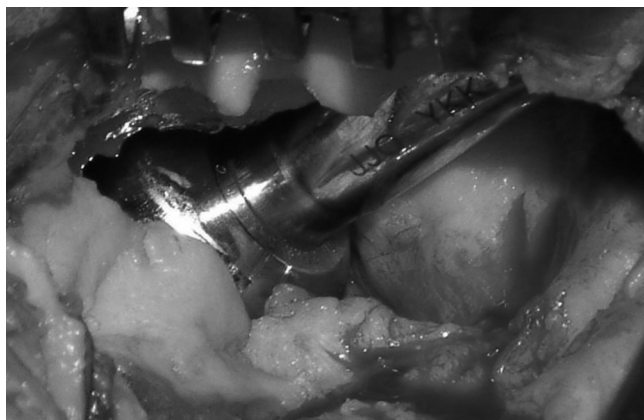


Fig. 2

Intra-operative photograph showing moderate soft-tissue changes with velvety-white synovium around the prosthesis.



Fig. 3

Intra-operative photograph showing severe changes with the upper femur denuded of soft-tissue attachments.

replacement as “not quite right”, usually at least two years post-operatively. Deterioration then followed with the predominant complaint being pain, typically felt in the groin and thigh. Of the 90 hips (13.8%) which underwent revision, 35 (38.9%) had normal radiographs and no pain, and 57 (69.5%) had soundly fixed implants (Table I). One implant fractured through the neck of the femoral component in a patient with a BMI of 42.5 kg/cm<sup>2</sup>. The patients with late dislocations were found to have avulsion of the tendons of the glutei and short external rotators. Two patients had palpable masses because of collections of fluid, and one had a sinus in the presence of a negative microbiological culture. However, positive cultures were found in the peri-implant soft tissues at revision. One patient developed late palsy of the sciatic nerve as a result of a soft-tissue reaction involving the nerve directly. This palsy did not improve after revision.

At revision a collection of opaque fluid was found, which was sometimes under pressure, and soft-tissue necrosis (Fig. 2). This could have been associated with avulsed gluteal tendons. The denuded upper femur was associated with avascularity of its surface (Fig. 3). The prosthesis was

typically well fixed with an intact cement mantle, except when a peri-prosthetic fracture had occurred. In all cases the stem was corroded where it was in contact with the cement (Fig. 4). **Infection.** In six patients positive bacterial cultures were found at revision. Three were due to *Staphylococcus aureus* (two being methicillin-resistant), one to *Streptococcus* group G, one to *Enterococcus*, and one to a coliform species. A total of eight patients had histological findings consistent with infection, of whom one was a proven case with *Enterococcus*. Three had been treated for infection before revision. None of these eight patients has required another revision. Therefore nine patients had infected implants giving an overall rate of infection for the whole series of 0.1% and for the revision series of 10.0%.

**Implant survival and revision rates.** The entire series had a failure rate of 13.8% (95% CI 9.3 to 19.5) at five years.

**Peri-prosthetic fractures.** These occurred in 16 patients. Three were of type Vancouver AG, one of which was noted intra-operatively, but not on the plain radiographs. One was Vancouver AL, also found intra-operatively, and the



Fig. 4

Photograph of the explanted femoral stem showing severe corrosion confined to the interface of the stem and acrylic cement.



Fig. 5a



Fig. 5b

Post-operative anteroposterior radiographs showing a) the normal post-operative appearance in a symptomatic patient with abnormal MR findings and b) a typical Vancouver-B1 peri-prosthetic fracture with well-fixed components (different patient).

rest were Vancouver B1, the typical pattern being involvement of the medial femoral cortex (Fig. 5) and the lesser trochanter, usually comminuted. All except two patients were men. This differed significantly from the pattern in the total series (chi-squared test,  $p < 0.001$ ). One of the female patients had a transverse fracture at the tip of the stem and pre-fracture radiographs showed lateral cortical thickening suggesting a stress reaction. In two men the fracture followed a fall from a height, but the rest were caused by low-energy trauma.

**Other complications in the main series.** Details of complications not requiring revision surgery were available for 574 hips (88.0%). There were 15 cases of venous thromboembolism but no pulmonary emboli, and nine hips had a

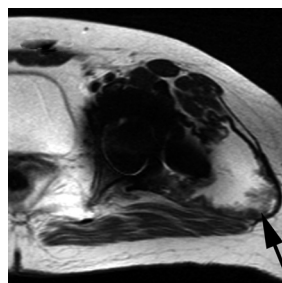


Fig. 6a

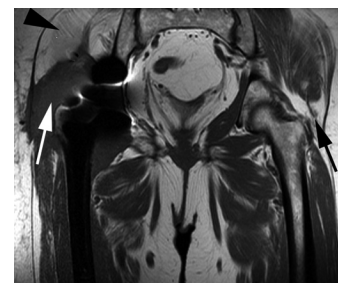


Fig. 6b

Metal artefact reduction sequence MR axial scans showing a) a peri-prosthetic high-signal collection of fluid surrounded by an irregular low-signal capsule (arrow) (T2-weighted). The wall of an abscess cavity would typically be intermediate to high signal on a T2-weighted image) and b) a peri-prosthetic collection extending laterally through the fascia lata (white arrow) (T1-weighted).

persistent wound leakage. Four patients described squeaking or clunking. There was one case of iliotibial band defect.

**Renal function.** Of the 82 patients (90 revised hips) full results were available for 67, of whom 12 had abnormal results before the primary procedure which had subsequently returned to normal. A further seven had abnormal results before their revision procedure, which subsequently returned to normal post-operatively. In four patients their results were persistently abnormal and one had chronic renal failure.

**Radiological findings.** Of the 90 revised hips the immediate primary post-operative radiographs were available for 83. The Barrack cement grade was A in 67 (80.7%), B in 13 (15.7%) and D in three (3.6%) and none were found to be grade C. The mean acetabular inclination ( $n = 82$ ) was  $46^\circ$  ( $27^\circ$  to  $60^\circ$ ) with 62 hips lying between  $30^\circ$  and  $50^\circ$ . The mean height of the acetabular component ( $n = 82$ ) was  $+3$  mm ( $-15$  to  $+22$ ). The mean lateral offset ( $n = 81$ ) was  $2.8$  mm ( $-30$  to  $+28$ ). The mean leg-length discrepancy ( $n = 81$ ) was  $0.4$  mm ( $-25$  mm to  $+18$  mm). On the AP radiograph ( $n = 83$ ), the mean position of the stem was varus  $2^\circ$  ( $3^\circ$  valgus to  $8^\circ$  varus). A total of 16 patients had evidence of loosening with radiolucent lines; three associated with peri-prosthetic fracture and four with late dislocation. Of the nine with painful aseptic loosening, migration was not noted (Table I).

In 43 patients (45 hips) MRI had been undertaken before revision. The mean age of these patients was 62 years (38 to 84). The mean time from the primary THR to MRI was 40 months (11 to 85) and the mean time to revision from the MRI was 5.6 months (0.1 to 30.3).

In 41 of the 45 MR scans a fluid-filled peri-prosthetic cavity was observed (Fig. 6a). It was lined by a ragged wall up to 1 cm thick and was contiguous with the neck of the femoral component. In 40 patients the cavity was most commonly positioned superolateral to the femoral neck extending into the gluteal compartment and deep to the fascia lata overlying the greater trochanter (Fig. 6b). In one it

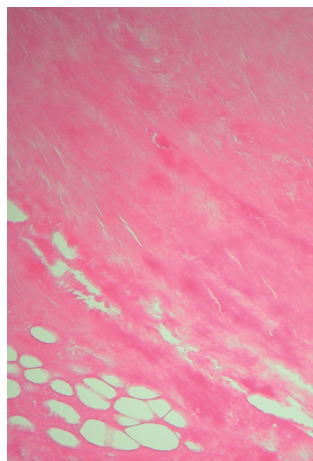


Fig. 7a

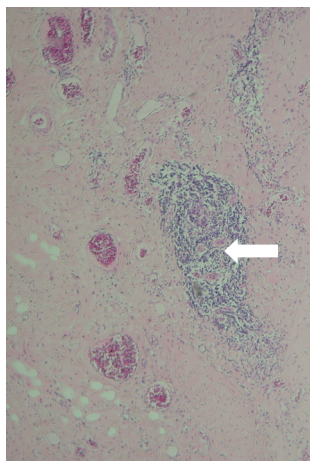


Fig. 7b

Photomicrographs showing a) bland necrosis with recognisable adipose tissue (bottom left) and b) perivascular lymphoid infiltrate (arrowed) at the interface between heterotopic tissue on the right and viable tissue on the left (haematoxylin and eosin  $\times 40$ ).

was positioned anteroinferiorly, replacing the proximal vastus intermedius. In five patients the peri-prosthetic cavity extended through the fascia lata into the subcutaneous fat. In four the primary cavity extended into the quadriceps compartment. Other peri-prosthetic fluid-containing structures included a seroma, a psoas bursa and a sinus track.

In these 45 hips atrophy of the gluteus medius ( $n = 22$ ), gluteus minimus ( $n = 20$ ) and at least one of the short external rotators ( $n = 21$ ) were common findings, with avulsion of the tendons of gluteus medius ( $n = 11$ ), gluteus minimus ( $n = 12$ ), and at least one of the short external rotators ( $n = 5$ ).

Bone-marrow oedema in the proximal femur was demonstrated in the greater trochanter in 22, the lesser trochanter in 21, the anterior inter-trochanteric region in one and posteriorly in five hips. In two patients a fracture of the medial calcar and one of the greater trochanter was seen.

**Histological findings.** Peri-prosthetic tissue was available for histological examination from 76 hips. Of these, 62 hips provided specimens with an appearance 'typical' of MoM necrotic reaction, of which 57 had an inner surface layer of acellular eosinophilic amorphous fibrinoid material. In 24 only necrotic tissue, including bone, was seen (Fig. 7a) within which original structures such as blood vessels and adipose tissue could be identified. The remaining 38 contained necrotic tissue and viable peripheral tissue. This was fibrotic, and in 35 of these there was a perivascular and diffuse lymphoid infiltrate of variable density (Fig. 7b). Histocytes were present and in 19 of the 38 hips there was an ill-defined granulomatous reaction at the interface of the necrotic and viable tissue. Four specimens had an appreciable eosinophilic component. Plasma cells were sparse and neutrophils were not a feature except in those hips with possible infection. Seven specimens had visible metal particles in the necrotic tissue while 16 had microscopic metal particles either free or in macrophages in the fibrinous sur-

face layer. In three, probable metal particles were present in peripherally-situated macrophages. The non-typical samples occurred early in the series when the specimens were small and/or superficial. Of the 76 specimens, eight showed additional active inflammatory changes which were suspicious of infection, five in otherwise typical MoM cases and three in non-typical cases.

One of the patients who had been revised to a metal-on-polyethylene implant went on to a further revision for recurrent dislocation. Histological analysis of the further retrieved tissue showed that there was an ongoing reaction to metal debris typical of this series, despite removal of the MoM implant.

## Discussion

Our study has shown failure of the Ultima MoM THR system in which the predominant feature was pain in the presence of normal radiographs. Most reports of failure of MoM implants describe aseptic loosening<sup>26,28,29</sup> with histological evidence of hypersensitivity.

The bearing surfaces were macroscopically pristine, with corrosion confined to the area of the stem in contact with the acrylic cement. Massive collections of fluid under pressure appeared to be the cause of the pain, since it was relieved temporarily on aspiration of the fluid. We attribute the high rate of late dislocations to avulsion of the gluteal and short external rotator tendons. No obvious abnormality was present on the plain radiographs, although careful review may suggest soft-tissue changes consistent with a large fluid collection, since there was no reaction at the cement-bone interface, and no demineralisation was evident even when the bone was dead. However, the osteonecrosis did lead to peri-prosthetic fractures as one of the mechanisms of failure. Jacobs et al<sup>40</sup> reported a comparative multicentre study using an identical acetabular component and bearing couple, but with a titanium alloy cementless S-ROM femoral component (Johnson and Johnson, Raynham, Massachusetts) with a CoCrMo 28 mm head comparing a ZTT II acetabular component (Johnson and Johnson) with a polyethylene liner and also the S-ROM femoral component. At follow-up at three to six years one hip of 95 in the MoM group had loosened at four years, with one case of dislocation and 12 cases of trochanteric bursitis. Of the 76 hips in the polyethylene-metal articulation group, none had loosened or dislocated, and three had trochanteric bursitis. These results are notably different from our findings.

Massive metallosis has been reported in one patient with a ceramic-on-CoCr bearing with a hydroxyapatite-coated titanium femoral component,<sup>40</sup> who presented with pain and an expanding mass in the proximal thigh. The CT showed a mass 14 cm in size involving the acetabulum, gluteus medius and maximus and extending into the thigh, with the histological examination revealing necrosis with massive metallosis. A further two patients with groin cysts and normal plain radiographs have also recently been



reported.<sup>42</sup> Histological examination showed diffuse or sometimes nodular dense lymphocytic infiltration, mostly around small venous blood vessels with only rare plasma cells, in between many macrophages with phagocytosed metal particles. In our series, the soft-tissue response produced characteristic changes on metal artefact reduction sequence MRI sequences and on histological examination. MRI demonstrated fluid collections of variable size with an irregular wall, generally extending into the gluteal compartment, with atrophy of gluteus medius and minimus being common. Oedema of muscle and bone marrow was noted. Histological examination showed extensive tissue necrosis and a dense perivascular lymphocytic infiltrate. A network of fibrin surrounded 'ghosts' of dead cellular material, with lymphocytic proliferation around blood vessels in the pseudocapsule of the hip. These were similar findings to those described by Davies et al.<sup>24</sup> It seems that this response is directly due to the massive ionic load. Similar changes have been reported in failed MoM hip resurfacings.<sup>43</sup>

High metal loss could be due to poor surgical technique. However, the cementing was satisfactory in most of the failed implants. Abnormal inclination of the acetabular component has been proposed as a cause.<sup>12,13,44,45</sup> In our revision series the acetabular inclination in 62 of the 82 hips with available radiographs was within the safe zone proposed by Langton et al<sup>46</sup> for metal-on-metal polyethylene THR. However, we accept that we have not measured the anteversion angle. Without knowing the position of the acetabular component for the whole series, it is not possible to state whether this correlated with failure. Brodner et al<sup>36</sup> showed no correlation between the inclination of the acetabular component and serum cobalt levels. Neither was chronic renal failure associated with a poor outcome, although the authors noted that their study was underpowered.<sup>36</sup>

MoM bearings are known to shed metal ions over a very long period. Sauvé et al<sup>47</sup> showed that patients with well-fixed Ring MoM prostheses had five times the reference serum levels for cobalt and three times for chromium. However, a long-term study of an original McKee MoM bearing did not show necrosis in bone or soft tissue in an asymptomatic patient.<sup>48</sup> Neither were cobalt or chromium ions found in the soft tissues.

The problem in our series may be one of galvanic corrosion caused by the combination of the titanium alloy shell and the CoCr alloy stem. A titanium stem with a CoCr alloy bearing has been shown to lead to corrosion and failure.<sup>49</sup> Scratching of the titanium surface worsens this effect.<sup>50</sup> This causes a capacitive impedance of nearly zero, favouring an electrochemical dissolution process at the surface of the implant. Wolner et al<sup>50</sup> noted in the laboratory that the passivation, repassivation and corrosion of a MoM slide in combination with a partly damaged titanium alloy surface are quite complex. Between the CoCrMo alloy and the titanium alloy, an electrical potential can be established,

the value of which depends on the electrolyte composition and the temperature which results in a dissolution process of components of the Ti-A1-Nb alloy. However, a cathodic partial reaction at cathodic active areas in the CoCrMo alloy can cause anodic dissolution reactions on anodic active parts at the tribologically treated MoM slide. This results in enhanced cobalt and chromium ion concentrations in solution.


However, other units which have used this implant combination in approximately 400 THRs or the same acetabular component and modular head with a different design of CoCr stem have not recorded this problem. Further work is underway to attempt to replicate these unusual corrosion results *in vitro* by appropriate laboratory studies. This should assist in furthering our understanding of the facts leading to these effects when implant systems are used within the human body in conditions of dynamic loading. We are also uncertain of the significance of the high-carbon CoCrMo alloy insert articulating with the low-carbon CoCrMo alloy head.<sup>51</sup>

An idiosyncratic immune response may be another explanation. The mechanism of the immune reaction seen in these patients is unknown, but it does fall within a stereotype which has been reported for MoM THRs.<sup>24,27</sup> It has been suggested that there may be a subtype of orthopaedic patients which is particularly prone to mounting an immune reaction against the metals in the implant. This may be analogous to the contact sensitivity in the skin which is seen in some patients after exposure to the same metals.<sup>23</sup>


It should be noted that we have not independently tested the bearing surface of the retrievals, but DePuy International has performed tests and reported that no significant wear could be found in the retrieved heads and cups. Unlike other reports of metallosis in MoM bearings where the loss has been at the bearing surface, in this implant configuration the metal loss has been principally from the femoral component.

We have reported a high revision rate of 13.8% (95% CI 9.3 to 19.5) at five years for the Ultima hybrid THR using an uncemented MoM acetabular component and cemented CoCr tapered polished femoral component. The failures were due to necrosis of peri-implant bone and soft tissues from massive metallosis, the cause of which was uncertain.

### Supplementary material

 A further opinion by Ms L. Biant is available with the electronic version of this article on our website at [www.jbjs.org.uk](http://www.jbjs.org.uk)

### Listen live

 Listen to the abstract of this article at [www.jbjs.org.uk/interactive/audio](http://www.jbjs.org.uk/interactive/audio)

We gratefully acknowledge the contribution and participation of the members of the Norwich Metal-on-Metal Study group as follows: Dr J. Cahir, Dr T. J. Marshall, Miss S. Neda, Dr S. Hammouche, Dr M. Elsaid from the Norfolk & Norwich University Hospital; Dr C. L. Peters from the Avon Orthopaedic Centre, Bristol; Professor I. Clark from the University of East Anglia, Norwich; Professor

A. Neville from the University of Leeds; and M. Borroff and R. Farrar from DePuy International.

The author or one or more of the authors have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article. In addition, benefits have been or will be directed to a research fund, foundation, educational institution, or other non-profit organisation with which one or more of the authors are associated.

## References

- McKee GK, Watson-Farrar J. Replacement of arthritic hips by the McKee-Farrar prosthesis. *J Bone Joint Surg [Br]* 1966;48-B:245-59.
- August AC, Aldam CH, Pynsent PB. The McKee-Farrar hip arthroplasty: a long-term study. *J Bone Joint Surg [Br]* 1986;68-B:520-7.
- Charnley J. The long-term results of low-friction arthroplasty of the hip performed as "a" primary intervention. *J Bone Joint Surg [Br]* 1972;54-B:61-76.
- Charnley J. *Low friction arthroplasty of the hip*. Berlin: Springer-Verlag, 1979:3.
- Schmalzried TP, Peters PC, Maurer BT, Bragdon CR, Harris WH. Long-duration metal-on-metal total hip arthroplasties with low wear of the articulating surfaces. *J Arthroplasty* 1996;11:322-31.
- Campbell P, Urban RM, Catelas I, Skipor AK, Schmalzried TP. Autopsy analysis thirty years after metal-on-metal total hip replacement: a case report. *J Bone Joint Surg [Am]* 2003;85-A:2218-22.
- Dunstan E, Sanghrajka AP, Tilley S, et al. Metal ion levels after metal-on-metal proximal femoral replacements: a 30-year follow-up. *J Bone Joint Surg [Br]* 2005;87-B:628-31.
- Saito R, Ryu J, Watanabe M, Ishii T, Saigo K. Midterm results of Metasul metal-on-metal total hip arthroplasty. *J Arthroplasty* 2006;21:1105-10.
- Müller ME. The benefits of metal-on-metal total hip replacements. *Clin Orthop* 1995;311:54-9.
- Ladon D, Doherty A, Newson R, et al. Changes in metal levels and chromosome aberrations in the peripheral blood of patients after metal-on-metal hip arthroplasty. *J Arthroplasty* 2004;19(Suppl 3):78-83.
- Rieker CB, Schön R, Kötting P. Development and validation of a second-generation metal-on-metal bearing: laboratory studies and analysis of retrievals. *J Arthroplasty* 2004;19(Suppl 3):5-11.
- Brodner W, Gröble A, Jankovsky R, et al. Cup inclination and serum concentration of cobalt and chromium after metal-on-metal total hip arthroplasty. *J Arthroplasty* 2004;19(Suppl 3):66-70.
- Vendittoli PA, Mottard S, Roy AG, Dupont C, Lavigne M. Chromium and cobalt ion release following the Durom high carbon content, forged metal-on-metal surface replacement of the hip. *J Bone Joint Surg [Br]* 2007;89-B:441-8.
- Grammatopoulos G, Pandit H, Glyn-Jones S, et al. Optimal acetabular orientation for hip resurfacing. *J Bone Joint Surg [Br]* 2010;92-B:1072-8.
- de Haan R, Pattyn C, Gill HS, et al. Correlation between inclination of the acetabular component and metal ion levels in metal-on-metal hip resurfacing replacement. *J Bone Joint Surg [Br]* 2008;90-B:1291-7.
- Ollivere B, Darrah C, Barker T, Nolan J, Porteous MJ. Early clinical failure of the Birmingham metal-on-metal hip resurfacing is associated with metallosis and soft-tissue necrosis. *J Bone Joint Surg [Br]* 2009;91-B:1025-30.
- Keel JB, Kuster MS. Massive wear of an incompatible metal-on-metal articulation in total hip arthroplasty. *J Arthroplasty* 2004;19:638-42.
- Zahiri CA, Schmalzried TP, Ebrahmdadeh E, et al. Lesson learned from loosening of the McKee-Farrar metal-on-metal total hip replacement. *J Arthroplasty* 1999;14:326-32.
- Howie DW, McCalden DW, McCalden RW, Nawana NS, et al. The long-term wear of retrieved McKee-Farrar metal-on-metal total hip prostheses. *J Arthroplasty* 2005;20:350-7.
- Schmalzried TP, Shepherd EF, Dorey FJ, et al. The John Charnley Award: wear is a function of use, not time. *Clin Orthop* 2000;381:36-46.
- Heisel C, Silva M, Skipor AK, Jacobs JJ, Schmalzried TP. The relationship between activity and ions in patients with metal-on-metal bearing hip prostheses. *J Bone Joint Surg [Am]* 2005;87-A:781-7.
- Khan M, Takahashi T, Kuiper JH, et al. Current in vivo wear of metal-on-metal bearings assessed by exercise-related rise in plasma cobalt level. *J Orthop Res* 2006;24:2029-35.
- Gawkrödger DJ. Metal sensitivities and orthopaedic implants revisited: the potential for metal allergy with the new metal-on-metal joint prostheses. *Br J Dermatol* 2003;148:1089-93.
- Davies AP, Willert HG, Campbell PA, Learmonth ID, Case CP. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg [Am]* 2005;87-A:18-27.
- Hart AJ, Hester T, Sinclair K, et al. The association between metal ions from hip resurfacing and reduced T-cell counts. *J Bone Joint Surg [Br]* 2006;88-B:449-54.
- Park YS, Moon YW, Lim SJ, et al. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg [Am]* 2005;87-A:1515-21.
- Willert HG, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints: a clinical and histomorphological study. *J Bone Joint Surg [Am]* 2005;87-A:28-36.
- Korovessis P, Petsinis G, Repanti M, Repantis T. Metallosis after contemporary metal-on-metal total hip arthroplasty: five to nine-year follow-up. *J Bone Joint Surg [Am]* 2006;88-A:1183-91.
- Milosevi I, Trebse R, Kovac S, Cör A, Pisot V. Survivorship and retrieval analysis of Sikomet metal-on-metal total hip replacements at a mean of seven years. *J Bone Joint Surg [Am]* 2006;88-A:1173-82.
- Harris WH, McCarthy JC Jr, O'Neill DA. Femoral component loosening using contemporary techniques of femoral cement fixation. *J Bone Joint Surg [Am]* 1982;64-A:1063-7.
- Willert HG, Semlitsch M. Reactions of the articular capsule to wear products of artificial joint prostheses. *J Biomed Mater Res* 1977;11:157-64.
- Sabokbar A, Kudo O, Athanasou NA. Two distinct cellular mechanisms of osteoclast formation and bone resorption in periprosthetic osteolysis. *J Orthop Res* 2003;21:73-80.
- Toms AP, Marshall TJ, Cahir J, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008;63:49-58.
- No authors listed. Medical Devices Directive 93/42/EEC. <http://www.mhra.gov.uk/Howweregulate/Devices/MedicalDevicesDirective/index.htm> (date last accessed 21 September 2010).
- Carrington NC, Sierra RJ, Gie GA, et al. The Exeter Universal cemented femoral component at 15 to 17 years: an update on the first 325 hips. *J Bone Joint Surg [Br]* 2009;91-B:730-7.
- Brodner W, Bitzan P, Meisinger V, et al. Serum cobalt levels after metal-on-metal total hip arthroplasty. *J Bone Joint Surg [Am]* 2003;85-A:2168-73.
- Johnston RC, Fitzgerald RH Jr, Harris WH, et al. Clinical and radiographic evaluation of total hip replacement: a standard system of terminology for reporting results. *J Bone Joint Surg [Am]* 1990;72-A:161-8.
- Barrack RL, Mulroy RD, JC, Harris WH. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty: a 12-year radiographic review. *J Bone Joint Surg [Br]* 1992;74-B:385-9.
- Duncan CP, Masri BA. Fractures of the femur after hip replacement. *Instr Course Lect* 1995;44:293-304.
- Jacobs M, Gorab R, Mattingly D, Trick L, Southworth C. Three- to six-year results with the Ultima metal-on-metal hip articulation for primary total hip arthroplasty. *J Arthroplasty* 2004;19(Suppl 2):48-53.
- Valenti JR, Del Rio J, Amillo S. Catastrophic wear in a metal-on-ceramic total hip arthroplasty. *J Arthroplasty* 2007;22:920-2.
- Gruber FW, Böck A, Trattig S, Lintner F, Ritschl P. Cystic lesion of the groin due to metallosis: a rare long-term complication of metal-on-metal total hip arthroplasty. *J Arthroplasty* 2007;22:923-7.
- Boardman DR, Middleton FR, Kavanagh TG. A benign psoas mass following metal-on-metal resurfacing of the hip. *J Bone Joint Surg [Br]* 2006;88-B:402-4.
- Hart AJ, Sabah S, Henckel J, et al. The painful metal-on-metal hip resurfacing. *J Bone Joint Surg [Br]* 2009;91-B:738-44.
- Langton DJ, Sprowson AP, Joyce TJ, et al. Blood metal ion concentrations after hip resurfacing arthroplasty: a comparative study of articular surface replacement and Birmingham Hip Resurfacing arthroplasties. *J Bone Joint Surg [Br]* 2009;91-B:1287-95.
- Langton DJ, Jameson SS, Joyce TJ, et al. Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: a consequence of excess wear. *J Bone Joint Surg [Br]* 2010;92-B:38-46.
- Sauvé P, Mountney J, Khan T, et al. Metal ion levels after metal-on-metal Ring total hip replacement: a 30-year follow-up study. *J Bone Joint Surg [Br]* 2007;89-B:586-90.
- Clarke MT, Darrah C, Stewart T, et al. Long-term clinical, radiological and histopathological follow-up of a well-fixed McKee-Farrar metal-on-metal total hip arthroplasty. *J Arthroplasty* 2005;20:542-6.
- Böhler M, Kanz F, Schwarz B, et al. Adverse tissue reactions to wear particles from Co-alloy articulations, increased by alumina-blasting particle contamination from cementless Ti-based total hip implants: a report of seven revisions with early failure. *J Bone Joint Surg [Br]* 2002;84-B:128-36.
- Wolner C, Nauer GE, Trummer J, Putz V, Tschegg S. Possible reasons for the unexpected bad biocompatibility of metal-on-metal hip implants. *Mater Sci Eng C* 2006;26:34-40.
- Tipper JL, Firkins PJ, Ingham E, et al. Quantitative analysis of the wear and wear debris from low and high carbon content cobalt chrome alloys used in metal on metal total hip replacements. *J Mater Sci Mater Med* 1999;10:353-62.



## Case Report

# The Natural History of Metallosis From Catastrophic Failure of a Polyethylene Liner in a Total Hip

Riaz J.K. Khan, FRCS (Tr&Orth), FRACS (Orth),\*  
James Wimhurst, MChir, FRCS (Tr&Orth),† Sara Foroughi, MBBS,§ and  
Andoni Toms, FRCS, FRCR‡

---

**Abstract:** We report on a case of metallosis initially presumed to be heterotopic ossification based on radiologic findings. A 68-year-old man with a total hip arthroplasty experienced failure of the polyethylene liner, resulting in articulation of the ceramic head with the titanium acetabular shell. During revision surgery, extensive metallic debris was evident macroscopically throughout the periprosthetic tissue and was confirmed histologically to be metallosis. **Keywords:** heterotopic ossification, metallosis, total hip arthroplasty, osteolysis, revision hip arthroplasty. Crown Copyright © 2009 Published by Elsevier Inc. All rights reserved.

---

Heterotopic ossification (HO) is well recognized after hip surgery, with typical appearances on radiographic examination. We describe a case of catastrophic failure of a polyethylene liner, resulting in articulation of the ceramic head with the titanium acetabular shell, in which preoperative radiographs and computerized tomography (CT) were thought to be typical of HO. However, at operation, extensive metallosis was encountered instead.

## Case History

A 68-year-old man presented with a 2-year history of increasing pain and squeaking from his right total hip arthroplasty. He had an antalgic gait, and his range of motion was 0° to 90° of flexion and 15° of internal and external rotation.

Fourteen years earlier, he had undergone a CLS Spotorno (Centerpulse, Sulzer/Winterthur, Switzerland) ceramic on polyethylene uncemented primary total hip arthroplasty through a posterior approach. The acetabular component failed a year later because of aseptic loosening and was revised to a Norwich Porous Coated Cup (DePuy, Leeds, UK) with a polyethylene liner. He remained symptom free for 11 years after the revision surgery. The patient then began to experience symptoms of mild hip discomfort. Plain films taken at this time by the family doctor demonstrated features consistent with Brooker grade 2 HO (Fig. 1A). There was a delay in the referral process, and it was another 14 months before he was seen by an orthopedic surgeon. At this time, the patient was experiencing moderate pain and occasional squeaking. Plain x-rays demonstrated asymmetric polyethylene wear, and the

---

*From the \*Department of Pathology and Surgery, University of Western Australia, Perth, Western Australia; †Department of Trauma and Orthopaedics, Norfolk and Norwich University Hospital, Norwich, England; ‡Department of Radiology, Norfolk and Norwich University Hospital, Norwich, England; and §Department of Trauma and Orthopaedics, Sir Charles Gairdner Hospital, Perth, Western Australia.*

Submitted October 11, 2007; accepted September 3, 2008.

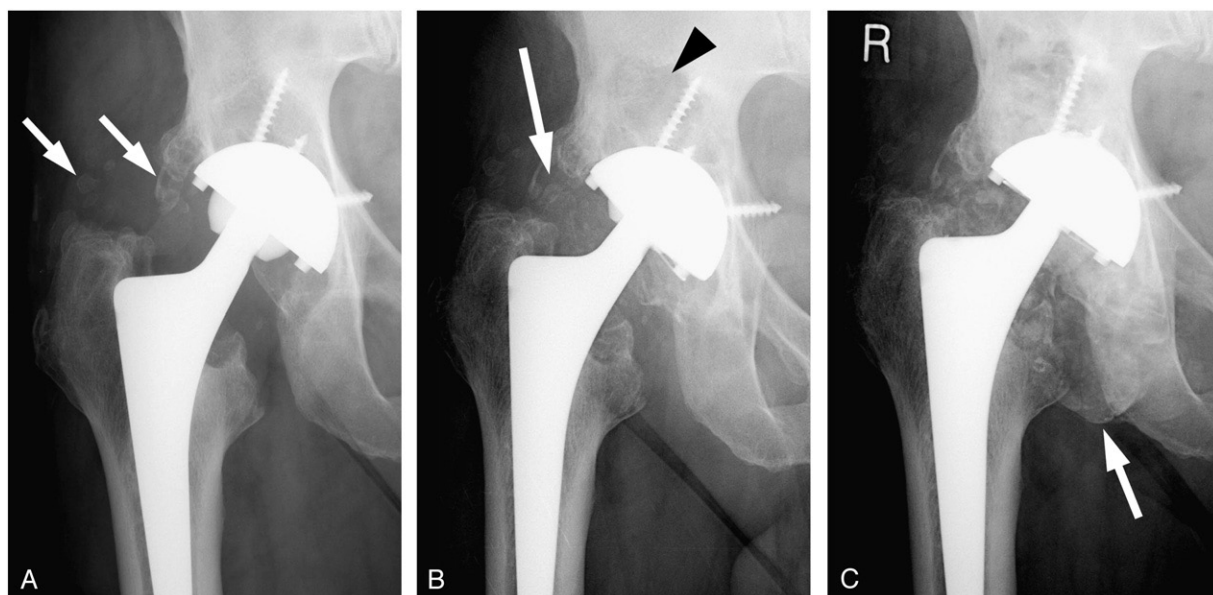
No benefits or funds were received in support of the study.

Reprint requests: Riaz J.K. Khan, FRCS (Tr&Orth), FRACS (Orth), Department of Trauma and Orthopaedics, Sir Charles Gairdner Hospital, Hospital Ave, Nedlands 6009, Perth, Western Australia.

Crown Copyright © 2009 Published by Elsevier Inc. All rights reserved.

0883-5403/08/2407-0028\$36.00/0

doi:10.1016/j.arth.2008.09.002

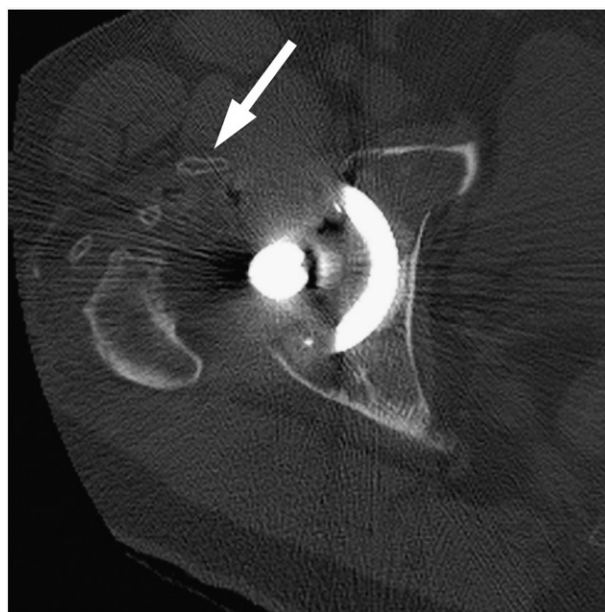


**Fig. 1.** A, Anteroposterior (AP) radiograph of the right total hip arthroplasty 10 years after surgery demonstrating well corticated HO, with central lucencies, projected over the superolateral soft tissues of the hip (arrows). B, AP radiograph of the right THR at presentation, 11 years after surgery, demonstrates that the head of the femoral component has migrated superomedially, indicating that the polyethylene liner has fractured. There are foci of HO (arrow) and a supra-acetabular lytic lesion (arrowhead). C, Five months after presentation, there is new opacification inferomedial to the neck of the femoral prosthesis (arrow).

patient was listed for revision surgery. The patient declined early surgery as he had planned a holiday. He presented to the preadmission clinic 5 months later with clinical deterioration and squeaking. Revision surgery was undertaken within 1 week.

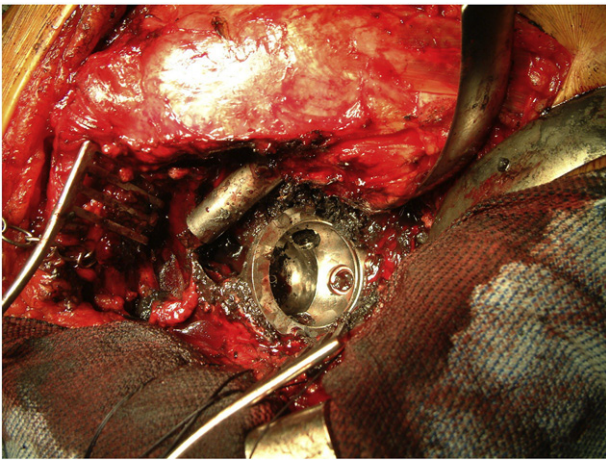
Plain radiographs at presentation (Fig. 1B) demonstrated marked polyethylene wear with an eccentrically placed head within the cup. A lytic area was noted in the supra-acetabular ilium. Heterotopic ossification, characterized by calcific opacities comprising a well-defined cortex and medulla, were projecting superolateral to the neck of the prosthesis. This HO had progressed when compared with a previous radiograph obtained 1 year previously to Brooker grade 3. A lytic area was noted in the supra-acetabular ilium. Five months after presentation, extensive opacification of calcific attenuation was present inferomedial to the neck of the prosthesis, with apparent ankylosis of the hip (Fig. 1C). A multidetector CT of the pelvis was obtained (120 kV, 350 mA, matrix 512 × 512, slice thickness 2.5 mm, General Electric Lightspeed, General Electric Healthcare, Chalfont St Giles, Bucks, UK). These demonstrated areas of amorphous increased attenuation within the pseudocapsule of the joint extending into areas of lysis within the supra-acetabular ilium and ischium. These areas were distinct from the clearly defined corticomedullary differentiation of HO (Fig. 2).

The patient was taken to the operating theater for revision of the hip arthroplasty. Intraoperative findings were of gross metallic debris throughout the soft tissues (Fig. 3). The pseudocapsule was thickened and black from extensive metallosis, and produced sparks



**Fig. 2.** Axial CT through the right THR, performed shortly after presentation, at the cranial level of the head of the femoral prosthesis demonstrating HO indicated by discrete cortical and medullary differentiation (arrow).





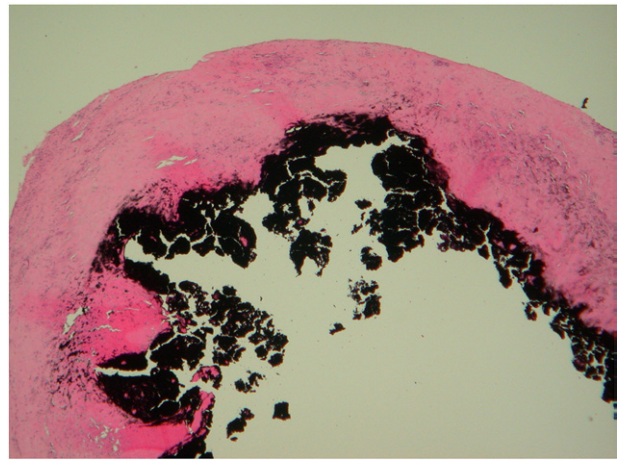
**Fig. 3.** Intraoperative view of the socket with head and liner removed. Black staining of tissues evident.

on electrocautery. The ceramic head had penetrated the polyethylene liner, eroded a hole through the metal shell, and worn away the heads of 2 screws securing the cup (Fig. 4). There was a large uncontained defect superiorly and contained defects in the ischium and pubis. The superior defect required mesh augmentation and impaction bone grafting. The femoral component was found to be stable and therefore retained. There was no HO present.

Microscopic analysis of the pigmented soft tissues demonstrated dense, hyalinized, fibrous tissue with large numbers of macrophages. There was heavy pigmentation by metallic particles (Fig. 5). There was no evidence of HO. Specimens of cortical bone from the acetabulum also showed black metallic



**Fig. 4.** Fractured polyethylene liner removed from socket.



**Fig. 5.** Microscopic analysis of the soft tissues demonstrating dense, hyalinized, fibrous tissue with large numbers of macrophages and heavy pigmentation by metallic particles.

particle-laden macrophages. There was no evidence of inflammation or infection.

## Discussion

Heterotopic ossification is the pathologic formation of bone in soft tissue and is a well-recognized complication of hip surgery. A review by Neal et al [1] found an incidence of HO after total hip replacement to be as high as 43%, with severe HO found in 9% of cases. The most common symptoms are of pain from abutting bone and decreased range of movement [2]. Predisposing factors include male sex, ankylosing spondylitis, severe osteoarthritis, posttraumatic arthritis, previous HO of the ipsilateral or contralateral hip [3-5], and traumatic brain or cord injury [6,7]. Heterotopic ossification is detectable on x-ray after 6 to 12 weeks [4]. The severity of HO is graded 1 to 4 by Brooker et al [8]. Prophylactic treatment includes radiation therapy and oral indomethacin [9]. Curative treatment involves surgical excision [2,10], but recurrence occurs in 10% to 20% of cases [2].

Metallosis is the infiltration of metallic wear debris in the periprosthetic tissues [11]. It occurs after catastrophic polyethylene liner failure due to dissociation [11], fracture or wear [12], or abnormal abrasion of the femoral head [13]. By virtue of the fact that it occurs after failure of componentry, it typically occurs after some years. The effect of metallosis is that of metallic staining of the tissues. Macroscopically, the hip pseudocapsule contains a black-stained synovium [13]. Microscopically, there is local necrosis with histiocytic and granulomatous reaction around the deposits of metallic particles [13,14]. In cases of severe metallosis, there is radiographic opacification of the periprosthetic soft tissue planes by the deposition of

metallic debris, known as the “bubble sign” [15], as well as osteolytic lesions [11]. Metallic debris is seen on CT [16]. With extensive metallosis, excision of the pseudomembrane using electrocautery can produce sparks [15], which are potentially hazardous.

The plain radiographic features of metallosis can be difficult to distinguish from HO, particularly when both processes are present. Layering of metallic debris around the pseudocapsule of the joint can mimic the peripheral cortical calcific opacity of HO when projected onto a 2-dimensional image. However, CT, with multiplanar reformats, should allow the differentiation of HO, with its clearly demarcated cortex and medulla, from the diffuse amorphous increased attenuation caused by metallosis. This distinction may be more difficult in the presence of fragmented cement. In addition, late and rapid radiographic deterioration of HO is uncommon, whereas late and rapid deterioration of metallosis is not.

In this case, the patient had symptoms for 2 years before definitive revision and was operated on within 6 months of presentation to an orthopedic surgeon. This delay allowed for progression of metallosis and subsequently complicated the revision surgery. This case demonstrates that once clinical symptoms such as squeaking occur in a total hip arthroplasty, any delay in revision can complicate surgery.

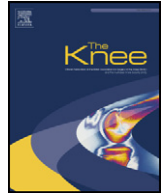
During the follow-up of hip arthroplasties, clinical evaluation is as important as radiography, and this is a pertinent example. The fact that this patient had low-grade HO present before the development of the metallosis further complicated the diagnosis, particularly because x-rays from the first arthroplasty and early revision were unavailable. However, the presence of a good range of movement despite supposedly high grade HO should have raised clinical suspicion that this was metallosis sooner. This case demonstrates that in the apparent presence of HO on radiographs of failed total hip arthroplasties, where liner failure has occurred, extensive metallosis should be considered as the primary diagnosis. Computed tomography may allow the diagnosis of metallosis to be made preoperatively.

### Acknowledgment

We are grateful to Prof Richard Ball of the Norfolk and Norwich University Hospital, Norwich, England, for his histologic advice.

### References

1. Neal B, Gray H, MacMahon S, et al. Incidence of heterotopic bone formation after major hip surgery. *ANZ J Surg* 2002;72:808.
2. Garland DE. A clinical perspective on common forms of acquired heterotopic ossification. *Clin Orthop Relat Res* 1991;13.
3. Spry NA, Dally MJ, Benjamin B, et al. Heterotopic bone affecting the hip joint is preventable in high risk patients by post-operative radiation. *Australas Radiol* 1995;39:379.
4. Ahrengart L. Periarticular heterotopic ossification after total hip arthroplasty. *Clin Orthop Relat Res* 1991;49.
5. Ahrengart L, Lindgren U. Heterotopic bone after hip arthroplasty. Defining the patient at risk. *Clin Orthop Relat Res* 1993;153.
6. Warren SB. Heterotopic ossification after total hip replacement. *Orthop Rev* 1990;109:603.
7. Pittenger DE. Heterotopic ossification. *Orthop Rev* 1991;100:33.
8. Brooker AF, Bowerman JW, Robinson RA, et al. Ectopic ossification following total hip replacement, Incidence and a method of classification. *J Bone Joint Surg* 1973;55A:1629.
9. Board TN, Karva A, Board RE, et al. The prophylaxis and treatment of heterotopic ossification following lower limb arthroplasty. *J Bone Joint Surg* 2007;89:434.
10. Garland DE, Orwin JH. Resection of heterotopic ossification in patients with spinal cord injuries. *Clin Orthop Relat Res* 1989;169.
11. Chang JD, Lee SS, Hur M, et al. Revision total hip arthroplasty in hip joints with metallosis. *J Arthroplasty* 2005;20:568.
12. Berry DJ, Barnes CL, Scott RD, et al. Catastrophic failure of the polyethylene liner of uncemented acetabular components. *J Bone Joint Surg Br* 1994;76-B:575.
13. Matsuda Y, Yamamuro T, Kasai R, et al. Severe metallosis due to abnormal abrasion of the femoral head in a dual bearing hip prosthesis. *J Arthroplasty* 1992;7(Suppl):439.
14. Keel JB, Kuster MS. Massive wear of an incompatible metal-on-metal articulation in total hip arthroplasty. *J Arthroplasty* 2004;19:638.
15. Su EP, Callander PW, Salvati EA. The bubble sign, a new radiographic sign in total hip arthroplasty. *J Arthroplasty* 2003;1:110.
16. Park JS, Ryu KN, Hong HP, et al. Focal osteolysis in total hip replacement: CT findings. *Skeletal Radiol* 2004;33:632.



# CT metal artefact reduction of total knee prostheses using angled gantry multiplanar reformation

Mark Lewis <sup>a,\*</sup>, Andoni P. Toms <sup>a</sup>, Karen Reid <sup>b</sup>, William Bugg <sup>a</sup>

<sup>a</sup> Radiology Academy, Cotman Centre, Colney Lane, Norwich, Norfolk, NR4 7UB, United Kingdom

<sup>b</sup> Department of Radiology, Norfolk & Norwich University Hospital, Colney Lane, Norwich, Norfolk, NR4 7UY, United Kingdom

## ARTICLE INFO

### Article history:

Received 27 November 2009

Received in revised form 18 February 2010

Accepted 19 February 2010

### Keywords:

CT  
Knee  
Metal  
Artefact  
Reduction

## ABSTRACT

This study was designed to determine whether or not acquiring CT images of total knee prostheses by using an angled gantry and multiplanar reformation can reduce beam hardening artefact. A CT phantom was created with a total knee prosthesis suspended in gelatine with a known attenuation. CT data was acquired with a gantry angled at 0°, 5°, 10° and 15° in both craniocaudal oblique planes. Axial images were then reformatted from these datasets. Two independent observers selected regions of interest to measure the mean and standard deviation (SD) of attenuation in the gelatine for all reformatted axial images. Artefact was measured as SD of the background attenuation and areas under the curve of SD for each gantry angle acquisition were compared. Inter-observer reliability was excellent (ICC = 0.89, CI 0.875–0.908). The most accurate mean attenuation values for tissues around a TKR were obtained with a CT gantry using 10° to 15° anteroposterior to posterosuperior angulation. The smallest area under the curve for SD of attenuation for the whole prosthesis, and the femoral component in isolation, was obtained with a 5° gantry angle in the same direction. The smallest area under the curve for the tibial component in isolation occurred with a gantry angle of 15°. We conclude that acquiring CT data with a gantry angle can reduce metal artefact around a TKR. Optimal overall metal artefact reduction can be achieved with a small angle from anteroposterior to posterosuperior. Further selective artefact reduction around the tibial component can be achieved with larger angles.

Crown Copyright © 2010 Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Metallic artefact can be a significant problem in the assessment of joint replacements using computed tomography (CT). The very high attenuation of most metallic implants can lead to severe streaking artefacts. These are caused by two principal mechanisms. The first problem is beam hardening which is caused by preferential absorption, by the implant, of lower energy photons in the beam which skews the attenuation profile for all pixels along the line of the beam resulting in streaks and dark bands projected in the reconstructed image [1]. This is compounded by the second problem, which is that the mean attenuation of the metal is outside the range expected by the computer resulting in incomplete attenuation profiles [1]. A number of techniques have been described for reducing both metallic artefact and beam hardening artefact. Improved software reconstruction algorithms can include expected attenuation profiles for metallic prostheses [2,3]. Increasing mAs and Kv will increase the number of photons, reducing noise, and narrow the profile of photon energies [4,5]. Increased slice thickness will improve the signal to noise ratio,

but can be associated with increased partial volume artefacts [4]. Increasing the CT scale will improve the appearance of streak artefact even if it does not improve the severity of the artefact [6]. CT imaging of total hip replacement has improved considerably using these adaptations. However in total knee replacement CT is still difficult because of the complex morphology and large volume of metal in the prosthesis. This aim of this study was to determine whether or not metal artefact reduction could be achieved in TKRs using multiplanar reformatted CT acquired with an angled gantry.

## 2. Materials and methods

A phantom was created using a chromium steel alloy TKR (DePuy International Ltd, Leeds UK), including the polyethylene tray. The TKR was suspended in gelatine containing iodinated contrast medium (Omnipaque™ GE Healthcare, Bucks, UK, diluted to 0.4% of its original concentration) with an approximate attenuation of 48 HU (Fig. 1). CT studies were performed using a GE Lightspeed plus (GE Healthcare, Bucks, UK) with the following imaging parameters: slice thickness 1.25 mm, 512 × 512 matrix, and 440 mA and 120 Kv. The whole prosthesis was covered in a single acquisition with the gantry vertical and repeated with the gantry tilted at 5°, 10° and 15° anterosuperior to posteroinferior and 5°, 10° and 15° posterosuperior to anteroposterior

\* Corresponding author. Tel.: +44 1603 286140; fax: +44 1603 286146.

E-mail address: mark.lewis@nnuh.nhs.uk (M. Lewis).



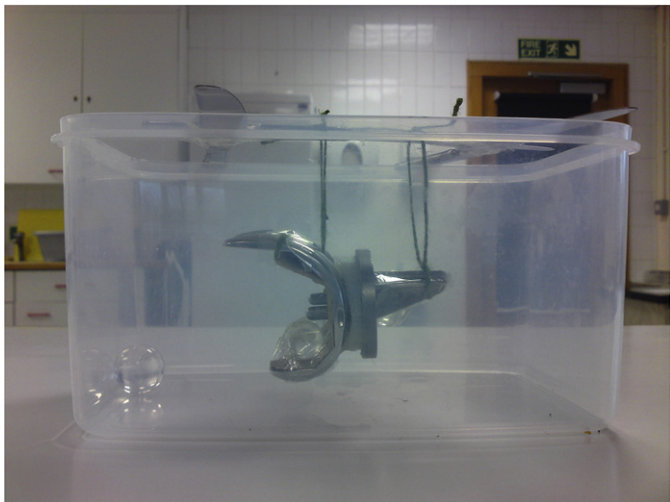


Fig. 1. Photograph of the phantom before addition of the gelatine suspension.

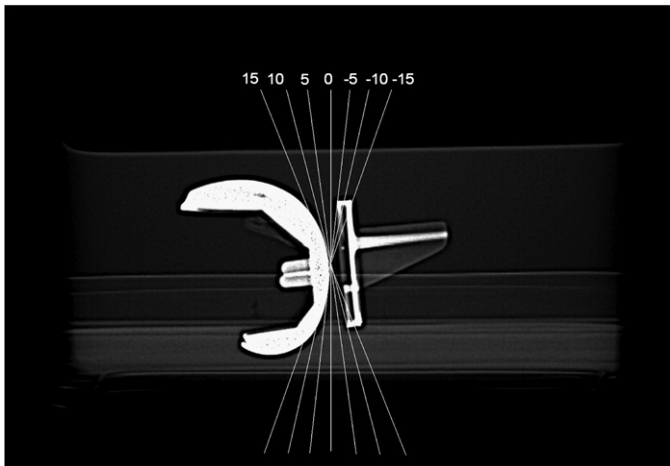


Fig. 2. CT lateral scout tomogram of the phantom with gantry tilt degrees superimposed.

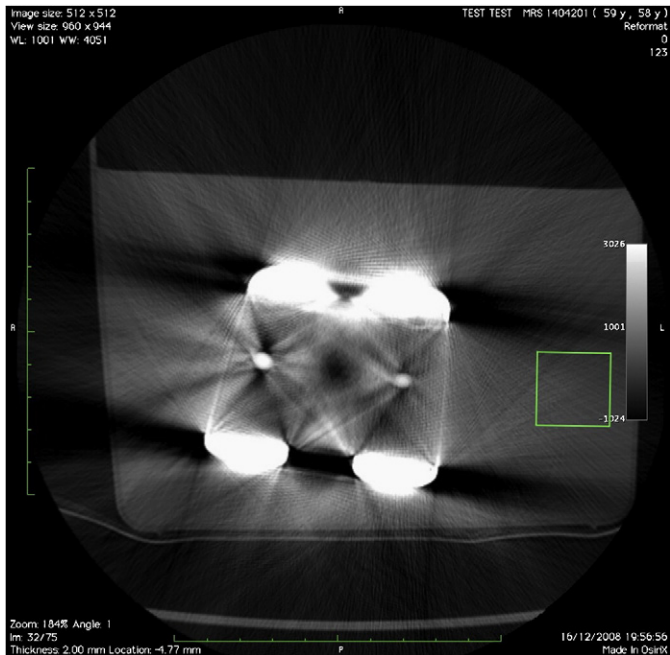


Fig. 3. A screenshot from the DICOM viewer (Osirix) demonstrating the selection of ROI to measure background attenuation.

(Fig. 2). Axial datasets were then reformatted using the preset bone algorithm for each acquisition. Reformatted datasets were then reviewed on a DICOM viewer (Osirix v3.1–32-bit) where two independent observers drew 10 cm<sup>2</sup> regions of interest (ROI) at the same position on the background gelatine and recorded the mean attenuation values and standard deviation (Fig. 3). The reliability of the observers' measurements was calculated using intra-class correlation (SPSS 16.0 for Windows). The mean values, of the two observers, for standard deviation were used as a measure of artefact [8]. The relative amount of artefact, for each gantry angle, was calculated by measuring the area under the curve (trapezoid method) generated by graphs of standard deviation of attenuation from the ROIs (Figs. 4 and 5).

3. Results

The intra-class correlation for the two observers was excellent  $r=0.89$  (95% CI, 0.875–0.908) (Table 1). The mean background attenuation was closest to the true value of 48 HU in those datasets acquired at  $-10$  and  $-15^\circ$  of gantry angle (Table 1). With a gantry angle in the anterosuperior to posteroinferior plane the mean attenuation rose to 59.2 HU at  $+15^\circ$  and 107.37 HU at  $+15^\circ$  for the whole phantom and for the femoral component respectively. For the tibial component the opposite occurred. Although the gantry angle of  $-10^\circ$  gave the attenuation closest to background (37.6 HU) in the opposite direction the mean attenuation fell to 1.9 HU at  $+15^\circ$  (Fig. 4). When the phantom was considered as a whole the minimum standard deviation for attenuation values (least artefact) in the background fat was achieved at a gantry tilt angle of  $-5^\circ$  (SD 64.3 HU) and the maximum (worst artefact) at  $+5^\circ$  (SD 81.7 HU). When the femoral component alone was considered optimal artefact reduction was similarly at  $-5^\circ$  (SD 87.65 HU) but much worse with increasing anteroinferior to posterosuperior gantry angle (SD 122.1 HU at  $-15^\circ$ ). Artefact around the tibial component appears to improve with any gantry angle and any in any direction when compared to directly acquired axial images (SD 50.27 HU) but artefact reduction was at its best with a gantry angle of  $-15^\circ$  (SD 31.21 HU).

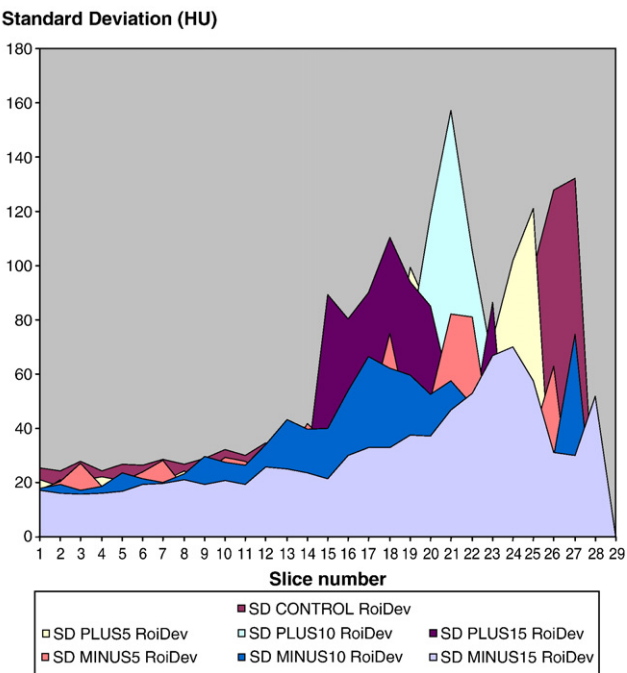


Fig. 4. Graph demonstrating the areas under the curve for measurements of standard deviation in background attenuation adjacent to the tibial component. The smaller the amplitude of the curve the less artefact there is.

**Table 1**

Demonstrating changes in mean attenuation and the standard deviation for various parts of the phantom with varying gantry angles.

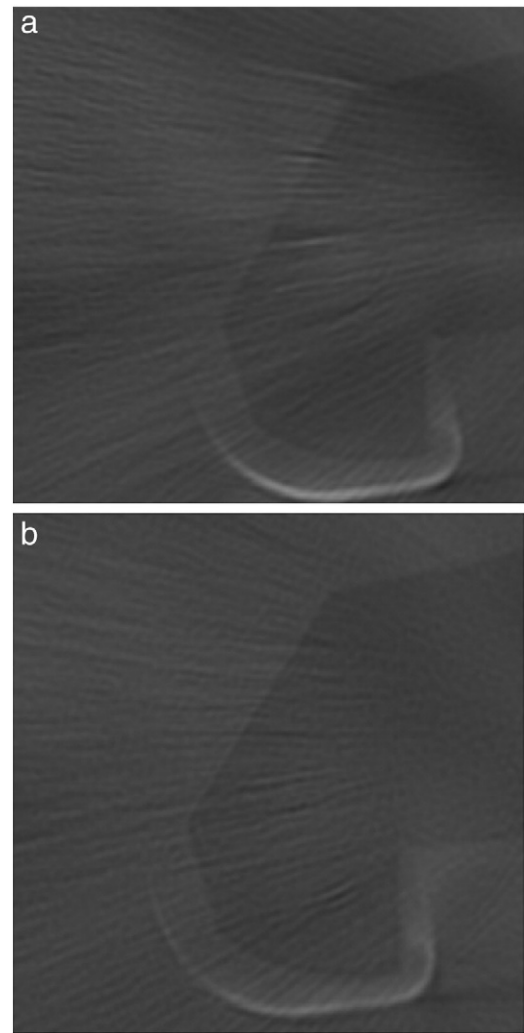
Attenuation (HU)		Gantry angle						
		−15°	−10°	−5°	0°	+5°	+10°	+15°
Whole phantom	Mean	35.0	44.9	80.3	71.0	44.7	55.5	59.2
	SD	81.7	66.9	64.3	75.8	81.7	75.5	74.9
Tibial component	Mean	33.7	37.6	36.5	16.4	10.5	13.3	1.9
	SD	31.2	34.6	37.9	50.3	43.5	40.5	43.7
Femoral component	Mean	33.9	61.1	119.8	116.2	75.3	91.8	107.4
	SD	122.1	90.2	87.7	95.6	115.1	107.5	105.3

#### 4. Discussion

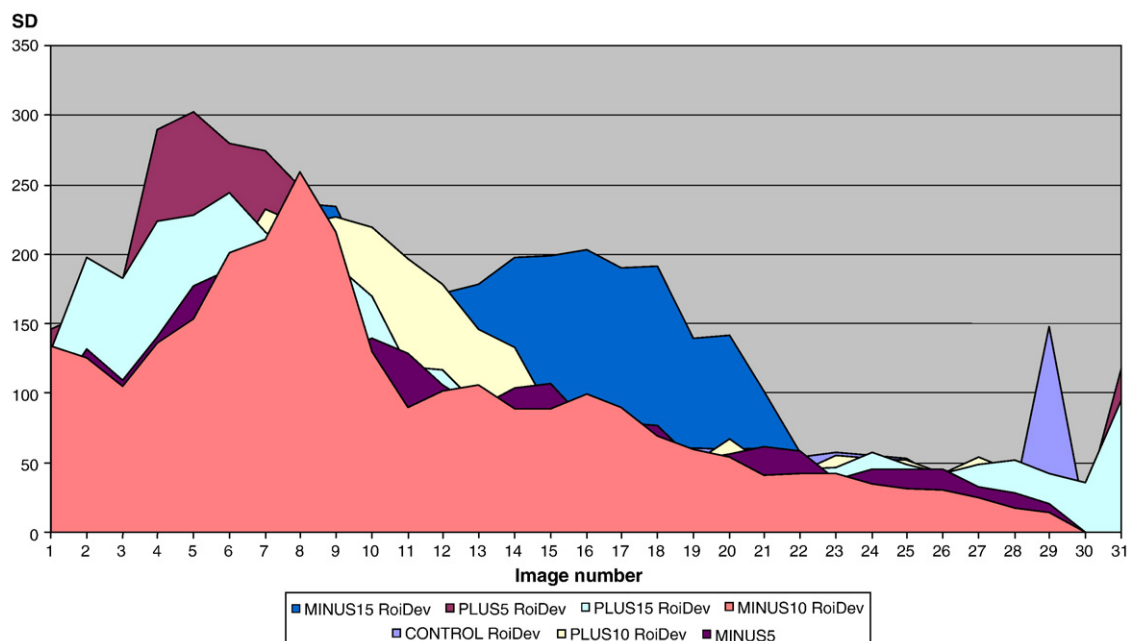
This study suggests that CT of total knee replacements may be improved by adding a gantry angle particularly in the anteroposterior direction and then reconstructing images in conventional orthogonal planes. Measures of artefact in the adjacent tissues can in general be reduced by a 5 to 10° tilt or artefact reduction can be more specifically targeted to the tibial component with an angle of 15°. In reality this is artefact displacement rather than true reduction. The angled gantry spreads the artefact over more of the reconstructed axial slices than if acquired as direct axial images but has the effect of reducing the amount of artefact in the areas of interest. The optimal gantry angle depends on the particular oblique axial plane that crosses the least amount of metal.

Angling the gantry tilt to reduce artefact in the region of interest to the radiologist is by no means a new technique [7] and is used in practice, for instance, when imaging the brain to avoid the dense bone at the base of the skull. However, we were unable to find an instance in the literature describing this technique in relation to joint replacements.

There are some limitations to this study. Demonstrating an improvement in measures of artefact in a phantom does not necessarily translate to subjectively improved clinical images or an improvement in the conspicuity of lesions such as areas of periprosthetic osteolysis. Despite the improved measures the mean standard deviations remain high. It is also recognised that the gantry tilt angles described are relative to the CT table not to the long axis of the prosthesis. In the phantom the TKR was set in slight extension (2–3°) although axial



**Fig. 6.** A cropped screenshot from GE PACS© demonstrating the difference between images taken at (a) 10° angulation and (b) −10° angulation (anteroposterior to posterolateral) at the level of the polyethylene riser.



**Fig. 5.** Graph demonstrating the area under the curve for measurements of standard deviation in background attenuation adjacent to the femoral component.

reconstructions were performed parallel to a plane through the tibial tray. Therefore the data from this study is useful only to describe trends in artefact reduction rather than absolute measures. In fact the trends are more important than absolute measures when trying to translate these findings into clinical practice. In clinical practice it is unlikely that a TKR would be in perfect alignment with the axis of the CT machine and there are, of course, numerous variations on the basic design of a TKR which are likely to influence any absolute measures. We have not fully assessed the reproducibility in multiple patients. This has not been addressed in this project and is a potential aspect for further research.

CT reconstruction and post processing algorithms have made significant contributions to artefact reduction in total hip replacements but the complex morphology and the volume of metal in a TKR have so far resisted similar advances. The tilted gantry angle may offer one approach to addressing this. It is possible to demonstrate the effect that changing the gantry angle has on image sharpness at key areas (Fig. 6).

## 5. Conclusion

The most accurate attenuation values for tissues around a TKR can be obtained with a CT gantry using 10 to 15° anteroinferior to posterosuperior angulation. Optimal overall metal artefact reduction

for the TKR, and for the femoral component in isolation, can be obtained with a small (approximately 5°) angulation in the same direction. Specific CT imaging of the tibial component can be further improved by using even larger gantry angles (approximately 15°).

## References

- [1] Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. *Radiographics* 2004;24(6):1679–91.
- [2] Kalender WA, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. *Radiology* 1987;164(2):576–7.
- [3] Wang G, Frei T, Vannier MW. Fast iterative algorithm for metal artifact reduction in X-ray CT. *Acad Radiol* 2000;7(8):607–14.
- [4] Lee M, Kim S, Lee S, Song H, Huh Y, Kim D, et al. Overcoming artifacts from metallic orthopedic implants at high-field-strength MR imaging and multi-detector CT. *Radiographics* 2007;27(3):791–803.
- [5] Haramati N, Staron RB, Mazel-Sperling K, Freeman K, Nickoloff EL, Barax C, et al. CT scans through metal scanning technique versus hardware composition. *Comput Med Imaging Graph* 1994;18(6):429–34.
- [6] Link TM, Berning W, Scherf S, Joosten U, Joist A, Engelke K, et al. CT of metal implants: reduction of artifacts using an extended CT scale technique. *J Comput Assist Tomogr* 2000;24(1):165–72.
- [7] White LM, Buckwalter KA. Technical considerations: CT and MR imaging in the postoperative orthopedic patient. *Semin Musculoskelet Radiol* 2002;6(1):5–17.
- [8] Park W, Kim K, Shin H, Lee S. Reduction of metal artifact in three-dimensional computed tomography (3D CT) with dental impression materials. *Engineering in Medicine and Biology Society*, 2007. EMBS 200729th Annual International Conference of the IEEE; 2007. p. 3496–9.



# MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips

A.P. Toms<sup>a,\*</sup>, T.J. Marshall<sup>a</sup>, J. Cahir<sup>a</sup>, C. Darrah<sup>b</sup>, J. Nolan<sup>b</sup>,  
S.T. Donell<sup>c</sup>, T. Barker<sup>d</sup>, J.K. Tucker<sup>b</sup>

Departments of <sup>a</sup>Radiology, <sup>b</sup>Orthopaedics, and <sup>d</sup>Pathology, Norfolk and Norwich University Hospital, and <sup>c</sup>Institute of Health, University of East Anglia, Norwich, Norfolk, UK

Received 11 December 2006; received in revised form 17 May 2007; accepted 22 July 2007

**AIM:** To perform a retrospective review of all the conventional radiographic and magnetic resonance imaging (MRI) studies performed in patients with early postoperative pain following cobalt–chrome metal-on-metal total hip arthroplasty.

**METHODS:** A retrospective review of the radiology, surgical findings and histology in nineteen patients who had undergone a total of 20 hip arthroplasties using a cobalt-chromium on cobalt-chromium alloy prosthesis was undertaken.

**RESULTS:** Measures of implant placement on the immediate postoperative radiographs were all within the normal ranges ( $n = 20$ ). Where more than one postoperative radiograph was available statistical analysis revealed no evidence of progressive change before the MRI examination (14). The median postoperative time to MRI was 35 months (range 11–63 months). Abnormalities were demonstrated using MRI in all symptomatic hips ( $n = 20$ ). These comprised: periprosthetic fluid collections (20), which were isointense to muscle on T1-weighted images in 19 cases and hyperintense on T2-weighted images in 18 cases, periprosthetic bone marrow oedema ( $n = 6$ ), muscle oedema ( $n = 4$ ), avulsion of the gluteus minimus and medius tendons ( $n = 5$ ), atrophy of piriformis ( $n = 15$ ) and obturator internus ( $n = 17$ ), and fracture of the medial calcar ( $n = 1$ ). Operative findings in patients who had undergone revision surgery ( $n = 15$ ) included: fluid-filled cavities ( $n = 11$ ), soft tissue necrosis ( $n = 8$ ), gluteal tendon avulsion ( $n = 5$ ), proximal femoral diaphyseal necrosis ( $n = 4$ ), and pitting and corrosion of the femoral stems ( $n = 8$ ), which were, in all cases, firmly fixed to the cement mantle. Histology revealed viable tissue in six hips with necrosis ( $n = 12$ ) and fibrin deposition ( $n = 15$ ) being the predominate findings. Other findings included a perivascular lymphocytic infiltrate ( $n = 5$ ), features of active inflammation ( $n = 4$ ), and metallosis ( $n = 1$ ).

**CONCLUSION:** A significant number of patients with metal-on-metal hip replacements presented with early postoperative pain because of an abnormal soft-tissue reaction. MRI can demonstrate characteristic soft-tissue disease in these patients where conventional radiographs are frequently normal.

© 2007 The Royal College of Radiologists. Published by Elsevier Ltd. All rights reserved.

## Introduction

Magnetic resonance imaging (MRI) is increasingly being recognized as a useful tool for imaging the complications of total hip replacements (THR).<sup>1–4</sup> The introduction of metal artefact reduction

sequences (MARS) has reduced the amount of susceptibility artefact and increased the conspicuity of peri-prosthetic soft tissues.<sup>5–10</sup> Despite this, a clearly defined role for MRI is yet to be found in the imaging of THRs. For decades conventional radiographs have provided the mainstay of imaging of symptomatic THRs. Complications, such as infection and loosening, produce characteristic, but often overlapping, appearances.<sup>10,11</sup> Cross-sectional imaging [computed tomography (CT) and MRI] has not yet been shown to be useful in differentiating these disease processes, and

\* Guarantor and correspondent: A.P. Toms, Department of Radiology, Norfolk & Norwich University Hospital, Colney Lane, Norwich NR4 7UY, UK. Tel.: +44 1603 286104; fax: +44 1603 286077.

E-mail address: andoni.toms@nnuh.nhs.uk (A.P. Toms).

therefore, is not routinely incorporated into diagnostic algorithms.

Metal-on-metal total hip arthroplasties are enjoying a renaissance with a second generation of prostheses. Technical developments in computer numeric control (CNC) machining and the use of hip simulators have allowed the design and manufacture of components with tight tolerances on diametral clearance with improved precision compared with historical prostheses, offering the potential for bearings with more reproducible low-wear performance. These prostheses are being pursued in an effort to overcome the problems of small-particle disease associated with polyethylene wear, and in the hope of prolonging the survival time of the implant by reducing wear rates. In our institution a small proportion of patients with a second-generation metal-on-metal implant have presented with pain in the presence of normal inflammatory markers and normal plain radiographic imaging. The exact mechanism for the failure is not yet fully understood but histological analysis reveals a unique lymphocytic perivascular infiltration,<sup>12</sup> which maybe the result of a hypersensitivity reaction.<sup>13,14</sup>

The present study is a retrospective review of the radiological, surgical and histological findings in 19 patients who experienced pain early in the postoperative course after a total of 20 primary metal-on-metal hip replacements, all of whom had subsequent abnormal MRI examinations.

## Materials and methods

A retrospective review of the radiology, surgical findings, and histology in 19 patients who had undergone a total of 20 hip arthroplasties using a cobalt-chromium on cobalt-chromium alloy prosthesis (De Puy International Ltd., Leeds UK) was undertaken. The first 20 hip replacements considered to be abnormal on MRI examination were included in the study. These 20 hips comprise 3% of 648 of these hips implanted between 1997 and 2005. Approval from Research Governance and Ethics Committees is not required for this type of study at our institution. Eleven women and eight men, aged 37–74 years old, who presented with early postoperative pain (defined as pain occurring within 3 years of the arthroplasty), were included in the study. The THRs had been performed by one of three surgeons each of whom used their preferred approach. The posterior approach included re-attachment of the short external rotators. The antero-lateral approach was also used, but less commonly. All 20 hips were imaged with MRI at one of two local institutions between March 2002 and November 2005.

Thirteen of the patients had at least two sets of posteroanterior (PA) radiographs of the pelvis and a lateral radiograph of 14 hip prostheses between the operation and the MRI examination. Five patients had only one set of radiographs after operation and before MRI, one patient had a coned anteroposterior (AP) examination of the treated hip only. Hardcopy radiographs for four patients were digitized and archived to PACS, all the rest were digitally acquired as computerized radiographs. Where available the first and the last post-operative radiographs were evaluated by two observers: a musculoskeletal radiologist (AT) and an orthopaedic surgeon (SD). The following measures were acquired from each radiograph by consensus opinion: acetabular inclination, acetabular cup height, leg length, lateral offset and femoral stem alignment using standardized measures.<sup>15</sup> Measures for the digitized radiographs were extrapolated from the diameter of the acetabular cup, which was known for each patient. Differences in the immediate postoperative radiographs and those before the MRI examination were tested using a paired Student's *t* test where the distribution of results was normal, abnormal distributions of data were compared using Wilcoxon's signed ranks test (SPSS). The presence of stress shielding,<sup>16</sup> heterotopic ossification,<sup>17,18</sup> osteolysis and loosening,<sup>19–25</sup> and the quality of the cement mantle<sup>26</sup> were also noted and graded.

Twenty symptomatic hips were imaged using MRI on one of two 1.5 T machines (Signa, GE, General Electric Healthcare, Milwaukee, Wisconsin and Symphony, Siemens, Erlangen, Germany). MRI was performed at a mean of 35 (range 11–63) months after the primary arthroplasty. The first seven patients underwent examination with conventional fast spin-echo sequences: coronal T1-weighted (T1W) spin-echo [echo time (TE) 12 ms, repetition time (TR) 360 ms], axial T1W spin-echo (TE 12 ms, TR 340 ms), and inversion recovery sequences: coronal short tau inversion recovery (STIR; TE 37 ms, TR 5360 ms), axial STIR (TE 41 ms, TR 3000 ms) with a matrix size of up to 512 × 256 and a bandwidth of 128 MHz. For the latter 13 patients metal artefact-reduction sequences (MARS) were employed: coronal T1W turbo spin-echo (TE 23 ms, TR 669 ms) and STIR (TE 37 ms, TR 3840 ms), axial T1W (TE 23 ms, TR 534 ms) and T2W turbo spin-echo (TE 69 ms, TR 5600 ms) of the whole pelvis and a sagittal T2W turbo spin-echo of the hip (TE 69 ms, TR 2900 ms). Section thickness 5 mm, field of view 340 × 340 mm, matrix size up to 448 × 336, pixel bandwidth 620 MHz. The MRI examinations were evaluated by a consensus of three musculoskeletal radiologists (A.T., T.M., J.C.) for the

presence of a periprosthetic soft-tissue abnormality, noting its size, signal characteristics, and involvement of one of the following compartments: the gluteal, adductor, quadriceps, and hamstring. Involvement of a compartment was defined as involvement of any of the constituent muscles. The presence of muscle atrophy was defined as loss of volume and the presence of fatty replacement when compared with the contralateral side. This was recorded for periprosthetic muscles in any of the involved compartments, as well as the piriformis and obturator internus muscles. Avulsion of the muscle tendon was defined as a discontinuity of the low signal muscle attachment. All radiological data were viewed on a dedicated PACS 2 K monitor (Barco, Kortrijk, Belgium).

Fourteen out of the 19 patients (15 hips) have had revision arthroplasties after the MRI examination. Two observers (J.N., A.T.) reviewed the operative notes retrospectively providing a consensus opinion. The following points were recorded: the presence of abnormal peri-prosthetic soft tissue, fluid, necrosis, abnormal tissue vascularity, oedema, and tendon avulsions. The fixity of the prosthesis and the presence of any corrosion were noted. A histopathologist with a specialist interest in musculoskeletal disease (T.B.) reviewed the histology specimens. The histological findings were classified as demonstrating necrosis, infection, granuloma formation, fibrin deposition or metallosis,<sup>27</sup> or peri-vascular lymphocytic infiltrate.<sup>12</sup>

## Results

### Conventional radiographs (Table 1)

Acetabular inclination on the first postoperative radiograph measured between 31° and 65°. In those patients with two comparable radiographs prior to MRI ( $n = 14$ ) the change in acetabular inclination varied from 0° to 3°, which was not statistically significant ( $p = 0.55$ ). The acetabular cup height ( $n = 14$ ) ranged from -10 to 12 mm with temporal changes of between 0 and 9 mm

( $p = 0.33$ ). The range of leg length measurements ( $n = 14$ ) varied from -11 to 13 mm, these measures changed by 0 to 8 mm on subsequent radiographs ( $p = 0.40$ ). The lateral offset produced a range ( $n = 14$ ) from -10 to 16 mm. Changes in these measures ranged from 0 to 18 mm ( $p = 0.08$ ). The stem position on the AP radiograph ( $n = 14$ ) ranged from 6° valgus to 5° varus. These measures varied from between 0 and 3° ( $p = 0.38$ ).

Stress shielding was noted in the greater trochanter in six out of 14 follow-up radiographs (Table 2). In two hips this was combined with demineralization of the lesser trochanter. There was no osteolysis or heterotopic ossification on any of the radiographs. One patient had a small fracture of the tip of the greater trochanter on the immediate postoperative radiograph. On the immediate postoperative radiographs the cement mantle was classified as a grade A in 18, and B and C in one each of the hips. There were no changes in the grading on subsequent radiographs. Dislocations were noted to have occurred in three hips.

### MRI (Table 3)

A periprosthetic soft-tissue abnormality was present on all 20 MRIs. This predominantly comprised intermediate T1W signal (19/20), which encircled (18/20) or abutted (2/12) the neck of the femoral prosthesis and which measured, on average, 7 cm in maximal mediolateral (range 2.5–10 cm), 6.5 cm in maximal AP (range 1.5–10 cm) and 10 cm in maximal craniocaudal diameter (3–30 cm; Figs. 1 and 2). In one hip, the lesion was hyperintense to muscle on T1W sequences, in 18 out of 20 cases the abnormality consisted of hyperintense fluid-like signal on T2W, and in two cases the abnormal areas were isointense. All 18 cases that demonstrated hyperintense fluid-like T2 signal had an irregular low signal periphery varying from 1–5 mm in thickness. The periarticular soft-tissue changes extended into the gluteal compartment in 18 cases. The quadriceps compartment was involved in seven cases and the adductor compartment in two cases.

**Table 1** Table summarising the radiographic position of 14 of the hip replacements prior to MRI

	First postoperative radiograph	Last postoperative radiograph	Range of differences	Test of differences
Acetabular inclination	31–57°	30–57°	0–3°	$p = 0.44^*$
Acetabular cup height	-10–14 mm	-7–10 mm	0–9 mm	$P = 0.34^*$
Leg length	-11–13 mm	-14–11 mm	0–8 mm	$P = 0.30^*$
Lateral offset	-10–18 mm	-20–14 mm	0–18 mm	$p = 0.12^*$
Stem position	-5–6°	-6–5°	0–3°	$p = 0.39^†$

\* Two-tailed Wilcoxon signed rank test.

† Two-tailed Paired Student t test.

**Table 2** Table summarising the plain radiographic features related to bone and cement mantle quality of 14 of the hip replacements prior to MRI

Radiographic feature	Result
Stress shielding	6 out of 14 Greater trochanter 6 Lesser trochanter 2
Heterotopic ossification	Nil
Osteolysis	Nil
Fracture	1 – Tip of greater trochanter on immediate postoperative radiograph
Femoral cement mantle	A 18 B 1 C 1

In one patient, the soft-tissue abnormality communicated through a sinus laterally just posterior to the greater trochanter.

Atrophy was only noted in the gluteal muscles and the short external rotators. Gluteus maximus was involved in two cases, but in all cases the gluteus maximus insertion remained intact. Gluteus medius was atrophic in eight cases and gluteus minimus in nine cases. Avulsion of the gluteus medius and minimus tendons was present. Table summarising the abnormal MRI findings in 20 symptomatic metal-on-metal hip replacements in

**Table 3** Table summarising the abnormal MRI findings in 20 symptomatic metal-on-metal hip replacements

MRI feature	Result
Periprosthetic soft-tissue lesion	20 of 20
MRI signal	T1W: isointense 19, hyperintense 1 T2W: hyperintense 18, isointense 2
Size	Mediolateral: mean 7 cm (2.5–10) Anteroposterior: mean 6.5 cm (1.5–10) Craniocaudal: mean 10 cm (3–30)
Compartments	Gluteal 18 Quadriceps 7 Adductor 2
Muscle atrophy	Gluteus maximus 2 Gluteus medius 8 Gluteus minimus 9 Piriformis 15 Obturator internus 17
Tendon avulsion	Gluteus medius and minimus 5
Muscle oedema	Obturator externus 1 Rectus femoris 1 Vastus lateralis 1 Adductor brevis 1
Bone marrow oedema	Greater and lesser trochanters 6
Fracture	Medial calcar 1

five patients (Fig. 3). Atrophy was recorded in 17 obturator internus and 15 piriformis muscles. Muscle oedema was noted in three atrophic gluteus medius and two gluteus minimus muscles, and once in each of the following: obturator externus, rectus femoris, vastus lateralis, and adductor brevis.

Bone marrow oedema was recorded in the greater trochanter and lesser trochanters in six of 15 hips where the trochanters could be adequately visualized (Fig. 4). A fracture of the medial calcar was demonstrated in one patient who had had a normal radiograph 3 months previously (Fig. 2).

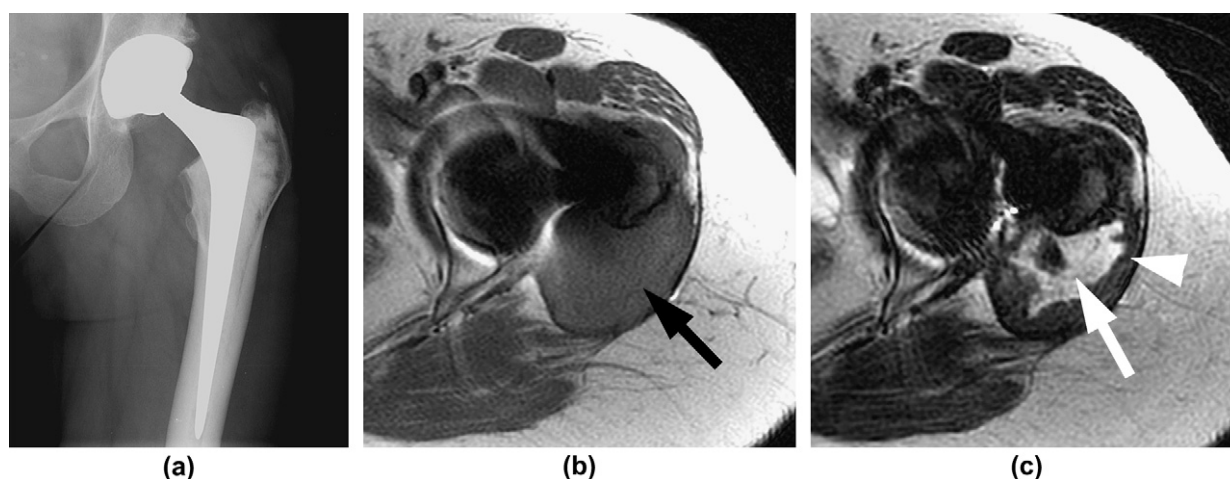
To date 15 patients have undergone revision total hip arthroplasty. The following findings were recorded in the operative notes (Table 4). Discrete peri-prosthetic soft-tissue thickening, a soft-tissue mass, a fluid-filled cavity, or a combination of these findings, was noted in all revised hips. Macroscopic soft-tissue necrosis was present in eight cases, avascular (but not frankly necrotic) soft tissue in 10 cases, avascular proximal femur in four cases, and gluteal tendon avulsion in five cases (Fig. 5). The femoral and acetabular implants were firmly fixed in all 15 cases. Corrosion and severe pitting was demonstrated on the femoral component in eight hips (Fig. 6) but all the acetabular components were normal. Samples taken from the macroscopically abnormal peri-prosthetic soft tissues from all 15 revised hips were sent for microbiological assessment but revealed no infective organisms in any case.

Only six of the 15 samples contained viable tissue (Table 5). The predominant histological finding was one of necrosis (12/15) and fibrin deposition (15/15; Fig. 7a). A perivascular lymphocytic infiltrate occurred in five patients (Fig. 7b). This was accompanied by eosinophils in four patients. Four samples contained granulomas (in two of the samples the granulomas were typical of rheumatoid nodules) and four samples had evidence of active inflammation consistent with infection. In one sample only were there features of metallosis.

## Discussion

The aetiology of failure in the new generation of cobalt–chromium metal-on-metal total hip arthroplasties is not yet fully understood, but is the subject of ongoing investigation. This paper presents the radiological findings in the first 20 hips to be imaged with, and considered to be abnormal on, MRI. A significant number of patients are continuing to present with pain early in their postoperative



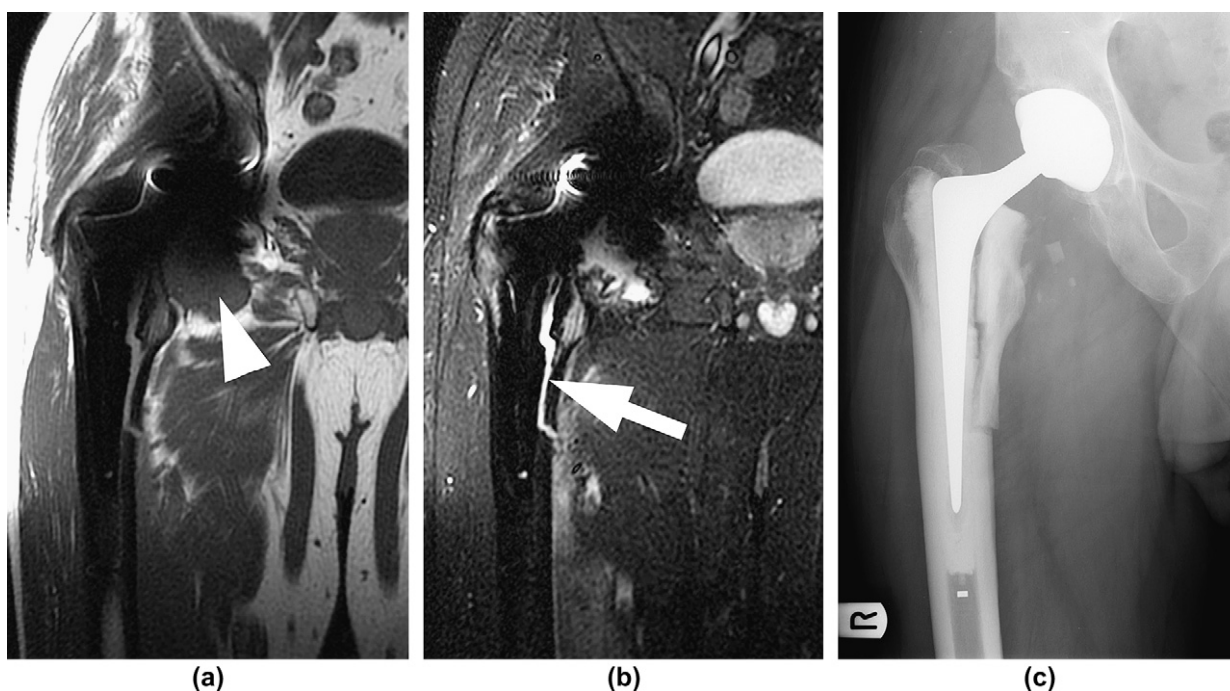


**Figure 1** (a) Normal plain radiograph of a metal-on-metal THR in a patient with a 20-week history of increasing hip pain. (b) Axial T1W image demonstrates a soft-tissue abnormality of intermediate signal (arrow) extending from the neck into the peri-prosthetic soft tissues. (c) Axial STIR image demonstrates a ragged intermediate signal rim (arrow-head) surrounding a fluid signal centre (arrow). At operation a cavity, lined by sheet-like avascular fibrous material, containing milky-white fluid was found.

course following these particular metal-on-metal prostheses. The findings at MRI described in this paper continue to be used in our institution to confirm or exclude the presence of peri-articular soft-tissue disease and influence the decision to revise the prosthesis. Early published impressions are that this is a disease process unique to metal-on-metal

bearings occurring as a result of a hypersensitivity reaction that causes a characteristically dense peri-vascular lymphocytic infiltrate.<sup>12</sup>

This histological appearance was found in only five out of the 15 samples, but this is probably unrepresentative because it was present in five out of the six samples containing viable tissue (Fig. 7b).

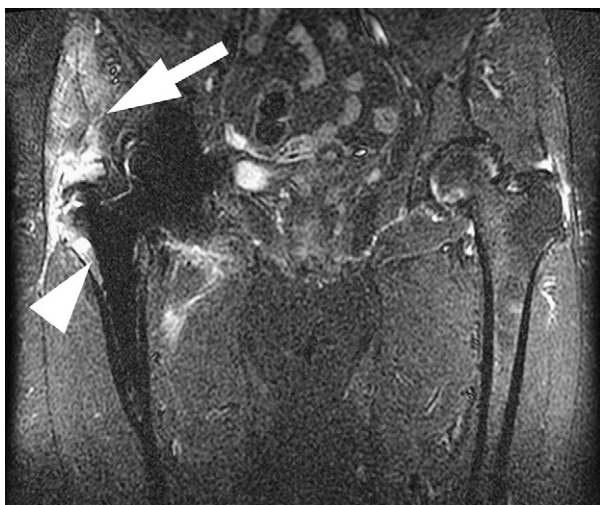


**Figure 2** Coronal T1W (a) and STIR (b) MRI demonstrating an inflammatory mass abutting the neck of the femoral prosthesis (arrowhead) with fluid signal tracking distally (arrow) deep to the fractured medial calcar. A fracture had not been suspected clinically and the previous radiographs were normal. A subsequent radiograph (c) confirmed the fracture of the medial calcar, which at surgery was found to be devitalized.



**Figure 3** Coronal T1W MRI of the left hip demonstrating a soft-tissue abnormality, comprising a hyperintense central area (arrow) and an intermediate intensity rim (arrowhead), lying superior and lateral to the left greater trochanter in place of the gluteus medius and minimus tendons. At operation the gluteal tendons, which were devitalized, were found avulsed from the greater trochanter within a collection of milky white fluid (Fig. 5).

All 15 samples contained extensive fibrin deposition, which was many millimetres thick on the prosthetic side of the sample. It is unlikely that this represents the 1 mm layer of fibrin previously noted surrounding "stable" prostheses.<sup>28</sup> In two of the samples, the perivascular lymphocytic

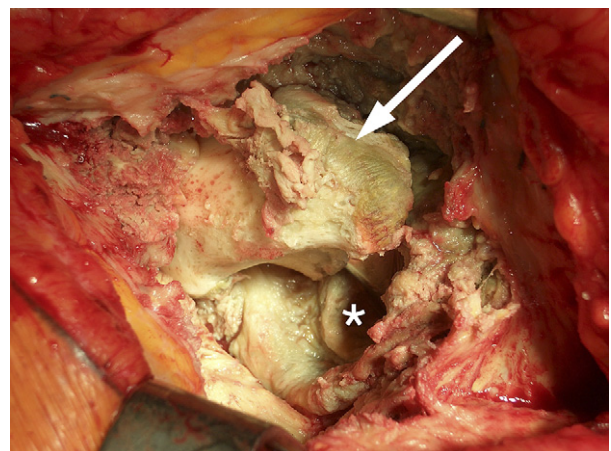


**Figure 4** Coronal STIR MRI demonstrating oedema within gluteus maximus (arrow) extending down to the greater trochanteric insertion. There is also bone marrow oedema within the greater trochanter (arrowhead).

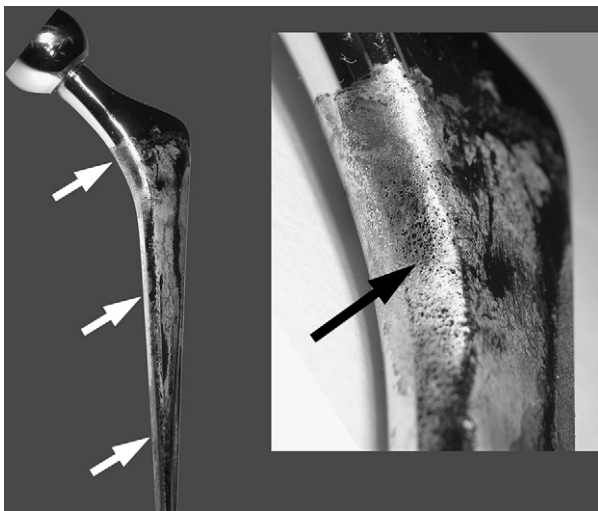
**Table 4** Table summarising the macroscopic operative findings in 15 patients at revision surgery

Surgical finding	Number of cases ( <i>n</i> = 15)
Abnormal peri-prosthetic soft-tissue thickening or mass	10
Fluid	11
Soft-tissue necrosis	8
Avascular soft tissue	10
Avascular bone	4
Tendon avulsion	5
Soft-tissue oedema	1
Peri-prosthetic fracture	1
Prosthetic corrosion	Femoral component = 8 Acetabular component = 0

infiltrate was accompanied by evidence, in the form of neutrophils and plasma cells, of active inflammation. In both these cases there was no evidence of necrosis, and therefore, these appearances could be due to infection, although no organisms were grown from any of the 15 revised hips. Another two cases of perivascular lymphocytic infiltrate were accompanied by granulomas with features that suggested rheumatoid nodules, and therefore, an inflammatory process not caused by the prosthesis. Only one sample out of 15 had evidence of metallosis despite pitting of the femoral stem being recorded in eight cases. It may be that this is a disease process similar to metal particulate debris described by other authors,<sup>29,30</sup> although there was no evidence of this in any but one of the samples. This may be because the particle size in this particular process is too small to be demonstrated by conventional light microscopy. There are certainly features that correspond to the previously described metal-on-metal soft-tissue reactions,<sup>12,13</sup> but histological sampling is



**Figure 5** Intra-operative photograph demonstrating devitalized proximal femoral diaphysis (asterisk) and necrotic avulsed gluteal tendons (arrows).



**Figure 6** Photograph of cobalt–chromium alloy femoral prosthesis removed from symptomatic patient demonstrating pitting corrosion (black arrow) affecting the portion of the prosthesis in contact with the cement mantle (white arrow).

incomplete, necessitating a change in the sampling technique used at revision hip surgery in these patients. In the viable samples, the histological features present are a little more heterogeneous than previous descriptions, which may be the result of other disease processes, such as infection or pre-existing connective tissue disease.

Analysis of the conventional radiographs produced a range of measures that did not change significantly from the first postoperative radiograph to the pre-MRI film (Table 1). The smallest range of changes in measurements occurred for acetabular inclination and stem position, whereas

the greatest range of changes occurred for the measures of lateralization of the hip. This occurred because the patients' pelvises were not always projected squarely on to the radiograph, and therefore, the femur was in varying positions of varus and valgus angulation. This will affect measures of lateralization the most and measures of acetabular inclination and stem position the least. Observation of stress shielding and cement grading (Table 2) were unremarkable, although there is some doubt as to whether these features can be reliably evaluated on plain radiographs.<sup>16,31</sup> Our interpretation of the plain radiographic findings is that there is no statistical, or subjective, evidence of progressive plain radiographic abnormalities and certainly no indications of the extensive soft-tissue disease demonstrated on MRI and at surgery.

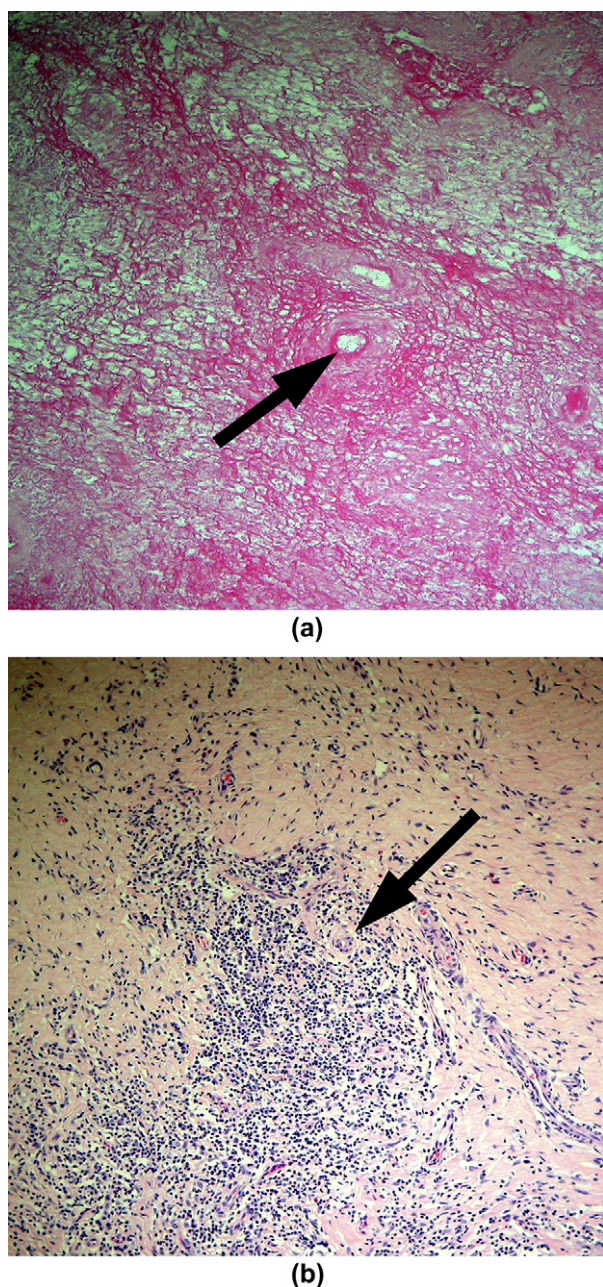
The most noticeable abnormality on MRI, present in all 20 hips, was the presence of soft-tissue abnormalities that were predominantly of intermediate T1W signal (19/20) and a high T2W signal (18/20) and consistent with fluid collections. These collections were of variable size, but always maintained contact with the neck of the prostheses. The collection, surrounded by its irregular wall, most commonly extended in to the gluteal compartment (18/20) where muscle atrophy of gluteus minimus and medius was noted in nearly half of cases. These findings would be consistent with an infected peri-prosthetic abscess, but in every case inflammatory markers were normal, and where the hips have been explored, all samples have proven to be sterile.

Muscle oedema was present in four peri-articular muscles (Table 3). The significance of muscle

**Table 5** Table summarising the histological findings in samples taken at the time of revision surgery

Patient	Histology							
	Viable tissue	Necrosis	Fibrinous tissue	Perivascular lymphocytes	Active inflammation	Eosinophils	Granulomas	Metallosis
1		X	X					
2	X	X	X		X			
3		X	X					
4		X	X					
5	X	X	X	X		X	X	
6	X	X	X	X		X	X	
7		X	X					
8	X	X	X	X		X		
9		X	X					
10	X		X	X	X	X		
11	X		X	X	X			
12		X	X				X	
13		X	X					X
14		X	X					
15			X		X			





**Figure 7** Photomicrographs of a haematoxylin and eosin stained histological section demonstrating (a) extensive tissue necrosis and (b) a dense perivascular lymphocytic infiltrate. (a) A network of fibrin surrounds "ghosts" of now dead cellular material such as the vascular channel marked with an arrow. (b) Lymphocyte proliferation surrounds a blood vessel (arrow) in the pseudocapsule of the hip joint.

oedema following THR is not clear. Greater trochanteric bursitis is a recognized complication of THR,<sup>32</sup> and therefore, is likely to be associated in some cases with gluteal muscle oedema. Avulsion of the hip abductors is well described on MRI and is associated with symptoms following hip replacement,<sup>4,33</sup> and it would be reasonable conjecture that avulsion

might be preceded by muscle oedema. Muscle oedema may also be dependent on the operative approach, antero-lateral, lateral, or posterior, but there are few data describing either the distribution or longevity of this in normal asymptomatic patients.

Bone marrow oedema in the proximal femoral diaphysis was a common finding that appears to correlate in some patients with operative descriptions of pale, bloodless, necrotic bone. Again the significance of bone marrow oedema is uncertain because of a lack of data in asymptomatic patients with THRs. In the present series a single patient had bone marrow oedema in their contralateral asymptomatic hip, but a single, retrospectively reviewed case is not helpful. It seems reasonable to presume that in the early postoperative stages oedema would be present and that this would resolve with time. Displaced marrow into the greater trochanter is considered by some authors to be a cause of increased radionuclide uptake on bone scintigraphy.<sup>34,35</sup> These appearances can be present for several years after the operation, and therefore, some MRI signal abnormalities might also be expected in asymptomatic patients. The biomechanical stresses of a loose prosthesis and infection would also be associated with periprosthetic oedema.

Avulsion of the hip abductors has recently been described in symptomatic patients following hip arthroplasty through anterolateral<sup>33</sup> and lateral transgluteal<sup>4</sup> approaches and is well demonstrated by MRI. Most of the present patients had their operations performed through a posterior approach. During closure the short external rotators (SERs) are reattached to the medial concavity of the greater trochanter. The MRI in 11 hips covered the short external rotator territory. In 10 of these 11 hips the SERs were abnormal. Where the periprosthetic soft-tissue mass was extensive it replaced the SERs. In those with smaller inflammatory masses at least one SER was avulsed. In these patients the individual SERs could be identified apart from the gemelli, which were demonstrated as one. Repair of the SERs is considered by many orthopaedic surgeons to protect against dislocation,<sup>36,37</sup> but this is not a universally held opinion.<sup>38</sup> There is some evidence that most repairs of the SERs fail.<sup>39</sup> As yet there are no descriptions of the normal MRI appearances of the SERs following surgery, and therefore, the present findings may have little to do with the disease process in these patients. The preoperative MRI appearances of the SERs also need to be documented because atrophy secondary to the severe osteoarthritis of the hip may affect the postoperative MRI findings.



The present study has a number of limitations. The radiological imaging and biopsies have not been performed to a standardized protocol nor have surgical observations been recorded systematically. Because of this it is not possible to directly correlate radiological findings with histology and operative findings because data acquisition has been heterogeneous. Therefore, the data have been analysed by consensus review to produce a radiological, surgical, and histological description of our clinical experience. A prospective study is required to evaluate the ability of MRI to accurately demonstrate this novel disease process. Although MRI has been a primary influence on the surgeons' decision to revise these prostheses, long-term follow-up is required to evaluate the negative predictive value of MRI.

The radiological findings are in contrast to hip arthroplasties that fail because of small particle disease or abnormal loading, and result in stress reactions, where the plain radiographs are often abnormal. Although metal artefact reduction sequences have been optimized for several years,<sup>5–10</sup> the role of MRI in imaging of the postoperative hip, particularly for infection and prosthetic loosening, has yet to be defined for these clinical indications. This series of patients clearly demonstrates a pivotal role for MRI in imaging symptomatic metal-on-metal hip replacements where the primary disease process occurs in the peri-prosthetic soft tissues and where, unlike other diseases of hip arthroplasties, conventional radiography is typically normal.

In conclusion, a significant number of patients treated with a new-generation metal-on-metal hip replacement present with pain after an early pain-free recovery period. In the authors' experience plain radiographs are normal but MRI, particularly when employing metal artefact-reduction sequences, demonstrates a range of abnormalities in these patients. A peri-prosthetic, fluid-filled, soft-tissue cavity is characteristic. Muscle and bone marrow oedema, and avulsion of the SERs are frequent findings, but their significance requires further research by comparing with the MRI findings in asymptomatic patients following hip arthroplasty.

## Acknowledgements

De Puy International Ltd, Leeds, UK, funded 15 of the MRI studies described in this paper. The authors thank Mrs Angela Britcher for her assistance during the preparation of this manuscript.

## References

1. Czerny C, Krestan C, Imhof H, et al. Magnetic resonance imaging of the postoperative hip. *Top Magn Reson Imaging* 1999;10:214–20.
2. Potter HG, Nestor BJ, Sofka CM, et al. Magnetic resonance imaging after total hip arthroplasty: evaluation of periprosthetic soft tissue. *J Bone Joint Surg Am* 2004;86-A:1947–54.
3. White LM, Kim JK, Mehta M, et al. Complications of total hip arthroplasty: MR imaging-initial experience. *Radiology* 2000;215:254–62.
4. Pfirrmann CW, Notzli HP, Dora C, et al. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology* 2005;235:969–76.
5. Chang SD, Lee MJ, Munk PL, et al. MRI of spinal hardware: comparison of conventional T1-weighted sequence with a new metal artefact reduction sequence. *Skeletal Radiol* 2001;30:213–8.
6. Olsen RV, Munk PL, Lee MJ, et al. Metal artefact reduction sequence: early clinical applications. *RadioGraphics* 2000;20:699–712.
7. Suh JS, Jeong EK, Shin KH, et al. Minimizing artefacts caused by metallic implants at MR imaging: experimental and clinical studies. *AJR Am J Roentgenol* 1998;171:1207–13.
8. Eustace S, Goldberg R, Williamson D, et al. MR imaging of soft tissues adjacent to orthopaedic hardware: techniques to minimize susceptibility artefact. *Clin Radiol* 1997;52:589–94.
9. Eustace S, Jara H, Goldberg R, et al. A comparison of conventional spin-echo and turbo spin-echo imaging of soft tissues adjacent to orthopedic hardware. *AJR Am J Roentgenol* 1998;170:455–8.
10. Tormanen J, Tervonen O, Koivula A, et al. Image technique optimization in MR imaging of a titanium alloy joint prosthesis. *J Magn Reson Imaging* 1996;6:805–11.
11. Allen AM, Ward WG, Haygood TM. Imaging of the total hip arthroplasty. *J South Orthop Assoc* 1995;4:96–123.
12. Davies AP, Willert HG, Campbell PA, et al. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am* 2005;87:18–27.
13. Willert HG, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am* 2005;87:28–36.
14. Park YS, Moon YW, Lim SJ, et al. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am* 2005;87:1515–21.
15. Johnston RC, Fitzgerald RHJ, Harris WH, et al. Clinical and radiographic evaluation of total hip replacement. A standard system of terminology for reporting results. *J Bone Joint Surg Am* 1990;72:161–8.
16. Engh CAJ, McAuley JP, Sychterz CJ, et al. The accuracy and reproducibility of radiographic assessment of stress-shielding. A postmortem analysis. *J Bone Joint Surg Am* 2000;82-A:1414–20.
17. Payne LZ, DeLuca PA. Heterotopic ossification after rhizotomy and femoral osteotomy. *J Pediatr Orthop* 1993;13:733–8.
18. Brooker AF, Bowerman JW, Robinson RA, et al. Ectopic ossification following total hip replacement. Incidence and a method of classification. *J Bone Joint Surg Am* 1973;55:1629–32.
19. Chambers IR, Fender D, McCaskie AW, et al. Radiological features predictive of aseptic loosening in cemented Charnley femoral stems. *J Bone Joint Surg Br* 2001;83:838–42.

20. Harris WH, McCarthy JCJ, O'Neill DA. Femoral component loosening using contemporary techniques of femoral cement fixation. *J Bone Joint Surg Am* 1982;**64**:1063–7.
21. Hodgkinson JP, Shelley P, Wroblewski BM. The correlation between the roentgenographic appearance and operative findings at the bone–cement junction of the socket in Charnley low friction arthroplasties. *Clin Orthop Relat Res* 1988;**228**:105–9.
22. Kobayashi A, Donnelly WJ, Scott G, et al. Early radiological observations may predict the long-term survival of femoral hip prostheses. *J Bone Joint Surg Br* 1997;**79**:583–9.
23. Loudon JR, Charnley J. Subsidence of the femoral prosthesis in total hip replacement in relation to the design of the stem. *J Bone Joint Surg Br* 1980;**62-B**:450–3.
24. Madey SM, Callaghan JJ, Olejniczak JP, et al. Charnley total hip arthroplasty with use of improved techniques of cementing. The results after a minimum of fifteen years of follow-up. *J Bone Joint Surg Am* 1997;**79**:53–64.
25. Mulroy RDJ, Harris WH. The effect of improved cementing techniques on component loosening in total hip replacement. An 11-year radiographic review. *J Bone Joint Surg Br* 1990;**72**:757–60.
26. Barrack RL, Mulroy RDJ, Harris WH. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. *J Bone Joint Surg Br* 1992;**74**:385–9.
27. Bullough PG. Tissue response to implanted prosthesis. *Orthopaedic pathology*. London: Mosby; 2003. p. 327–346.
28. Saito S, Ryu J, Watanabe M, et al. Midterm results of Meta-sul metal-on-metal total hip arthroplasty. *J Arthroplasty* 2006;**21**:1105–10.
29. Boardman DR, Middleton FR, Kavanagh TG. A benign psoas mass following metal-on-metal resurfacing of the hip. *J Bone Joint Surg Br* 2006;**88**:402–4.
30. Shetty VD, Villar RN. Development and problems of metal-on-metal hip arthroplasty. *Proc Inst Mech Eng [H]* 2006;**220**:371–7.
31. Harvey EJ, Tanzer M, Bobyn JD. Femoral cement grading in total hip arthroplasty. *J Arthroplasty* 1998;**13**:396–401.
32. Berquist TH, Bender CE, Maus TP, et al. Pseudobursae: a useful finding in patients with painful hip arthroplasty. *AJR Am J Roentgenol* 1987;**148**:103–6.
33. Twair A, Ryan M, O'Connell M, et al. MRI of failed total hip replacement caused by abductor muscle avulsion. *AJR Am J Roentgenol* 2003;**181**:1547–50.
34. Miles KA. The painful prosthetic joint. In: Peters AM, editor. *Nuclear medicine in radiological diagnosis*. London: Martin Dunitz; 2003. p. 55–64.
35. Palestro CJ, Love C, Tronco GG, et al. Combined labeled leucocyte and technetium 99m sulfur colloid bone marrow imaging for diagnosing musculoskeletal infection. *RadioGraphics* 2006;**26**:859–70.
36. Dixon MC, Scott RD, Schai PA, et al. A simple capsulorrhaphy in a posterior approach for total hip arthroplasty. *J Arthroplasty* 2004;**19**:373–6.
37. Weeden SH, Paprosky WG, Bowling JW. The early dislocation rate in primary total hip arthroplasty following the posterior approach with posterior soft-tissue repair. *J Arthroplasty* 2003;**18**:709–13.
38. Stahelin T, Vienne P, Hersche O. Failure of reinserted short external rotator muscles after total hip arthroplasty. *J Arthroplasty* 2002;**17**:604–7.
39. Kao JT, Woolson ST. Piriformis tendon repair failure after total hip replacement. *Orthop Rev* 1992;**21**:171–4.

## CASE REPORT

# Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR

<sup>1</sup>A P TOMS, FRCR, <sup>2</sup>J NOLAN, FRCS(Orth), <sup>3</sup>T BARKER, FRCPath, <sup>2</sup>C DARRAH, RGN Dip Sci and <sup>1</sup>P MALCOLM, FRCR

Departments of <sup>1</sup>Radiology, <sup>2</sup>Orthopaedics and <sup>3</sup>Pathology, Norfolk & Norwich University Hospital Trust, Colney Lane, Norwich, Norfolk NR4 7UY, UK

**ABSTRACT.** Metal-on-metal hip replacements are the subject of much current debate. There is some evidence that there may be a hypersensitivity reaction, specific to metal-on-metal total hip replacements (THRs), which is associated with early failure of these prostheses. It has to date only been described in total replacements and not in metal-on-metal hip resurfacing. We present the case of a 68-year-old man who underwent bilateral metal-on-metal hip resurfacing for osteoarthritis. The patient presented 6 months after surgery with pain and lateral thigh swelling. Pre-operative ultrasound and MRI demonstrated findings similar to those described in early failing metal-on-metal THRs, as well as evidence of lymphoreticular spread of metal debris. The operative findings included extensive aseptic soft-tissue necrosis. Histology revealed necrosis and a dense perivascular lymphocytic infiltrate along with metal debris within sinus histiocytes. The surgical, radiological and histological findings are similar to soft-tissue reactions described in metal-on-metal THRs.

Received 6 February 2008  
Revised 22 March 2008  
Accepted 16 April 2008

DOI: 10.1259/bjr/32928628

© 2009 The British Institute of  
Radiology

Metal-on-metal resurfacing of the hip has become an increasingly popular choice of arthroplasty for younger patients with osteoarthritis [1–3]. Initial reports have described promising results at short- and medium-term follow-up [1, 4–6]. However there are increasing reports of soft-tissue reactions, leading to failure of the prosthesis, that appear to be specific to metal-on-metal bearings [7, 8]. These have been described in total hip replacements (THRs) rather than resurfacing procedures. This case report describes the failure of a metal-on-metal resurfacing prosthesis due to metallosis 6 months after surgery. The diagnosis of a metal-on-metal soft-tissue reaction and lymphoreticular spread of metallic debris was suspected on pre-operative MRI.

## Case report

A 68-year-old retired large-animal vet underwent bilateral Birmingham resurfacing hip replacements for advanced osteoarthritis in January 2006. Prior to surgery, the patient was otherwise fit and well with no other past medical history of note. The surgery was uneventful and clinical progress was good in the early post-operative period. In July 2006 (6 months after surgery), the patient developed bilateral discomfort in the hips, particularly after activity. Clinical examination

revealed apparent bilateral defects in the fascia lata and bilateral palpable swelling arising from the lateral aspect of the thigh.

Conventional radiographs (Figure 1) demonstrated a normally aligned prosthesis that had not changed since the immediate post-operative films. The adjacent bones and soft tissues were normal.

Ultrasound examination of the palpable lumps revealed large anechoic fluid collections, which extended from a defect in the iliotibial tract proximally into the gluteal muscles. The fluid extended superiorly and medial to the greater tuberosity to lie adjacent to the prosthesis. Thick fronds of what were assumed to be fibrinous material lined the cavities (Figure 2).

MRI demonstrated bilateral large fluid-filled cavities that extended for approximately 13 cm cranio-caudad from the gluteal muscles distally, lateral to the greater trochanter, and into the subcutaneous fat of the lateral thigh through the iliotibial tract. The fluid was hyperintense on  $T_2$  weighted and isointense to muscle on  $T_1$  weighted images. The wall of the cavity was smooth, thin and of low signal on  $T_2$  weighted images. There was atrophy of both the gluteus medius and minimus. On the left, between the gluteus medius and minimus, approximately 6 cm caudad to the prosthesis, there lay a 1.5 cm in diameter lobulated lesion at the most caudad extent of the cavity. This lesion was isointense to muscle on  $T_1$  weighted images and of low signal, surrounded by an ill-defined bloom, on  $T_2$  weighted images (Figure 3).

At surgery, the joint was found to be lined by a pale membrane, which communicated with the cavity that

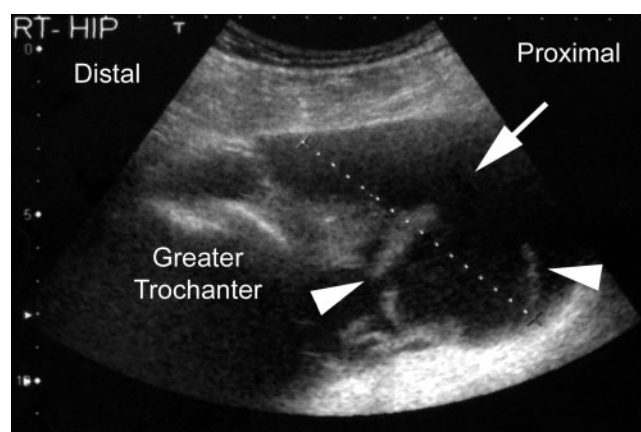
Address correspondence to: Andoni Toms, Department of Radiology, Norfolk & Norwich University, Hospital Trust, Colney Lane, Norwich, Norfolk NR4 7UY, UK. E-mail: andoni.toms@nnuh.nhs.uk



**Figure 1.** Conventional anteroposterior radiograph of the left hip demonstrating a Birmingham resurfacing hip replacement 6 months after surgery. Alignment is normal and there are no bone or soft-tissue abnormalities.

contained a large amount of milky aseptic fluid (Figure 4). The prosthesis was well fixed. The patient underwent a revision to a ceramic THR with subsequent complete relief of symptoms. The patient remains asymptomatic 12 months after revision. At the position demonstrated by MR, a pale smooth lobulated 1.5 cm in diameter lesion was excised.

The excised specimen was examined with MR to confirm that this was the lesion identified by the pre-operative imaging. This confirmed the pre-operative signal characteristics of a very low  $T_2$  weighted signal with a surrounding bloom and an intermediate signal on  $T_1$  weighted imaging (Figure 5). A series of MR examinations were performed with a repetition time



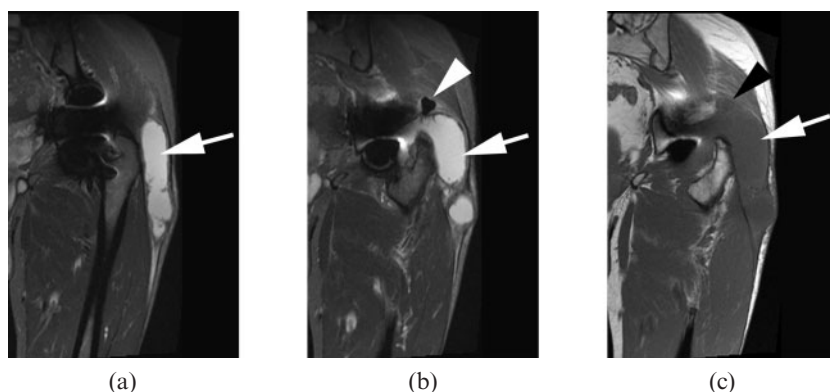
**Figure 2.** Ultrasound examination of the left hip demonstrating a coronal view of the buttock (superior is to the right of the image and inferior is to the left) and an anechoic fluid collection (arrow) containing thick echoic septations (arrowheads). The fluid collection extends medially (towards the bottom of the image) to the hip joint.

(TR) of 2220 ms and increasing echo time (TE) values from 8.8 ms, in intervals of 8.8 ms, up to 140.8 ms. Regions of interest were drawn for the lesion at each TE interval and the mean pixel value and standard deviation were recorded. The mean pixel value was plotted against the TE value to produce a decay curve which indicated that the  $T_2$  value of the tissue was in the order of 25 ms (Figure 6).

Histological assessment of the membrane lining the cavity demonstrated a sheet of necrotic fibrinous tissue. Deep to this necrotic layer was a more vascular layer with a predominantly lymphocytic perivascular infiltrate (Figure 7). The histology of the soft-tissue nodule revealed a central area of necrosis with a viable peripheral margin of lymphocytes and macrophages. Some of the macrophages contained particles of metal within their cytoplasm (Figure 8).

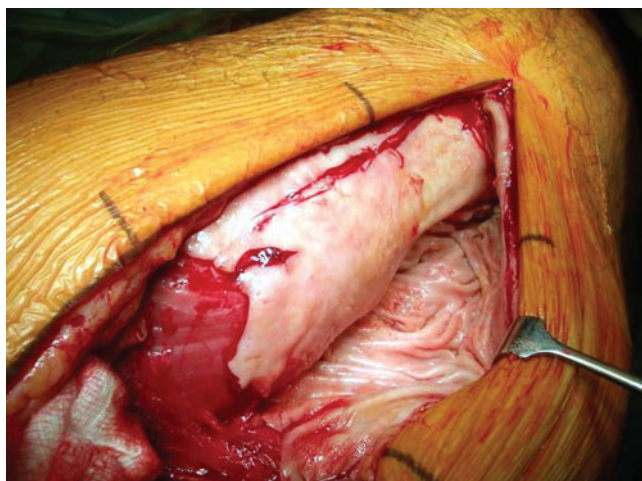
## Discussion

Metallosis has been previously described in metal-on-metal [8, 7], as well as in other, bearings [9]. To our knowledge, this is the first description of this mechanism of failure in a metal-on-metal resurfacing hip prosthesis. Metal-on-metal bearings have caused some concern over the past few years because it has been recognised that



**Figure 3.** (a,b) Coronal short tau inversion recovery (STIR) sequences and (c)  $T_1$  weighted fast spin echo demonstrate a fluid collection (arrows) that extends from the left buttock in to the left lateral thigh. A 1.5 cm lobulated lesion is demonstrated between the gluteus medius and minimus (arrowhead in (b,c)), which is conspicuous by its low signal bloom on the STIR image but difficult to appreciate on the  $T_1$  weighted image.

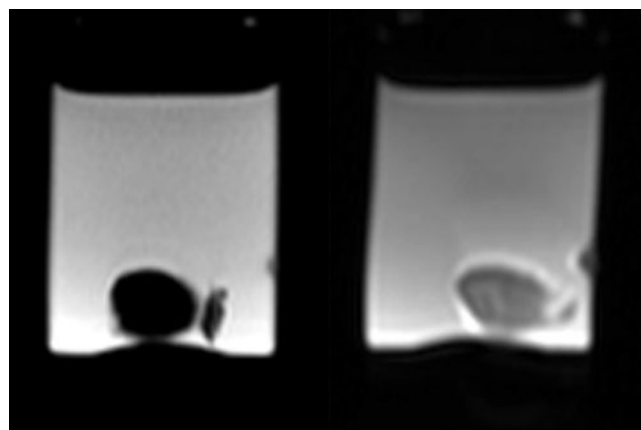




**Figure 4.** Intraoperative photograph demonstrating the exposed cavity lateral to the hip lined by a pale membrane comprising multiple folds.

some mechanisms of failure may be unique to second-generation metal-on-metal prostheses [10, 11]. Despite the increased machine tolerances of the second-generation metal-on-metal bearings, they still result in elevated ionic serum levels of cobalt and chromium [12–14]. Hypersensitivity to metals in the first-generation metal-on-metal bearings were recognised 30 years ago [15] and it has been suggested that the failures seen with the second-generation prostheses may have a similar aetiology [11], although differences in ion concentration and particle size of metal debris may modulate the presenting features. There have yet to be any published reports of similar mechanisms of failure in resurfacing prostheses.

However, it has been demonstrated that large metal-on-metal bearings result in higher levels of released metal ions than do smaller bearings [16], and that some of the larger bearings having a “running-in” effect, whereby increased levels of ions are released during the first post-operative year [17]. The possible local toxic effects have not been described in this type of prosthesis. In our experience of failures with one particular metal-

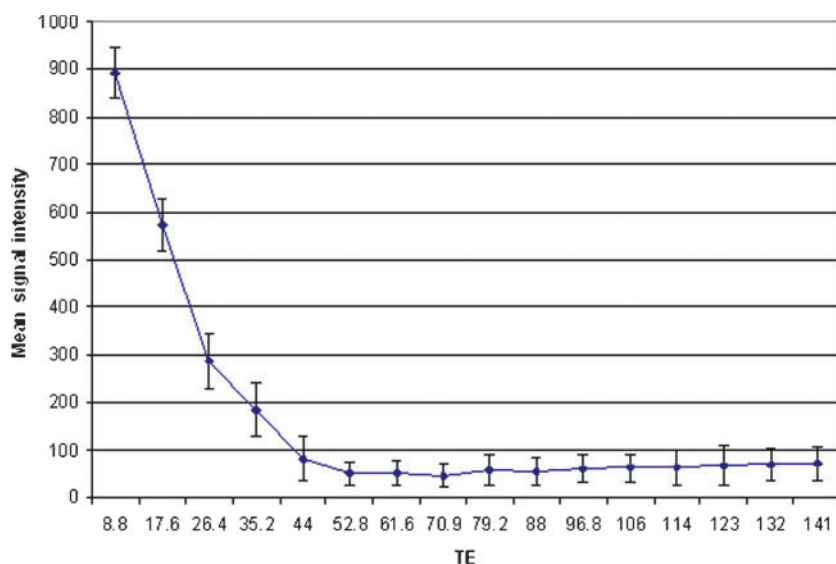


(a) (b)

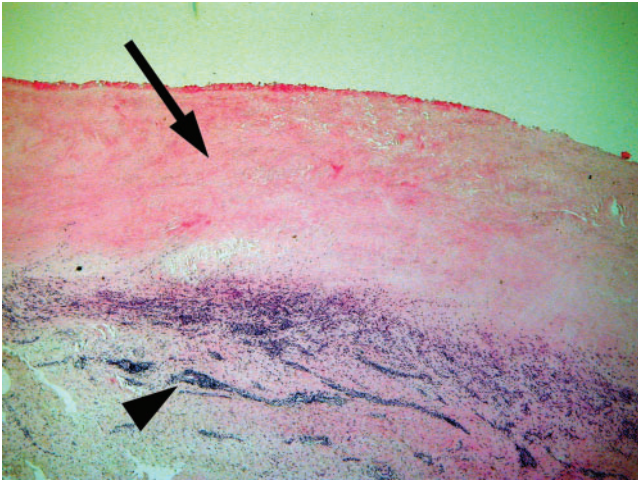
**Figure 5.** MRI of the excised nodule confirming similar MR characteristics with (a) very low signal on  $T_2$  weighted images and (b) intermediate signal on  $T_1$  weighted images.

on-metal total hip prosthesis, extensive soft-tissue necrosis, particularly of the gluteal muscles and tendons, and necrosis of the proximal femur are characteristic [18]. Aseptic ion-rich fluid-filled cavities about the failing prosthesis and it is assumed that the released ions cause local cell death either by direct toxicity or indirectly through a hypersensitivity reaction. The exact mechanism is still not understood. The MR appearances and the operative findings of the resurfacing arthroplasty in this case report are, in our experience, very similar to the findings in failing metal-on-metal THRs, suggesting that there may be a common cause for failure.

Other common causes for failure can be confidently excluded in this case. The post-operative course was unremarkable, the post-operative radiographs demonstrated a prosthesis in normal alignment, the patient was systemically well at presentation and there were no haematological or microbiological markers of infection. The extensive soft-tissue disease with normal radiography is, in our experience [18], typical of metal-on-metal soft-tissue reactions. Ultrasound is useful for identifying the presence of collections but is limited in fully evaluating the extent of disease. MRI, with metal artefact reduction



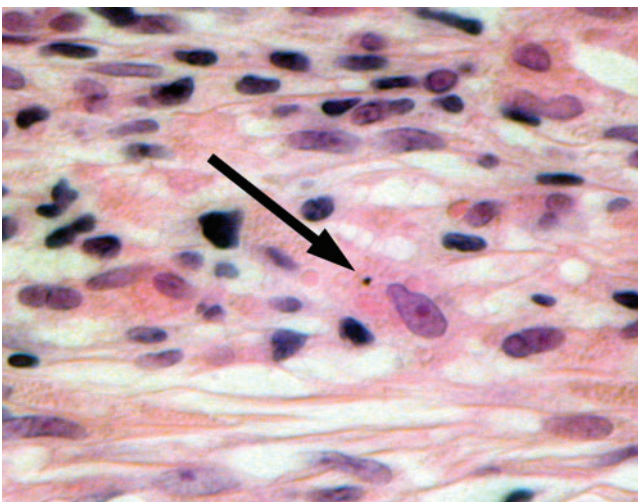
**Figure 6.** Graph demonstrating the mean signal characteristics of the excised nodule with increasing echo time (TE). By 50 ms, all of the signal noted on the  $T_1$  weighted sequences has been lost. MRI with a TE of longer than 50 ms (typical in  $T_2$  weighted sequences) gives rise to the very low signal noted in the nodule.



**Figure 7.** Low-power micrograph of a haematoxylin and eosin stained histological section through the periprosthetic soft-tissue biopsy. The tissue nearest to the fluid collection comprises necrotic fibrinous tissue (arrow). Deep to this lies a predominantly perivascular lymphocytic infiltrate (arrow-head).

sequences, is proving to be increasingly useful for assessing the periprosthetic soft tissues [18, 19]. The findings presented in this case — the large fluid collection with a low signal rim on  $T_2$  weighted images, with gluteal atrophy and avulsion — are also the most common abnormal findings on MRI in patients with early post-operative pain following metal-on-metal THR [18]. In this case, the diagnosis of lymphoreticular spread of metal debris was also suspected at pre-operative MRI.

Sinus histiocytosis in pelvic lymph nodes has been described in patients with a THR in which the histiocytes contain metallic particles of cobalt-chromium or titanium [20]. More distant spread of metallic debris to lymph nodes and organs, including the spleen and liver, has also been reported [21]. Iatrogenic labelling of the



**Figure 8.** High-power micrograph of a histological section (haematoxylin and eosin stained) through the periphery of the soft-tissue nodule, which comprised a centre of necrotic histiocytes with a margin of lymphocytes and macrophages. Some of the macrophages contained particles of metal within their cytoplasm (arrow).

lymphoreticular system is now an accepted method for enhancing the MRI of lymph nodes for cancer staging [22–24]. Iron particle preparations can be administered by a number of routes so that they are taken up by the lymphoreticular system. The iron particles become magnetised within the bore of the MRI machine, resulting in loss of signal, and therefore negative contrast enhancement in normal lymph node tissue. This case presents the first description of the same process occurring in histiocytes that have taken up metal particles, thus producing the characteristic signal loss on MRI. While large metal objects such as orthopaedic prostheses are obvious on MRI, the appearance of microscopic metallic particles is more subtle. They are characterised by signal loss that is ill-defined at the margins of the lesion and typically described as “blooming”. This occurs because metal particles become magnetised, forming their own magnetic dipoles which disrupt the local homogeneous magnetic field of the MR machine. However, signal loss is not necessarily present on all sequences. It is most obvious on sequences with long TEs such as  $T_2$  weighted images (Figure 5a). In sequences with short TEs, such as  $T_1$  and proton density weighted sequences, there may be no visible signal loss (Figure 5b), and therefore the lesion may not be appreciated. Multi-echo MR acquisitions can be helpful for demonstrating the characteristic fall in signal ( $T_2^*$  decay) caused by susceptibility artefacts from the metallic particles (Figure 6).

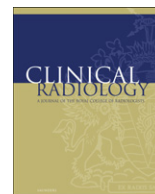
In summary, this is the first description of the radiological and histological appearances of a resurfaced hip arthroplasty failing because of metallosis. The periprosthetic soft-tissue changes, predominantly caused by necrosis, are similar to those described in some second-generation metal-on-metal THRs. Ultrasound and MR (with metal artefact reduction sequences) may demonstrate characteristic soft-tissue changes when plain radiographs are normal. The lymphoreticular spread of metal debris, indicating microscopic metallosis, can be diagnosed on MR sequences with a long TE by the presence of characteristic signal loss at the site of metal particle deposition.

## References

1. Daniel J, Pynsent PB, McMinn DJ. Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *J Bone Joint Surg Br* 2004;86:177–84.
2. Beaulé PE, Le DM, Campbell P, Dorey FJ, Park SH, Amstutz HC. Metal-on-metal surface arthroplasty with a cemented femoral component: a 7–10 year follow-up study. *J Arthroplasty* 2004;19:17–22.
3. Shetty VD, Villar RN. Development and problems of metal-on-metal hip arthroplasty. *Proc Inst Mech Eng [H]* 2006;220:371–7.
4. De Smet KA. Belgium experience with metal-on-metal surface arthroplasty. *Orthop Clin North Am* 2005;36:203–13, ix.
5. Treacy RB, McBryde CW, Pynsent PB. Birmingham hip resurfacing arthroplasty. A minimum follow-up of five years. *J Bone Joint Surg Br* 2005;87:167–70.
6. McMinn D, Treacy R, Lin K, Pynsent P. Metal on metal surface replacement of the hip. Experience of the McMinn prosthesis. *Clin Orthop Relat Res* 1996;:S89–98.
7. Davies AP, Willert HG, Campbell PA, Learmonth ID, Case CP. An unusual lymphocytic perivascular infiltration in

- tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am* 2005;87:18–27.
8. Boardman DR, Middleton FR, Kavanagh TG. A benign psoas mass following metal-on-metal resurfacing of the hip. *J Bone Joint Surg Br* 2006;88:402–4.
  9. Valenti JR, Del Rio J, Amillo S. Catastrophic wear in a metal-on-ceramic total hip arthroplasty. *J Arthroplasty* 2007;22:586–90.
  10. Park YS, Moon YW, Lim SJ, Yang JM, Ahn G, Choi YL. Early osteolysis following second-generation metal-on-metal hip replacement. *J Bone Joint Surg Am* 2005;87:1515–21.
  11. Willert HG, Buchhorn GH, Fayyazi A, Flury R, Windler M, Koster G, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am* 2005;87:28–36.
  12. Jacobs JJ, Skipor AK, Campbell PA, Hallab NJ, Urban RM, Amstutz HC. Can metal levels be used to monitor metal-on-metal hip arthroplasties? *J Arthroplasty* 2004;19:59–65.
  13. Brodner W, Bitzan P, Meisinger V, Kaider A, Gottsauner-Wolf F, Kotz R. Elevated serum cobalt with metal-on-metal articulating surfaces. *J Bone Joint Surg Br* 1997;79:316–21.
  14. Brodner W, Grubl A, Jankovsky R, Meisinger V, Lehr S, Gottsauner-Wolf F. Cup inclination and serum concentration of cobalt and chromium after metal-on-metal total hip arthroplasty. *J Arthroplasty* 2004;19:66–70.
  15. Jones DA, Lucas HK, O'Driscoll M, Price CH, Wibberley B. Cobalt toxicity after McKee hip arthroplasty. *J Bone Joint Surg Br* 1975;57:289–96.
  16. Clarke MT, Lee PT, Arora A, Villar RN. Levels of metal ions after small- and large-diameter metal-on-metal hip arthroplasty. *J Bone Joint Surg Br* 2003;85:913–7.
  17. Bowsher JG, Hussain A, Williams P, Nevelos J, Shelton JC. Effect of ion implantation on the tribology of metal-on-metal hip prostheses. *J Arthroplasty* 2004;19:107–11.
  18. Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donnell ST, et al. MR imaging of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008;63:49–58.
  19. White LM, Kim JK, Mehta M, Merchant N, Schweitzer ME, Morrison WB, et al. Complications of total hip arthroplasty: MR imaging-initial experience. *Radiology* 2000;215:254–62.
  20. Albores-Saavedra J, Vuitch F, Delgado R, Wiley E, Hagler H. Sinus histiocytosis of pelvic lymph nodes after hip replacement. A histiocytic proliferation induced by cobalt-chromium and titanium. *Am J Surg Pathol* 1994;18:83–90.
  21. Case CP, Langkamer VG, James C, Palmer MR, Kemp AJ, Heap PF, et al. Widespread dissemination of metal debris from implants. *J Bone Joint Surg Br* 1994;76:701–12.
  22. Saksena MA, Saokar A, Harisinghani MG. Lymphotropic nanoparticle enhanced MR imaging (LNMRI) technique for lymph node imaging. *Eur J Radiol* 2006;58:367–74.
  23. Misselwitz B. MR contrast agents in lymph node imaging. *Eur J Radiol* 2006;58:375–82.
  24. Harisinghani MG, Dixon WT, Saksena MA, Brachtel E, Blezek DJ, Dhawale PJ, et al. MR lymphangiography: imaging strategies to optimize the imaging of lymph nodes with ferumoxtran-10. *Radiographics* 2004;24:867–78.





## Original Paper

# MRI of asymptomatic patients with metal-on-metal and polyethylene-on-metal total hip arthroplasties

A. Mistry<sup>a</sup>, J. Cahir<sup>a</sup>, S.T. Donell<sup>b</sup>, J. Nolan<sup>b</sup>, A.P. Toms<sup>a,\*</sup>

<sup>a</sup> Department of Radiology, Norfolk & Norwich University Hospital, Norwich, Norfolk, UK

<sup>b</sup> Department of Orthopaedic Surgery, Norfolk & Norwich University Hospital, Norwich, Norfolk, UK

## ARTICLE INFORMATION

## Article history:

Received 15 October 2010

Received in revised form

31 January 2011

Accepted 2 February 2011

**AIMS:** To define and compare magnetic resonance imaging (MRI) findings of asymptomatic patients with metal-on-metal (MOM) and polyethylene-on-metal (POM) total hip replacements (THRs).

**MATERIALS AND METHODS:** Twenty-two THRs in 20 asymptomatic patients (seven men, 13 women, mean age 68 years, range 47–86 years) with normal hip radiographs were included in the study. These comprised 10 POM and 12 MOM bearings. Each patient underwent MRI with metal artefact reduction sequences (MARS) at a mean time of 46 months (POM) and 70 months (MOM) after surgery. Two musculoskeletal radiologists independently read each MRI examination for fluid collections, soft-tissue masses, muscle atrophy, and bone marrow signal changes.

**RESULTS:** A pre-MRI hip radiograph showed no significant differences from the post-operative radiograph regarding acetabular inclination, femoral stem angle, and stem mantle grade. There were eight periprosthetic collections (one POM, seven MOM). The majority of THRs had normal gluteal muscles. The ipsilateral piriformis and obturator internus muscles were more frequently abnormal in the MOM group. Overall, there were no significant differences in the number of abnormalities between the two types of bearings.

**CONCLUSION:** A range of MRI abnormalities are present in normal asymptomatic THRs but the increased frequency of these associated with MOM THR suggest that some of these changes might represent subclinical disease.

© 2011 The Royal College of Radiologists. Published by Elsevier Ltd. All rights reserved.

## Introduction

Conventional radiographs have been the mainstay of imaging symptomatic total hip replacements (THRs) for decades. Complications such as infection and loosening produce characteristic but overlapping appearances,<sup>1,2</sup> which are still difficult to differentiate with any imaging technique. MRI is increasingly being recognised as a useful tool for imaging complications of THR<sup>3–6</sup>. In particular,

metal artefact reduction sequences (MARS) have reduced the amount of susceptibility artefact and increased the conspicuity of peri-prosthetic soft tissues.<sup>7,8</sup>

In recent years there have been increasing reports in the literature of problems associated with the second generation metal on metal hip arthroplasty or metal on metal hip resurfacing<sup>9–11</sup> with MR imaging becoming key to the investigation of these patients. Patients present with early postoperative pain and despite normal plain radiographs and normal serum inflammatory markers, are found to have an unusual peri-prosthetic soft tissue collection or mass on MRI which have been described as acute lymphocytic vasculitis-associated lesions (ALVALs).<sup>12,13</sup> Histological analysis reveals a unique lymphocytic perivascular infiltration which may be the result of a hypersensitivity

\* Guarantor and correspondent: A.P. Toms, Department of Radiology, Norfolk & Norwich University Hospital, Colney Lane, Norwich, Norfolk NR4 7UY, UK. Tel.: +44 1603 286104; fax: +44 1603 286077.

E-mail address: [andoni.toms@nnuh.nhs.uk](mailto:andoni.toms@nnuh.nhs.uk) (A.P. Toms).



reaction.<sup>12</sup> Our experience, which has been previously published,<sup>9</sup> has found that nearly all symptomatic patients, with MOM arthroplasties, have at least one of a number of abnormalities including peri-prosthetic inflammatory masses, bone marrow oedema, muscle oedema and muscle atrophy and avulsion. Periprosthetic inflammatory masses are obviously abnormal findings but it could be argued that bone marrow oedema, muscle oedema and muscle atrophy might be part of the normal spectrum of MRI findings that occur after uncomplicated hip replacement. The aim of this study is to describe the MR appearances of THRs in asymptomatic patients with polyethylene on metal (POM) and metal on metal (MOM) arthroplasty.

## Materials and methods

This is a cross-sectional observational study of 20 patients (13 females, seven males) who had undergone a total of 22 total hip replacements; 10 THRs (in 10 patients) had a POM bearing (Exeter chromium steel alloy on polyethylene prosthesis; Stryker, Newbury UK) and 12 THRs (in 10 patients) had a MOM bearing (Ultima TPS cobalt–chromium on cobalt–chromium alloy, Depuy International Ltd, Leeds UK). Approval was given by the Research Governance and Ethics Committee and informed consent was obtained from all patients. The THRs had been performed by one of three surgeons using, according to their preference, either a posterior or anterolateral approach.

Patients were recruited to the study if they met the following clinical criteria: no symptoms of pain attributable to the affected hip, no hip pain on standing or walking, were not taking any analgesics for any reason, had no tenderness on palpation of and around the hip joint, and had full range of painless passive movement. Patients were excluded if the most recent radiograph of the hip prosthesis showed abnormal acetabular inclination, femoral stem alignment, and cement mantle grade C–D<sup>14</sup> and were also excluded if features of loosening or osteolysis were present. Acetabular inclination, femoral stem alignment, and cement mantle grade were recorded from the postoperative and pre-MRI radiograph of the pelvis. Measurements were made by a radiology registrar (A.M.) using a picture archiving and communication system (PACS; 2K) diagnostic workstation (Barco, Kortrijk, Belgium).

MRI was undertaken using a 1.5 T, Siemens Symphony Machine obtaining the following metal artefact reduction sequences (MARS): coronal T1-weighted (W) turbo spin-echo [echo time (TE) 23 ms, repetition time (TR) 669 ms] and short tau inversion recovery (STIR; TE 37 ms, TR 3840 ms), axial T1W (TE 23 ms, TR 534 ms) and T2W turbo spin-echo (TE 69 ms, TR 5600 ms) of the whole pelvis and a sagittal T2W turbo spin-echo of the hip (TE 69, TR 2900 ms). Using a section thickness of 5 mm, field of view of 340 × 340 mm, matrix size up to 448 × 336, and pixel bandwidth of 620 MHz.

Two experienced musculoskeletal radiologists (A.T., J.C.) independently reviewed the images. Radiological images were examined on a dedicated high-resolution PACS (2K)

diagnostic workstation (Barco). The MR images were examined for the following: periprosthetic soft-tissue collection (recording size and signal characteristics), presence of bone marrow oedema, and evaluation of the ipsilateral hip muscles. The gluteus muscles, piriformis and obturator internus were assessed for oedema (defined as presence of abnormal high T2 signal on STIR images), atrophy (defined as loss of volume and presence of fatty replacement), and tendon avulsion (defined as discontinuity of the low signal muscle attachment).

Statistical analysis using paired *t*-tests<sup>15,16</sup> comparing the postoperative to the most recent pre-MRI radiograph. Significance level was set at  $p < 0.05$ . Kappa statistic<sup>17</sup> was calculated to quantify the level of agreement between radiologists for the MRI observations.

## Results

### Demographics

A total of 22 THRs were inserted between 1998 to 2008 for degenerative osteoarthritis; 10 replacements with POM bearings and 12 with MOM bearings (Table 1). There were no significant differences in age between the asymptomatic POM and MOM groups, mean ages of 70.5 and 64.4 years, respectively. In the POM group there were eight female patients and five female patients in the MOM group (two female patients had bilateral MOM THRs).

The date of the THR operation was not available for one of the patients from the POM group and one from the MOM group. Therefore the time from surgery to the MRI examination was available for 20 of the 22 THRs. The mean time from operation to MRI for the POM group was 46 months (SD 33.2) and for the MOM group was 70 months (SD 13.8).

### Conventional radiographs

Acetabular inclination on the first postoperative radiograph measured between 27° and 56° (mean 42.5°, SD 8.9°). Twenty-one of the hips had a comparable radiograph prior to MRI, mean acetabular inclination of 43.8° (SD 9.8°). There was no statistically significant difference in acetabular inclination from the first postoperative radiograph to the most recent

**Table 1**

Demographics and characteristics of patients with polyethylene-on-metal (POM) and metal-on-metal (MOM) bearings for total hip replacement (THR).

	POM	MOM	<i>p</i> -Value
Mean age (range), years	70.5 (53–86)	64.4 (56–73)	0.17 <sup>a</sup>
Sex (F/M)	8/2	5/5	
Year of operation	1998–2008	1999–2003	
Side of THR (left/right)	5/5	7/5	
Surgical approach			
Posterior	8	9	
Anterolateral	0	2	
Information not available	2	1	
Mean time (SD) from THR to MRI, months	46 (33.2)	70 (13.8)	0.07 <sup>a</sup>

MRI, magnetic resonance imaging.

<sup>a</sup> Two-tailed paired Student's *t*-test.

**Table 2**

Summary of the radiographic position of the 22 hips in the study.

	First postoperative radiograph mean (SD)	Most recent pre-MRI radiograph mean (SD)	Range of differences	Test of differences <sup>a</sup>	
				POM	MOM
Acetabular inclination	42.5° (8.9)	43.8° (9.5)	0–9°	<i>p</i> = 0.16	<i>p</i> = 0.12
Stem position	2.0° varus (2.1)	2.2° varus (2.0)	0–4°	<i>p</i> = 0.14	<i>p</i> = 0.76
Cement mantle grade	A–B	A–B	–	–	–

POM, polyethylene-on-metal; MOM, metal-on-metal.

<sup>a</sup> Two tailed paired Student's *t*-test.

prior to MRI between the two types of THR bearings (*p* = 0.16 and *p* = 0.12, respectively). The anteroposterior (AP) femoral stem position ranged from 6° varus to 2° valgus, varying on the latest radiograph between 0° to 4°, which was not statistically different (*p* = 0.14 and *p* = 0.76). Cement mantle grade was classified A in 15 THRs and B in seven THRs. The cement mantle of one THR (POM bearing) changed from grade A to grade B. The cement mantle grade of the other bearings in the study did not change. Table 2 summaries the radiographic position of the THRs.

## MRI

The presence of an abnormal periprosthetic collection and its signal characteristics noted by the two readers are summarized in Table 3. Observer A noted one abnormal collection out of the 10 in the POM group and eight out of the 12 in the MOM group. Observer B also found one abnormal collection in the POM group and six collections in the MOM group. There was overall moderate agreement between the two observers for the presence and signal characteristics of an abnormal collection. The POM periprosthetic collection was a fluid-like, hyperintense signal on

T2W with a low signal rim. There were more abnormal collections observed in the MOM bearings compared with the POM bearings (seven compared with one). The MOM lesions were more heterogeneous in signal characteristics, most were fluid-like on T2W with low signal rims (Fig 1).

There was no agreement (Kappa = 0.01) between observers in the assessment of bone marrow oedema around the THR. Observer A noted bone marrow oedema in three of the 22 hips (one POM, two MOM) located in the ischial tuberosity, superior pubic ramus and posterior intertrochanter. Observer B noted bone marrow oedema in seven of the 22 hips (one POM, six MOM) all of the lesser trochanter.

There was good agreement between the observers (Kappa = 0.71) of the assessment of muscles around the THRs (Table 4). Both observers found intact and normal gluteus maximus muscle in all the POM and MOM THRs. The gluteus medius muscle was often found intact and normal for both groups (one or two hips noted atrophic in the POM group). Observer A recorded a normal gluteus minimus in all THRs. Observer B noted oedema or atrophy of the gluteus minimus in four THRs (two of each type of bearing; Fig 2). The piriformis and obturator internus muscle was observed to be atrophic in over half of both types of hip replacement (Fig 3). Table 5 summarizes the findings for both types of THR bearings.

**Table 3**

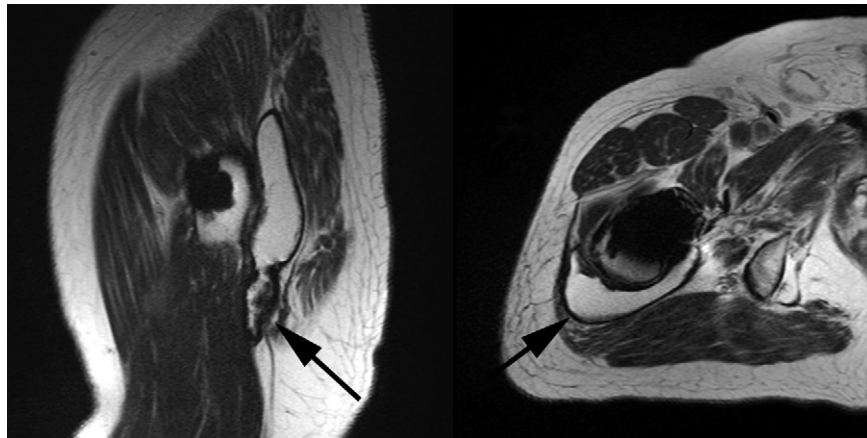
Summary of the findings for both total hip replacement (THR) groups.

	Observer A		Observer B		Kappa
	POM	MOM	POM	MOM	
Presence of periprosthetic collection	1	8	1	6	0.80 (0.55–1.06)
Signal intensity of abnormal soft-tissue collection:					0.54 (0.16–0.92)
T1					
High	0	1	0	1	
Isointense	1	7	0	3	
Low	0	0	1	2	
T2					
High	1	8	1	5	
Isointense	0	0	0	0	
Low	0	0	0	1	
Signal intensity of soft-tissue rim					0.81 (0.44–1.17)
T1					
High	0	0	0	0	
Isointense	1	5	1	2	
Low	0	3	0	4	
T2					
High	0	0	0	0	
Isointense	0	0	0	0	
Low	1	8	1	6	

POM, polyethylene-on-metal; MOM, metal-on-metal.

## Discussion

The results of this pilot study demonstrate that a range of soft-tissue changes can be found in asymptomatic patients, with both POM and MOM bearings, following THR. It also suggests that the incidence of soft tissue changes may be higher in patients with MOM articulations, although the difference between the two groups did not reach statistical significance. This is important because the Medicines and Healthcare Products Regulatory Agency (MHRA) have recently advised that all patients with MOM bearings should be screened for signs of ALVAL with either MRI or ultrasound.<sup>18</sup> In light of the findings of the study caution should be applied in over-interpreting some of the MRI findings in isolation in cases of suspected ALVAL. As this is a pilot study, the numbers are relatively small and, therefore, larger studies may demonstrate a more conclusive statistical difference between MRI findings in asymptomatic POM and MOM. Conversely, there is clearly an overlap in findings, which indicates that certain features will not be diagnostic of ALVAL on their own.



**Figure 1** Sagittal and axial T2W images of a right THR. The patient was asymptomatic with a MOM prosthesis. MRI with MARS demonstrates a collection of high T2W signal with a rim of low T2W signal intensity. The wall of this collection (arrows) was thick and irregular in areas.

To the authors' knowledge, this is the first prospective MRI study of asymptomatic THR that includes POM and MOM bearings. The results of the present study suggest that the MRI assessment of periprosthetic collections and of gluteal and short external rotator (SER) muscles was reliable with good or excellent measures of reproducibility.<sup>19</sup> Assessment of bone marrow oedema around the THR demonstrated very poor reliability with almost no

correlation between observers. The reason for this discrepancy between observers is not clear, but the most commonly reported site of bone marrow oedema was in the lesser trochanter, which is a difficult area to assess. Not only is the medial calcar a common site for disease in osteolysis, it also coincides with the transition from stem to neck of the femoral prosthesis, which is prone to hyperintense susceptibility artefact on T2W sequences (Fig 4).

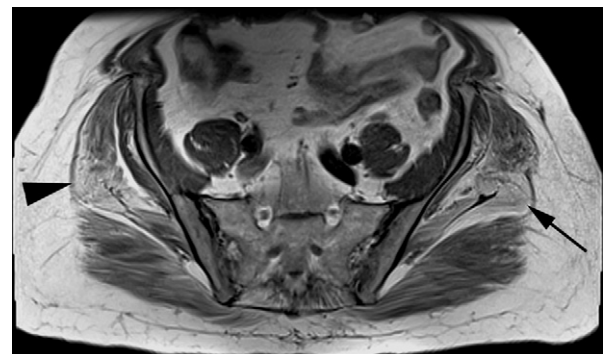
At least six of 12 THRs in the MOM group exhibited abnormal periprosthetic soft-tissue collections compared with one of 10 in the POM group. Although this at first glance appears to be a substantial difference, it does not reach statistical significance because the groups are small. The MOM periprosthetic collections were predominantly hyperintense fluid-like signal on T2W and intermediate or low on T1W and most were noted to have a thick or irregular wall as has been described previously in ALVAL.<sup>10</sup> The POM collection returned a fluid-like hyperintense signal on T2W and was noted to have a thin wall. It is difficult to be certain from the MRI signal characteristics alone whether all of these collections are secondary to reactive synovitis, particle disease, or ALVAL. Cooper et al.<sup>20</sup> have recently

**Table 4**

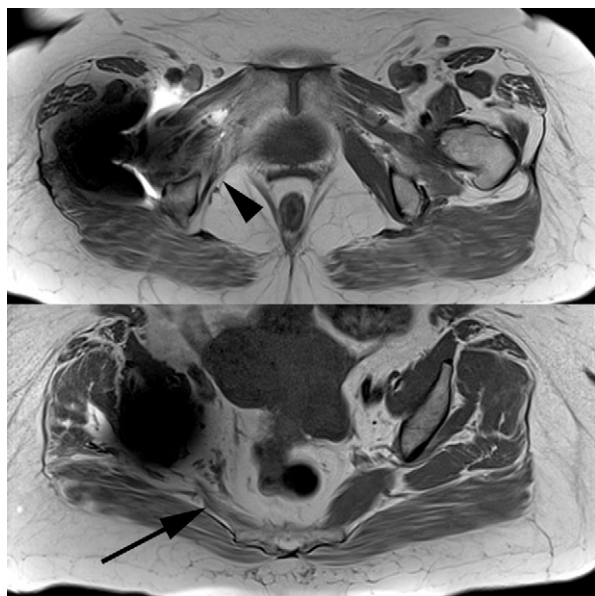
Summary of the observers findings of a periprosthetic collection and when present recording signal characteristics of the rim and main bulk.

	Observer A		Observer B		Kappa	
	POM	MOM	POM	MOM		
Gluteus maximus						
Normal	10	12	10	12	0.71(0.57–0.84)	
Oedema	0	0	0	0		
Atrophy	0	0	0	0		
Avulsion	0	0	0	0		
Gluteus medius						
Normal	9	12	8	12		
Oedema	0	0	0	0		
Atrophy	1	0	2	0		
Avulsion	0	0	0	0		
Gluteus minimus						
Normal	10	12	8	10		
Oedema	0	0	1	0		
Atrophy	0	0	1	2		
Avulsion	0	0	0	0		
Piriformis						
Normal	6	2	5	1		
Oedema	0	0	0	0		
Atrophy	4	10	5	11		
Avulsion	0	0	0	0		
Obturator Internus						
Normal	4	1	0	1		
Oedema	0	0	0	0		
Atrophy	6	11	10	9		
Avulsion	0	0	0	2		

POM, polyethylene-on-metal; MOM, metal-on-metal.



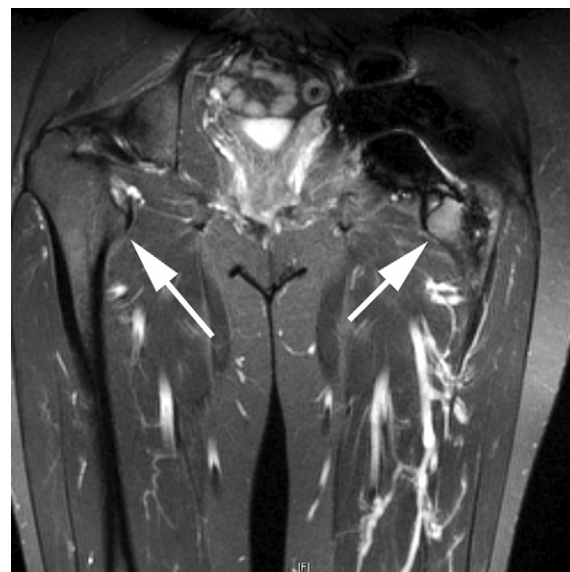
**Figure 2** Axial T2W MRI image demonstrating bilateral atrophy of the gluteus medius muscles (arrow and arrow head). The arrow head on the right is on the side of the MOM THR, the left hip (arrow) is osteoarthritic but has not been replaced and demonstrates the same pattern of atrophy.



**Figure 3** Axial T1W images of a patient with a right POM THR inserted using a posterior approach demonstrating muscle fatty atrophy of the right obturator internus muscle (arrow head) and right piriformis muscle (arrow).

published MRI data on a series of asymptomatic THRs using POM bearings, ceramic-on-ceramic, or ceramic-on-polyethylene for the presence of early reactive synovitis. Reactive synovitis (defined as intermediate to decreased signal intensity in the synovial lining with distension of the joint capsule) was found in 13 of 33 THRs. In contrast to the present study, two observers recorded MRI findings but the reliability between observers was not tested. Collections were reported to have an average volume of 2046 mm<sup>3</sup>, which are an order of magnitude less than those previously described in MOM bearings with ALVAL. Again without histological analysis, which is not ethically feasible in asymptomatic patients, the aetiology of these collections is not known. A prospective long-term study with regular interim MRI examinations of asymptomatic patients may provide an insight into whether a periprosthetic collection is subclinical disease that will later contribute to hip failure and subsequent revision or a true incidental finding.

In the present study, none of the patients from either group demonstrated atrophy of the gluteus maximus, and the majority demonstrated normal appearances of the gluteus



**Figure 4** Coronal STIR image of a patient with a left MOM THR. Although there is high marrow signal return from the left lesser trochanter, it is difficult to know whether this represents bone marrow oedema or is the result of susceptibility artefact.

medius and minimus muscles associated with an ipsilateral THR. Seventeen of 22 patients underwent THR with a posterior approach. This involves blunt dissection and splitting of gluteus maximus fibres and joint capsule exposed by incision close to the femoral attachment of the SER muscles. Once the prosthesis is inserted, closure is achieved by reattachment of the SER muscles to the medial concavity of the greater trochanter. Therefore, SER muscle atrophy may be expected with the posterior surgical approach. The predominant finding in most of the cases was atrophy of the SER muscles, particularly of the obturator internus. Two patients in the MOM group, who underwent an anterolateral surgical approach, were found to have a normal obturator internus and piriformis muscles. Pellicci et al.<sup>21</sup> investigated the posterior soft tissues after THR using MRI to assess the integrity of the SER muscles. At 3 months they found that 77% of hips (23/30) demonstrated mild or moderate atrophy of the piriformis muscle and 80% (24/30) had mild or moderate atrophy of the obturator internus muscle. The findings of the present study support the fact that SER atrophy is an incidental finding in uncomplicated THR where a posterior approach has been used.

This observational study is relatively small but provides useful information about the expected MRI findings in uncomplicated THR. The surgical approach and pre-existing osteoarthritis are likely causes for patterns of muscle atrophy, which should not necessarily be attributed to more important complications such as infection or ALVAL. Small simple fluid collections, probably caused by seromas or organized haematomas, may well also be normal, but infrequent findings. For the moment demonstrating bone marrow oedema around THRs on MRI does not appear to be a reliable finding.

In conclusion, the appearance of short external rotator muscle atrophy and small simple fluid collections are

**Table 5**

Summary of observer assessment of the ipsilateral gluteal and short external rotator muscles.

	POM (n = 10)	MOM (n = 12)
Periprosthetic collection	1	6/8
Bone marrow oedema	1	2/6
Abnormal		
Gluteus maximus	0	0
Gluteus medius	1/2	0
Gluteus minimus	0/2	0/2
Piriformis	4/5	10/11
Obturator internus	6/10	11

POM, polyethylene-on-metal; MOM, metal-on-metal.



probably normal findings of THR at MRI. The incidence of abnormal fluid collections in asymptomatic MOM THR may be higher than in POM THR and may represent subclinical disease.

## References

- Allen AM, Ward WG, Haygood TM. Imaging of the total hip arthroplasty. *J South Orthop Assoc* 1995;**4**:96–123.
- Törmänen J, Tervonen O, Koivula A, et al. Image technique optimization in MR imaging of a titanium alloy joint prosthesis. *J Magn Reson Imaging* 1996;**6**:805–11.
- Blumenfeld TJ, Bargar WL, Campbell PA. A painful metal-on-metal total hip arthroplasty a diagnostic dilemma. *J Arthroplasty* 2009; 4 Dec.
- Cahir JG, Toms AP, Marshall TJ, et al. CT and MRI of hip arthroplasty. *Clin Radiol* 2007;**62**:1163–71. discussion 1172–1173.
- Pfarrmann CWA, Notzli HP, Dora C, et al. Abductor tendons and muscles assessed at MR imaging after total hip arthroplasty in asymptomatic and symptomatic patients. *Radiology* 2005;**235**:969–76.
- Johnston C, Kerr J, Ford S, et al. MRI as a problem-solving tool in unexplained failed total hip replacement following conventional assessment. *Skeletal Radiol* 2007;**36**:955–61.
- Olsen RV, Munk PL, Lee MJ, et al. Metal artifact reduction sequence: early clinical applications. *RadioGraphics* 2000;**20**:699–712.
- Suh JS, Jeong EK, Shin KH, et al. Minimizing artifacts caused by metallic implants at MR imaging: experimental and clinical studies. *AJR Am J Roentgenol* 1998;**171**:1207–13.
- Toms AP, Marshall TJ, Cahir J, et al. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008;**63**:49–58.
- Toms AP, Nolan J, Barker T, et al. Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR. *Br J Radiol* 2009;**82**:e87–91.
- Fang CSJ, Harvie P, Gibbons CLMH, et al. The imaging spectrum of peri-articular inflammatory masses following metal-on-metal hip resurfacing. *Skeletal Radiol* 2008;**37**:715–22.
- Willert H, Buchhorn GH, Fayyazi A, et al. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg Am* 2005;**87**:28–36.
- Davies AP, Willert HG, Campbell PA, et al. An unusual lymphocytic perivascular infiltration in tissues around contemporary metal-on-metal joint replacements. *J Bone Joint Surg Am* 2005;**87**:18–27.
- Barrack RL, Mulroy RD, Harris WH. Improved cementing techniques and femoral component loosening in young patients with hip arthroplasty. A 12-year radiographic review. *J Bone Joint Surg Br* 1992;**74**:385–9.
- Fleiss J. *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley & Sons; 1981.
- Altman DG. *Practical statistics for medical research*. 1st ed. Oxford, UK: Chapman & Hall; 1990.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Measure* 1960;**20**:37–46.
- Medicines and Healthcare Products Regulatory Agency (MHRA) WM. Medical device alert: all metal-on-metal (MoM) hip replacements (MDA/2010/033). Available from: <http://www.mhra.gov.uk/Publications/Safetywarnings/MedicalDeviceAlerts/CON079157>. [Accessed 25 07 2010].
- Koch GG, Landis JR, Freeman JL, et al. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics* 1977;**33**:133–58.
- Cooper HJ, Ranawat AS, Potter HG, et al. Early reactive synovitis and osteolysis after total hip arthroplasty. *Clin Orthop Relat Res* 2010; 25 Apr.
- Pellicci PM, Potter HG, Foo LF, et al. MRI shows biologic restoration of posterior soft tissue repairs after THA. *Clin Orthop Relat Res* 2009;**467**:940–5.

# Silent soft tissue pathology is common with a modern metal-on-metal hip arthroplasty

## Early detection with routine metal artifact-reduction MRI scanning

Henry Wynn-Jones<sup>1</sup>, Rory Macnair<sup>2</sup>, James Wimhurst<sup>2</sup>, Nish Chirodian<sup>2</sup>, Brian Derbyshire<sup>1</sup>, Andoni Toms<sup>2</sup>, and John Cahir<sup>2</sup>

<sup>1</sup>The Centre for Hip Surgery, Wrightington Hospital, Lancashire; <sup>2</sup>Department of Trauma and Orthopaedics, Norfolk and Norwich University Hospital NHS Foundation Hospital, Norwich, Norfolk, UK

Correspondence: hwj@doctors.net.uk

Submitted 10-06-06. Accepted 10-12-16

**Background and purpose** Adverse reactions to metal debris have been reported to be a cause of pain in metal-on-metal hip arthroplasty. We assessed the incidence of both symptomatic and asymptomatic adverse reactions in a consecutive series of patients with a modern large-head metal-on-metal hip arthroplasty.

**Methods** We studied the early clinical results and results of routine metal artifact-reduction MRI screening in a series of 79 large-head metal-on-metal hip arthroplasties (ASR; DePuy, Leeds, UK) in 68 patients. 75 hips were MRI scanned at mean 31 (12–52) months after surgery.

**Results** 27 of 75 hips had MRI-detected metal debris-related abnormalities, of which 5 were mild, 18 moderate, and 4 severe. 8 of these hips have been revised, 6 of which were revised for an adverse reaction to metal debris, diagnosed preoperatively with MRI and confirmed histologically. The mean Oxford hip score (OHS) for the whole cohort was 21. It was mean 23 for patients with no MRI-based evidence of adverse reactions and 19 for those with adverse reactions detected by MRI. 6 of 12 patients with a best possible OHS of 12 had MRI-based evidence of an adverse reaction.

**Interpretation** We have found a high early revision rate with a modern, large-head metal-on-metal hip arthroplasty. MRI-detected adverse reactions to metal debris was common and often clinically “silent”. We recommend that patients with this implant should be closely followed up and undergo routine metal artifact-reduction MRI screening.

Metal-on-metal (MoM) total hip replacements have been used since the 1960s. Failure in early designs was attributed to mechanical loosening caused by poor bearing tolerances producing high friction (Amstutz and Grigoris 1996, Kothari

et al. 1996). Improved manufacturing and engineering techniques enabled the development of a new generation of MoM hip replacements. In the 1990s, the Birmingham Hip Resurfacing (BHR) was developed, and good early to medium-term results have been published (Daniel et al. 2004, Treacy et al. 2005, Heilpern et al. 2008). Similar implants, both resurfacings and large MoM bearings, coupled with standard femoral stems were subsequently developed and marketed by other manufacturers.

The development of magnetic resonance imaging (MRI) metal artifact reduction (MAR) sequences has enabled good visualization of the periprosthetic tissues (Toms et al. 2008), and been reported to be a clinically useful part of the assessment of painful MoM hip replacements (Hart et al. 2009). A number of authors have described the appearance of collections of fluid and inflammatory masses around painful MoM hip arthroplasties (Boardman et al. 2006, Pandit et al. 2008, Toms et al. 2008). These have been grouped under a variety of headings such as “aseptic lymphocyte-dominated vasculitis-associated lesions” (Willert et al. 2005), “pseudotumors” (Pandit et al. 2008), or “adverse reactions to metal debris (ARMD)” (Langton et al. 2010). Although these lesions have been previously described in patients investigated for pain, there have been no studies on the overall incidence of these lesions in an unselected series of patients, including those with no, or few, symptoms. It is not known whether these lesions may occur in the absence of symptoms.

At our institution, we have a policy of offering routine MAR MRI imaging to patients who have undergone MoM total hip replacement or resurfacing. We determined the early clinical outcome, revision rate, and incidence of ARMD using MAR MRI screening in a consecutive series of patients with an ASR THR or resurfacing (ASR; DePuy, Leeds, UK).

## Patients and methods

The ASR system was used at our institution between February 2005 and March 2008. This study is a report of the results of our standard follow-up and imaging protocol. 79 hip arthroplasties using ASR components were performed in 68 patients by 5 surgeons. 17 ASR resurfacing procedures were performed in 14 patients. 62 THRs were performed in 54 patients using an ASR acetabular component, a matched cobalt-chrome ASR XL head, and a Corail titanium hydroxyapatite-coated uncemented stem (DePuy, Leeds, UK). 14 head sizes were available, ranging from 39 mm to 63 mm in diameter in 2-mm increments. For the purposes of comparative analysis, we designated femoral head component sizes in the range 39–49 mm as “small”, and 51–63 mm as “large”.

The mean age of the 79 cases (56 males) at the time of surgery was 55 (30–76) years. The mean time from the primary procedure to last follow-up or revision was 32 (14–51) months. No patients had died or were lost to follow-up. Indications for surgery were primary osteoarthritis (OA) ( $n = 70$ ), OA secondary to dysplasia ( $n = 3$ ), post trauma ( $n = 2$ ), avascular necrosis ( $n = 2$ ), and OA secondary to Perthes' ( $n = 2$ ).

## Implants

The median size of the femoral head component was 49 (43–57) mm). For the resurfacing group, the median size was 51 (45–57) mm and for the THR group it was 49 (43–55) mm. 31 cases had a “large” femoral head (51–63 mm) and 48 cases a “small” head (38–49 mm).

## Follow-up assessments

The departmental policy at our institution is that all patients who have undergone a MoM hip replacement should remain under review and be assessed annually. The review involves clinical assessment, radiological assessment, and a patient-based self-assessment questionnaire. The questionnaire includes the Oxford hip score (Dawson et al. 1996) (OHS) (where 12 = best score and 60 = worst score), an assessment of the patients' satisfaction with the outcome of their hip replacement (possible responses: Yes, No, Uncertain), and rating the result of their hip replacement on a visual analog scale (VAS) from 0 (unsatisfactory) to 10 (perfect).

Patients who were scheduled for revision were asked to complete a questionnaire before to revision surgery. All patients at our hospital with a MoM hip replacement are also routinely invited to undergo an MRI scan, even if they are asymptomatic (provided there are no contraindications).

## Plain radiographs

Plain radiographs were assessed by one of the authors (HWJ) on a diagnostic PACS workstation. The acetabular implant orientation, leg length, offset, and femoral component alignment were measured. Two techniques were used for measurement of acetabular component orientation. Acetabular inclination

angle was measured manually (on the earliest postoperative, anteroposterior, supine pelvis radiograph of sufficient quality) with reference to the inter-teardrop line using tools on the PACS workstation. Acetabular component orientation was also measured using Wrightington cup orientation software. This enables measurement of inclination and version, and corrects for angular artifact due to the central X-ray beam offset from the hip (Derbyshire and Porter 2003, Derbyshire 2008). 2 of the authors (HWJ and BD) tested the inter-observer and intra-observer reliability of this software using standard statistical techniques (Bland and Altman 1983, Ranstam et al. 2000). Using these measurements, we designated acetabular components as being within or outside Lewinnek's “safe zone” (anteversion 5–25 degrees, and inclination 30–50 degrees) (Lewinnek et al. 1978).

Serial radiographs were compared to assess for periprosthetic osteolysis, lucent lines, bone loss, prosthesis migration, and soft tissue swelling. We noted osteolysis and radiolucent lines greater than 1 mm around the acetabular component in the zones of DeLee and Charnley (DeLee and Charnley 1976) as modified by Beaulé et al. (2004). Radiolucency around the femoral stem was recorded using the zones of Amstutz et al. (2004) for resurfacing arthroplasty and of Gruen for stemmed total hip arthroplasty.

## Metal artifact-reduction MRI

All MR examinations were performed on a 1.5T machine (Siemens Symphony; Siemens Healthcare, Erlangen, Germany) using sequences adapted for metal artifact suppression. All images were reviewed by two musculoskeletal radiologists (each had 5 years' experience in reporting MRI findings around MoM hip prostheses of various designs) and consensus findings were recorded.

Findings were categorized as: normal (Cahir et al. 2007), abnormal and typical of an adverse reaction to metal debris (Fang et al. 2008, Pandit et al. 2008, Toms et al. 2009), or abnormal but typical of a disease other than a metal-on-metal reaction—e.g. infection (Cahir et al. 2007). For those cases with characteristic findings of ARMD, they were further classified into mild, moderate, or severe disease (Figures 1–3). Mild changes constituted periprosthetic collections less than 5 cm in diameter, moderate comprised soft tissue masses of fluid collections greater than 5 cm in diameter, gluteal muscle atrophy or bone marrow edema and severe changes including extension through deep fascia, tendon avulsion, bone marrow replacement or fracture, or neurovascular involvement. This grading system has been shown to be reliable (Anderson et al. 2011).

## Histopathology

The tissue specimens in those patients who were revised, or underwent a biopsy, were assessed by a histopathologist experienced in evaluation of metal debris-related periprosthetic tissue reactions.



Figure 1. Mild adverse reaction to metal debris. Sagittal T2W MR through the femoral stem (S) of a Corail total hip replacement demonstrating mild periprosthetic disease. A small fluid-filled cavity (asterisk) surrounding the neck of the prosthesis is encapsulated by a thick, ragged low-signal rim (white arrow).

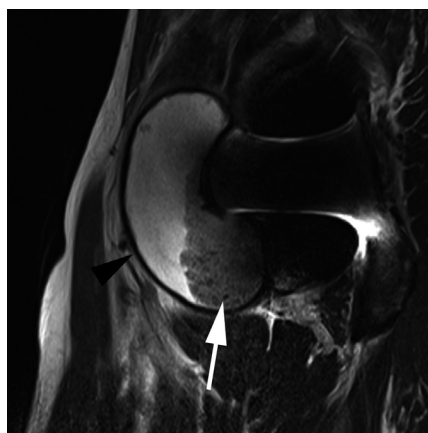


Figure 2. Moderate adverse reaction to metal debris. A sagittal T2W MR positioned just medial to the acetabular cup demonstrates moderate periprosthetic disease with a large cystic collection, demarcated by a low signal wall (black arrow), and filled with debris (white arrow) extending proximally in the line of the iliopsoas bursa. The relatively thick low signal wall and the debris are not typical of conventional iliopsoas bursae.

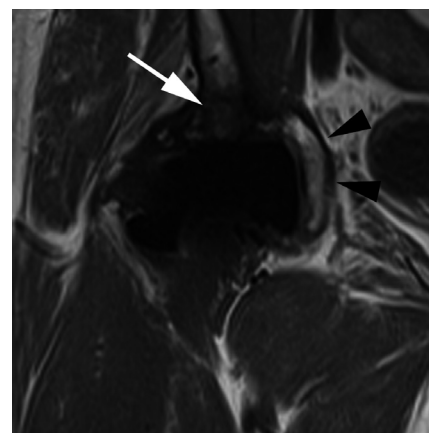


Figure 3. Severe adverse reaction to metal debris. Coronal T1W MR through the mid-coronal plane of the femoral head (black arrows indicate the medial wall of the acetabulum), demonstrating severe periprosthetic disease with bone marrow replacement in the acetabular roof (white arrow).

## Statistics

Statistical analysis was performed using StatsDirect statistical software version 2.7.7 (StatsDirect Ltd., Altrincham, UK). Dichotomous variables were analyzed using Fisher's exact test. Continuous parametric data were analyzed using unpaired t-tests and non-parametric data were assessed with the Mann-Whitney U-test.

## Results

### Plain radiographs

The mean cup inclination angle measured manually on digital plain radiographs was 50 (36–74) degrees. The acetabular orientation using Wrightington cup orientation software had a mean inclination of 50 (34–75) degrees and mean anteversion of 12 (2.3–39) degrees.

The intra- and inter-observer reliabilities of Wrightington cup orientation software for measuring ASR acetabular orientation were satisfactory. The intra-observer repeatability for version was  $\pm 0.55$  degrees, and it was  $\pm 0.49$  degrees for inclination. The inter-observer limits of agreement (95%) for version were  $-1.9$  to  $6.6$  degrees, and for inclination they were  $-2.8$  to  $2.4$  degrees. All 79 acetabular components appeared to be well fixed, with good bone ongrowth on the last follow-up radiograph. None of the acetabular components had osteolysis or radiolucent lines greater than 1 mm in any of the three Charnley DeLee zones.

16 of the 17 resurfacing femoral components had no evidence of loosening, migration, neck thinning, or radiolucent lines around the stem in any of Amstutz zones. One patient

had neck thinning, with resorption of the superior aspect of the femoral neck on the anteroposterior radiograph. There was no lysis, and no radiolucent lines around the stem of the femoral resurfacing component.

52 of the 62 Corail stems appeared well fixed on the latest radiographs, and had no radiolucent lines in any of the 7 Gruen zones. 10 hips had radiolucent lines in 1 or more Gruen zone. In 7 of these hips, a radiolucent line was seen only in Gruen zone 1. 3 hips also had lucent lines in Gruen zone 7. In all 10 hips, the Corail femoral component appeared well fixed from zones 2 to 6, and had not migrated.

### MAR MRI examinations

75 patients had MAR MRI examinations (59 THR, 16 resurfacing) at a mean of 31 (12–52) months after surgery. 4 hips were not scanned because 2 patients (3 hips) had a contraindication to MRI (1 pacemaker and 1 spinal cord stimulator) and 1 patient declined to be scanned as he was claustrophobic (Table 1). 42 MRI scans were classified as consistent with normal postoperative appearances (including seromas and atrophy of the short external rotators). 33 scans were considered to be abnormal, of which 3 were not thought to be typical of an adverse reaction to metal debris, including: infection ( $n = 1$ ), iliopsoas bursa ( $n = 1$ ), and osteolysis ( $n = 1$ ). 27 scans were considered to be abnormal and demonstrated features consistent with an adverse reaction to metal debris. 5 cases were considered to be mild, 18 were considered moderate, and 4 were classified as severe. The typical appearance was of a fluid signal collection extending from, and surrounding, the bearing that was demarcated by a very low-signal capsule, which was often ragged. Debris and a heterogeneous



Table 1. Summary of metal artifact-reduction (MAR) MRI findings

A	B	C	MAR MRI findings						
			D	E	F	G	H	I	J
ASR resurfacing	7	0	2	3	1	6	3	1	17
Corail with XL head and ASR cup	35	3	3	15	3	21	0	3	62
Total	42	3	5	18	4	27	3	4	79
A Prosthesis									
B Normal									
C Abnormal, not ARDM									
D Mild ARMD									
E Moderate ARMD									
F Severe ARMD									
G All ARDM									
H Unclassifiable									
I Not scanned									
J Total									

Table 2. Metal artifact-reduction MRI findings in relation to potential risk factors for metal debris-related reactions

	MRI classification		p-value
	Not ARMD (A and B)	ARMD (C1, C2, C3)	
Sex, F / M	13 / 35	9 / 18	0.4
Mean head size (mm)	50	49	0.2
Mean acetabular inclination (°)	50	51	0.4
Mean acetabular anteversion (°)	13	12	0.2
Head, small / large	26 / 22	19 / 8	0.2
Cup in Lewinnek's "safe zone", yes / no	24 / 21	10 / 16	0.2

signal were common findings within the fluid collections. The patients with severe disease included 3 cases with bone marrow replacement around either the acetabulum ( $n = 1$ ) or the proximal femur ( $n = 2$ ), and 1 patient had encasement of the sciatic nerve. The radiologists were not able to classify 3 of the MRI examinations and recommended follow-up with repeat imaging after a further 6 months. There were 2 cases of atrophy of the gluteus medius and minimus but no cases of gluteal avulsion. There were 7 cases of bone marrow edema in the proximal femur without any other abnormal findings. The significance of bone marrow edema in the proximal femur is unknown, but may be part of the spectrum of normal MAR MRI appearances in the absence of other changes. These cases were classified as normal postoperative appearances.

In patients with normal MAR MRI findings or abnormalities that were not an adverse reaction, the mean corrected cup inclination angle was 50 (34–75) degrees, the mean anteversion was 13 (2.3–33) degrees, and mean head size was 50 (43–57) mm. In those patients with an adverse reaction to metal debris, the mean cup inclination was 50 (37–60) degrees, mean anteversion was 12 (2.3–39) degrees, and mean head size was 49 (45–57) mm. The MAR MRI findings, and potential risk factors associated with ARMD, including sex, head size (large:

Table 3. Summary of revisions

A	B	C	D	E	F	G	H	I	J	K
1	51	F	P	34	30	47	58	20	+	+
2	59	M	P	40	40	49	49	24	+	+
3	60	M	S	34	15	51	60	13	+	+
4	55	F	P	38	46	51	56	13	+	–
5	57	F	P	32	14	51	57	3	(+)	+
6	67	F	P+S	23	22	45	53 <sup>a</sup>	<sup>a</sup>	+	+
7	60	M	P	16	57	49	56	11	–	–
8	55	M	P	29	55	47	53	7	+	+

<sup>a</sup> Only conventional radiograph available; cup inclination measured manually.

A Case

B Age

C Sex

D Indication for revision

P pain

S squeaking

E Time to revision (months)

F OHS

G Head size (mm)

H Cup inclination

I Cup version

J MRI findings

– normal

(+) mild

+ moderate

K Histology (+ ARMD)

> 50 mm, small: < 50 mm) and acetabular orientation (inclination, version, and location within or outside Lewinnek's "safe zone" (Lewinnek et al. 1978)) are summarized in Table 2. There was an increased risk of MRI-detected adverse reaction to metal debris with small femoral heads and cup orientation outside Lewinnek's "safe zone", but this increase was not statistically significant.

### Implant survival

At a mean follow-up of 32 (14–51) months, 8 revisions had been performed in 8 patients (4 female) (Table 3). The cumulative revision rate at 40 months with revision for any reason was 11% (95% CI: 4–18).

All the cases that were revised had an ASR acetabular component with an XL head and a Corail femoral component. 6 revisions were performed for pain (1 of these patients also reported squeaking). MRI confirmed an adverse reaction to metal debris before revision in 4 of these patients. 2 patients had minimal pain (1 had a squeaking hip), but screening MRI revealed changes consistent with a moderate adverse reaction to metal debris in one case and mild in the other case that was squeaking. Both patients elected to undergo revision. The plain radiographs were unremarkable, with no osteolysis in 7 of these patients. In 1 patient, there were radiolucencies with the appearance of a neocortex in Gruen zones 1 and 7. At the time of revision, the proximal stem was found to be loose with necrotic tissue, metal-stained debris, and fluid between the stem and the bone. The stem was well fixed distally. This

patient underwent revision of both components. In the remaining revisions, the femoral component was preserved and the acetabular component revised to an uncemented acetabular component with a polyethylene or ceramic liner. The XL heads were exchanged to appropriate ceramic heads to match the acetabular components.

### Histopathology

All 6 patients with a MAR MRI diagnosis of ARMD had histopathological findings (in the tissue taken at the time of revision) consistent with an ARMD. The findings were similar in the 6 cases: a fibrous capsular wall was seen, showing fibroid proliferation, with surface necrosis. A wide band of bland necrosis was seen. Perivascular lymphocytic infiltration was seen with macrophages or histiocytes containing small metal particles.

The histology in the 2 patients who were revised for pain, but with normal MRI scans, revealed normal fibrous tissue with no evidence of inflammation or adverse reaction to metal debris.

### Patient-related outcome

All hips were assessed with a hip questionnaire and an OHS at mean 32 (14–51) months after the primary procedure. The assessment scores were those at the latest follow-up or last assessment prior to revision.

66 patients were satisfied with their hip replacements whereas 7 were not, and 6 were doubtful. The mean overall “success” rating by patients of their hip replacements (on a VAS from 1 to 10) was 8.

The mean OHS in all patients—either at the latest follow-up or before revision—was 21. In the 8 patients who had a revision, the mean OHS before revision was 37. In patients without MRI-based evidence of an adverse reaction to metal debris, the mean OHS was 23, and in those with MRI-based evidence of ARMD it was 19 ( $p = 0.3$ ) (Table 4).

51 patients had an OHS at latest follow-up of 20 or less, and 18 of these patients had MRI-based evidence of an adverse reaction to metal debris. 26 patients had a “perfect” OHS of 12 at latest follow-up, and 6 of these had MRI-based evidence of ARMD.

### Discussion

We found MRI-detected metal debris-related abnormalities in one third of patients with a modern MoM bearing. Previous studies have concentrated on the MRI findings in patients investigated for painful prostheses (Boardman et al. 2006, Fang et al. 2008, Pandit et al. 2008, Toms et al. 2008).

One of our most concerning findings was that MRI-based evidence of an adverse reaction to metal debris does not appear to correlate with symptoms. In fact, some of the highest levels of satisfaction were in those patients with the

Table 4. Patient-related outcome in relation to MAR MRI findings

A	B	C	D	E	F	G
Normal	42	33	5	4	7.9	23
Abnormal (not ARMD)	3	3	0	0	8	25
Mild ARMD	5	5	0	0	9	18
Moderate ARMD	18	14	1	2	7.6	21
Severe ARMD	4	4	0	0	9.5	13
Unclassifiable	3	3	0	0	9	15
Not scanned	4	4	0	0	9.5	12
All	79	66	6	6	8.2	21

A MRI classification  
 B No. of cases  
 Patient satisfaction:  
 C Yes  
 D Doubtful  
 E No  
 F VAS (0–10): 0 = unsatisfactory; 10 = perfect.  
 G OHS: 12 = best; 60 = worst.

worst MAR MRI findings. One quarter of patients with a best possible OHS (12) had MRI-based evidence of ARMD. This suggests that even a policy of frequent clinical review would not detect patients developing soft tissue complications until extensive damage had occurred. It is unclear why there is often no pain.

A comparison can be made with the problem of silent osteolysis, which is well documented in patients with uncemented acetabular components with a polyethylene liner (Hozack et al. 1996, Utting et al. 2008). It is generally accepted that patients with such implants should be routinely assessed from plain radiographs—even in the absence of symptoms—in order to detect osteolysis before it becomes extensive. The difference with metal-on-metal related pathology is that soft tissue pathology is of particular concern, and this is not visible on a plain radiograph. We believe it is preferable to detect ARMD soft tissue damage and fluid-filled cavities at an early stage before the damage becomes extensive and irreversible. Grammatopolous et al. (2009) reported that resurfacing prostheses revised for pseudotumors have a poor outcome. This may well be because, in their series, patients only presented once they had become symptomatic and the disease had become extensive. Our experience with an earlier-generation 28-mm bearing MoM prosthesis, used in the 1990s, was that it functioned well for several years and then some patients suddenly presented with severe extensive soft tissue and bone necrosis, which was often undetectable on plain radiographs (Nolan et al. 2008).

The pattern of disease seen in our series on MRI shares similarities with those previously described for other prostheses, but there are also key differences. The pseudocysts in this group of patients commonly contained debris resulting in heterogeneous signal patterns (Figure 2), whereas those described with other prostheses were typically homogeneous fluid-filled cavities (Fang et al. 2008, Toms et al. 2008, 2009).

Gluteal myositis, atrophy, and avulsion have been described on MR with metal-on-metal-associated disease (Toms et al. 2008) but these were not common findings in our series of patients. This may be because MRI has been performed on asymptomatic patients and patients earlier in their postoperative course than previously described.

A number of factors, including female sex, small prosthetic head size, “poor” acetabular component orientation, and component design may contribute to ARMD. Hart et al. (2009) have shown that in a series of 16 failed large-head MoM prostheses, 13 were positioned outside the Lewinnek “safe zone”. We have found MRI-based evidence of MoM disease in 41% of prostheses with “small heads” (38–49 mm) and cup orientations outside Lewinnek’s “safe zone”. With “large heads” and a cup within the “safe zone”, the incidence of MoM disease was still one fifth. Lewinnek’s safe zone originally related to dislocation risk in metal-on-polyethylene hip replacements. There have been no prospective studies of acetabular component position to confirm whether there is actually a safe zone for prevention of ARMD. It is possible that all MoM prostheses, in any orientation, would develop a reaction.

Our cumulative revision rate at 40 months of 11% is much higher than that for a conventional THR. In our series, the overall revision rate for ARMD was 8%. Other authors have reported a high early revision rate with the ASR. Langton et al. (2010) reported poor early results with the ASR system, with a revision rate for symptomatic ARMD of 3% at 3 years. The revision rate in the subgroup of patients in their series with a total hip replacement (ASR cup, ASR XL head, and a Corail stem) rather than resurfacing was higher, at 6%. The Australian National Joint Replacement Registry 2009 report (AOA NJRR 2009) found that the number of revisions per 100 observed component years for the ASR was 2.3 as compared to 0.8 for the Birmingham Hip Resurfacing (BHR).

The design of the ASR acetabular component may be one of the reasons for the high failure rate. The cup comprises between 148 and 160 degrees of a sphere, whereas the BHR ranges from 158 to 166 degrees. This means that for any given cup position, more of the ASR head is uncovered. This may lead to increased edge loading and wear of the ASR cup. In a comparative study of the ASR and BHR, Langton et al. (2009) found that serum chromium and cobalt levels from ASR prostheses were more strongly influenced by the effect of the orientation of the acetabular component. There was an increase in metal ions at inclinations  $> 45^\circ$  and anteversion angles of  $< 10^\circ$  and  $> 20^\circ$  with the ASR, whereas these levels were only increased in the BHR group when the acetabular components were implanted at inclinations  $> 55^\circ$ .

We conclude that in our series of patients, the ASR Corail THR had a high rate of early revision due to MoM-related soft tissue problems. Furthermore, the incidence of MRI-detected MoM disease was high in both ASR Corail THRs and ASR resurfacings. Many patients with lesions revealed by MAR MRI were asymptomatic. We recommend that all patients

with this implant be carefully followed up on a regular basis. We believe that routine assessment of these implants should include soft tissue imaging.

HWJ: Main author, data collection, data analysis, radiographic analysis. RM: Coordinated imaging and MRI scans, data collection and radiographic analysis. JW: Clinical follow-up and manuscript preparation. NC: Clinical follow-up. BD: Radiographic analysis. AT: MRI scanning, interpretation and classification. JC: MRI scanning, interpretation and classification

No competing interests declared.

- Amstutz H C, Grigoris P. Metal on metal bearings in hip arthroplasty. *Clin Orthop* (Suppl) 1996; (329): S11–34.
- Amstutz H C, Beaulé P E, Dorey F J, Le Duff M J, Campbell P A, Gruen T A. Metal-on-metal hybrid surface arthroplasty: two to six-year follow-up study. *J Bone Joint Surg (Am)* 2004; 86 (1): 28–39.
- Anderson H, Toms A P, Cahir J G, Goodwin R W, Wimhurst J, Nolan J F. Grading the severity of soft tissue changes associated with metal-on-metal hip replacements: reliability of an MR grading system. *Skeletal Radiol* 2011; 40 (3): 303–7.
- Australian Orthopaedic Association National Joint Replacement Registry. Annual Report. Adelaide: AOA; (2009).
- Beaulé P E, Dorey F J, LeDuff M, Gruen T, Amstutz H C. Risk factors affecting outcome of metal-on-metal surface arthroplasty of the hip. *Clin Orthop* 2004; (418): 87–93.
- Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1983; (i): 307–10.
- Boardman D R, Middleton F R, Kavanagh T G. A benign psoas mass following metal-on-metal resurfacing of the hip. *J Bone Joint Surg (Br)* 2006; 88 (3): 402–4.
- Cahir J G, Toms A P, Marshall T J, Wimhurst J, Nolan J. CT and MRI of hip arthroplasty. *Clin Radiol* 2007; 62 (12): 1163–71; discussion 72–3.
- Daniel J, Pynsent P B, McMinn D J. Metal-on-metal resurfacing of the hip in patients under the age of 55 years with osteoarthritis. *J Bone Joint Surg (Br)* 2004; 86 (2): 177–84.
- Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg (Br)* 1996; 78 (2): 185–90.
- DeLee J G, Charnley J. Radiological demarcation of cemented sockets in total hip replacement. *Clin Orthop* 1976; (121): 20–32.
- Derbyshire B. Correction of acetabular cup orientation measurements for X-ray beam offset. *Med Eng Phys* 2008; 30 (9): 1119–26.
- Derbyshire B, Porter M L. An image analysis technique for the radiographic measurement of all-plastic acetabular cup orientation. *Proceedings of the 5th EBRA Meeting, Sheffield, Hip International* 2003.
- Fang C S, Harvie P, Gibbons C L, Whitwell D, Athanasou N A, Ostlere S. The imaging spectrum of peri-articular inflammatory masses following metal-on-metal hip resurfacing. *Skeletal Radiol* 2008; 37 (8): 715–22.
- Grammatopoulos G, Pandit H, Kwon Y M, Gundle R, McLardy-Smith P, Beard D J, Murray D W, Gill H S. Hip resurfacings revised for inflammatory pseudotumour have a poor outcome. *J Bone Joint Surg (Br)* 2009; 91 (8): 1019–24.
- Hart A J, Sabah S, Henckel J, Lewis A, Cobb J, Sampson B, Mitchell A, Skinner J A. The painful metal-on-metal hip resurfacing. *J Bone Joint Surg (Br)* 2009; 91 (6): 738–44.
- Heilpern G N, Shah N N, Fordyce M J. Birmingham hip resurfacing arthroplasty: a series of 110 consecutive hips with a minimum five-year clinical and radiological follow-up. *J Bone Joint Surg (Br)* 2008; 90 (9): 1137–42.

- Hozack W J, Mesa J J, Carey C, Rothman R H. Relationship between polyethylene wear, pelvic osteolysis, and clinical symptomatology in patients with cementless acetabular components. A framework for decision making. *J Arthroplasty* 1996; 11 (7): 769-72.
- Kothari M, Bartel D L, Booker J F. Surface geometry of retrieved McKee-Farrar total hip replacements. *Clin Orthop (Suppl)* 1996; (329): S141-7.
- Langton D J, Sprowson A P, Joyce T J, Reed M, Carluke I, Partington P, Nargol A V. Blood metal ion concentrations after hip resurfacing arthroplasty: a comparative study of articular surface replacement and Birmingham Hip Resurfacing arthroplasties. *J Bone Joint Surg (Br)* 2009; 91 (10): 1287-95.
- Langton D J, Jameson S S, Joyce T J, Hallab N J, Natsu S, Nargol A V. Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: A consequence of excess wear. *J Bone Joint Surg (Br)* 2010; 92 (1): 38-46.
- Lewinnek G E, Lewis J L, Tarr R, Compere C L, Zimmerman J R. Dislocations after total hip-replacement arthroplasties. *J Bone Joint Surg (Am)* 1978; 60 (2): 217-20.
- Nolan J F, Darrah C, Donell S T, Wimhurst J, Toms A, Marshall T, Barker T, Case C P, Peters C, Tucker J K. Metal on metal hip replacement. *J Bone Joint Surg (Br) (suppl III)* 2008; 90: 531.
- Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons C L, Ostlere S, Athanasou N, Gill H S, Murray D W. Pseudotumours associated with metal-on-metal hip resurfacings. *J Bone Joint Surg (Br)* 2008; 90 (7): 847-51.
- Ranstam J, Ryd L, Onsten I. Accurate accuracy assessment: review of the basic principles. *Acta Orthop Scand* 2000; 71 (1): 106-8.
- Toms A P, Marshall T J, Cahir J, Darrah C, Nolan J, Donell S T, Barker T, Tucker J K. MRI of early symptomatic metal-on-metal total hip arthroplasty: a retrospective review of radiological findings in 20 hips. *Clin Radiol* 2008; 63 (1): 49-58.
- Toms A P, Nolan J, Barker T, Darrah C, Malcolm P. Early failure of a Birmingham resurfacing hip replacement with lymphoreticular spread of metal debris: pre-operative diagnosis with MR. *Br J Radiol* 2009; 82 (977): e87-91.
- Treacy R B, McBryde C W, Pynsent P B. Birmingham hip resurfacing arthroplasty. A minimum follow-up of five years. *J Bone Joint Surg (Br)* 2005; 87 (2): 167-70.
- Utting M R, Raghuvanshi M, Amirfeyz R, Blom A W, Learmonth I D, Banister G C. The Harris-Galante porous-coated, hemispherical, polyethylene-lined acetabular component in patients under 50 years of age: a 12- to 16-year review. *J Bone Joint Surg (Br)* 2008; 90 (11): 1422-7.
- Willert H G, Buchhorn G H, Fayyazi A, Flury R, Windler M, Koster G, Lohmann C H. Metal-on-metal bearings and hypersensitivity in patients with artificial hip joints. A clinical and histomorphological study. *J Bone Joint Surg (Am)* 2005; 87 (1): 28-36.