THE INVESTIGATION AND MANAGEMENT OF METAL-ON-METAL HIP ARTHROPLASTY PATIENTS

A DISSERTATION SUBMITTED TO THE UNIVERSITY OF OXFORD



IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

GULRAJ SINGH MATHARU

WOLFSON COLLEGE

OXFORD

HILARY TERM 2017

ABSTRACT

THE INVESTIGATION AND MANAGEMENT OF METAL-ON-METAL HIP ARTHROPLASTY PATIENTS

Gulraj Singh Matharu **Doctor of Philosophy**
Wolfson College **Hilary Term 2017**

Many metal-on-metal hip arthroplasties (MoMHAs) were implanted worldwide in the form of hip resurfacing arthroplasty (HRA) and total hip arthroplasty. These devices experienced unexpectedly high short-term failure rates, with many revisions performed for adverse reactions to metal debris (ARMD) and pseudotumour (a specific type of ARMD). To identify problems early, worldwide regulatory authorities recommend most MoMHA patients require regular follow-up. However these recommendations are not evidence-based and are very costly. The broad aims of this thesis were to determine the clinical impact of pseudotumour, and to develop evidence-based guidance for the investigation and management of MoMHA patients.

A prospective single centre cohort of 1,429 HRAs identified pseudotumour as the commonest indication for revision surgery. The overall 15 year pseudotumour revision rate was high at 14.0%. Young females had the highest risk of pseudotumour revision, whilst males with Birmingham Hip Resurfacing (BHR) implants had the lowest risk.

Two large prospective cohort studies investigated the role of blood metal ions during follow-up. Patients with BHR implants (n=494) who had blood metal ions below newly devised thresholds were at low risk of ARMD (cobalt: 2.15 μg/l for unilateral BHRs; maximum cobalt or chromium: 5.52 μg/l for bilateral BHRs). Current fixed regulatory authority thresholds missed more patients with ARMD compared with the new thresholds. These findings were subsequently validated in 416 BHR patients at two external European centres.

Three observational studies investigated the role of hip imaging during follow-up. Radiographic factors predictive of HRAs with a pseudotumour included acetabular component malposition, osteolysis, and acetabular loosening. A retrospective diagnostic accuracy study demonstrated that ultrasound and magnetic resonance imaging (MRI) both had a role prior to HRA revision surgery. However combined ultrasound and MRI was most effective for both identifying and excluding intra-operative pseudotumours. Thus combined imaging may be useful in complex cases. In a prospective longitudinal cohort of 140 asymptomatic HRAs, 11% developed new pseudotumours within five years of initial assessment. No patients with a normal initial ultrasound and low blood metal ions (<2 μg/l) developed new pseudotumours (36% of hips). These patients do not need regular follow-up.

Two large retrospective observational studies were performed involving MoMHA revision procedures recorded in the National Joint Registry for England and Wales. Contrary to previous observations, patients revised for ARMD had half the risk of re-revision and death compared with matched patients undergoing non-ARMD revision surgery (such as for fracture, loosening, and infection). Modifiable predictors of re-revision surgery following ARMD revision were modular component only revisions, and using ceramic-on-ceramic revision bearings.

The clinical impact of pseudotumour following HRA is significant and has continued to increase into the second decade. This thesis has developed an evidence-based approach for the investigation and management of MoMHA patients. No single investigation can universally be used during patient surveillance. However important information has now been provided for risk stratifying patients for surveillance, interpreting blood metal ions, clarifying the roles of imaging, and providing parameters to safely exclude many asymptomatic patients from regular surveillance. The outcomes following ARMD revision surgery have improved since the threshold for performing revision was lowered. However outcomes were influenced by modifiable factors (type of revision procedure and bearing), suggesting that surgeons could reduce failure rates further following ARMD revision procedures.

This thesis is dedicated to my wife, Vicky, and my son, Joshan.

For all the love and support they have given me during my academic journey.

ACKNOWLEDGEMENTS

This academic journey would not have been possible without the overwhelming help and support that I have received from so many people, for which I am eternally grateful. Firstly, I would like to thank my three supervisors, Professor David Murray, Professor Hemant Pandit, and Associate Professor Andrew Judge. I thank them all immensely for their patience with me as well as their invaluable guidance, advice, and support during my time in Oxford. I am deeply indebted to them all for being excellent supervisors and for teaching me so much more than what is in this thesis.

This thesis would not have been possible without the daily advice and generous support provided by Barbara Marks and Jo Brown, with no task ever being too large to deal with. I have consistently been impressed by their organisational skills, and timely management of any problems that I have presented them with. Jo has kept a meticulous clinical hip database that was heavily relied upon in this thesis, for which I am extremely grateful.

I wish to give thanks to my seniors in Birmingham who I have continued to work closely with, namely Professor Paul Pynsent, Mr Ronan Treacy, Mr David Dunlop, and Dr Fiona Berryman. They have continued to provide me with excellent guidance and support for all of the collaborative work we have undertaken together during this thesis.

I must also thank all of the colleagues I met when I first arrived in Oxford as well as those I have met along my academic journey, who have now all become good friends. There are too many people to mention individually, but I would particularly like to thank Stephen Mellon, Elise Pegg, and Thomas Hamilton. They all introduced me to the department when I arrived back in August 2014 and would happily spend their own time answering all of my research and non-research related questions. For this I am truly appreciative. I wish to also thank Jenny Ferris, Louise Strickland, Cathy Jenkins, and Afsie Sabokbar for all their assistance and support during my time in Oxford.

Great thanks must also go to all of the funders who have kindly supported my research, including The Royal College of Surgeons of England, Arthritis Research UK, The Royal Orthopaedic Hospital Hip Research and Education Charitable Fund, and The Orthopaedics Trust.

Finally, I would like to thank all of the patients who have participated in the research contained within this thesis and without whom none of this work would have been possible.

ABBREVIATIONS

ALTR = adverse local tissue reaction

ALVAL = aseptic lymphocytic vasculitis associated lesion

ARMD = adverse reactions to metal debris

ASA = American Society of Anesthesiologists

ASR = Articular Surface Replacement

AUC = area under the curve

BHR = Birmingham Hip Resurfacing

BMI = body mass index

CI = confidence interval

Co = cobalt

CoC = ceramic-on-ceramic

CoM = ceramic-on-metal

CoP = ceramic-on-polyethylene

Cr = chromium

CRP = C-reactive protein

CT = computed tomography

DOF = depth of field

EBRA = Einsel-Bild-Roentgen analysis

EDTA = ethylenediaminetetraacetic acid

EPR = electronic patient record

ESR = erythrocyte sedimentation rate

F = female

FDA = Food and Drug Administration

FOV = field of view

FSE = fast spin echo

HHS = Harris Hip Score

HO = heterotopic ossification

HOOS = hip disability and osteoarthritis outcome score

HR = hazard ratio

HRA = hip resurfacing arthroplasty

IQR = interquartile range

LMWH = low molecular weight heparin

M = male

MARS-MRI = metal artefact reduction sequence magnetic resonance imaging

MDS = minimum dataset

MHRA = Medical and Healthcare products Regulatory Agency

MI = multiple imputation

MoC = metal-on-ceramic

MoM = metal-on-metal

MoMHA = metal-on-metal hip arthroplasty

MoP = metal-on-polyethylene

MRI = magnetic resonance imaging

NA = not applicable

NHS = National Health Service

NICE = National Institute for Health and Care Excellence

NJR = National Joint Registry

NOC = Nuffield Orthopaedic Centre

NPV = negative predictive value

NS = not stated

OA = osteoarthritis

ODEP = Orthopaedic Data Evaluation Panel

OHS = Oxford Hip Score

ONS = Office for National Statistics

OR = odds ratio

PACS = picture archiving and communication system

PPV = positive predictive value

Ref = reference group

ROC = receiver operating characteristic

ROH = Royal Orthopaedic Hospital

RSC = Research Sub-committee

SD = standard deviation

SHR = sub-hazard ratio

STIR = short tau inversion recovery

T1W = T1 weighted image

T2W = T2 weighted image

TE = time to echo

TGA = Therapeutic Goods Administration

THA = total hip arthroplasty

TLO = tertiary lymphoid organ

TR = time to resonate

UCLA = University of California, Los Angeles Activity Score

UK = United Kingdom

US = United States

VTE = venous thrombo-embolism prophylaxis

WOMAC = Western Ontario and McMaster Universities Arthritis Index

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1. Overview 1

1.2. Hip arthroplasty terminology and cohorts assessed 4

1.3. Broad aims 5

1.4. Specific objectives 5

1.5. Structure of the thesis 6

1.6. Data conventions 7

1.7. Declaration 7

2. LITERATURE REVIEW 9

2.1. Metal-on-metal hip arthroplasty 9

2.1.1. Rationale, development, and usage 9

2.1.2. Implant revision rates for all-causes 12

2.1.3. Modes of implant failure 14

2.2. ARMD and pseudotumour 16

2.2.1. Terminology 16

2.2.2. Aetiology and pathogenesis 18

2.2.3. Clinical presentation 21

2.2.3.1. Local presentation 21

2.2.3.2. Systemic presentation 22

2.2.4. Differential diagnosis 23

2.2.5. Revision rates for ARMD and pseudotumour 24

2.2.6. Risk factors for ARMD and pseudotumour revision 25

2.2.6.1. Patient risk factors 26

2.2.6.2. Implant risk factors 27
2.2.6.3. Surgical risk factors 28
2.3. Investigative modalities 29
2.3.1. Blood metal ion sampling 29
2.3.1.1. Rationale, methods, and interpretation 29
2.3.1.2. Blood metal ions in healthy controls and well-functioning hips 30
2.3.1.3. Factors influencing blood metal ion concentrations 30
2.3.1.4. Blood metal ion thresholds for clinical use 31
2.3.1.5. Repeat blood metal ion sampling 33
2.3.2. Radiographs 33
2.3.3. Cross-sectional imaging 35
2.3.3.1. Ultrasound 35
2.3.3.2. MARS-MRI 37
2.3.3.3. Ultrasound vs. MARS-MRI 41
2.3.3.4. Computed tomography 43
2.3.4. Other investigations 43
2.4. Revision surgery for ARMD 45
2.4.1. Threshold for revision surgery 45
2.4.2. Intra-operative findings and surgical strategy 46
2.4.3. Outcomes following ARMD revision surgery 48
2.4.3.1. Evidence available and patient demographics 48
2.4.3.2. Complications and re-revision surgery 52
2.4.3.3. Functional outcomes 56
2.4.3.4. Predicting re-revision and guiding surgical decisions 56
2.5. Regulatory authority guidance for patient surveillance 57

2.5.1. Variation in regulatory authority guidance 57
2.5.1.1. Type of metal-on-metal hip arthroplasty 61
2.5.1.2. Follow-up regularity 61
2.5.1.3. Investigative modalities 62
2.5.1.4. Blood metal ion thresholds 63
2.5.1.5. Threshold for revision surgery 63
2.5.2. Regulatory authority guidance versus available evidence 63
2.5.2.1. Do all patients require follow-up? 64
2.5.2.2. Should symptoms decide follow-up? 64
2.5.2.3. What is the role of blood metal ion sampling? 65
2.5.2.4. What is the role of cross-sectional imaging? 65
2.5.2.5. When should follow-up be repeated? 66
2.5.2.6. What is the threshold for revision surgery? 66
2.5.3. Financial analysis of regulatory authority guidance 67
2.5.4. Conclusions regarding currently proposed surveillance 68
2.6. Summary 70
3. CLINICAL IMPACT OF PSEUDOTUMOURS 72
3.1. Introduction 72
3.2. Objectives 73
3.3. Patients and methods 73
3.3.1. Study design 73
3.3.2. Data source 74
3.3.3. Surgery performed 74
3.3.4. Patient follow-up 75
3.3.4.1. Cross-sectional imaging 76

3.3.4.2. Blood metal ion sampling 77
3.3.5. Outcomes 77
3.3.6. Statistical analysis 78
3.4. Results 79
3.4.1. Outcomes 79
3.4.2. 15 year revision rates 81
3.4.3. Predicting pseudotumour revision: whole cohort 82
3.4.4. Predicting pseudotumour revision: stratified by gender 85
3.4.5. Effect of year of primary surgery 86
3.5. Study strengths and limitations 88
3.6. Discussion 89
3.6.1. 15 year revision rates 89
3.6.2. Predicting pseudotumour revision 92
3.6.3. Clinical implications of findings 95
3.7. Summary 97
4. INVESTIGATION – BLOOD METAL IONS 99
4.1. Introduction 99
4.2. Objectives 100
4.3. Patients and methods 101
4.3.1. Study design 101
4.3.2. Data source 102
4.3.2.1. Derivation cohort 102
4.3.2.2. Validation cohort 103
4.3.3. Surgery performed 104
4.3.4. Patient follow-up 105

4.3.4.1. Cross-sectional imaging....108
4.3.4.2. Blood metal ion sampling....108
4.3.5. Study inclusion and exclusion criteria....109
4.3.6. Outcomes....112
4.3.7. Statistical analysis....113
4.4. Results....115
4.4.1. Patient demographics and outcomes....115
4.4.2. Blood metal ion concentrations....116
4.4.3. Derivation cohort: formulating unilateral BHR thresholds....119
4.4.4. Derivation cohort: formulating bilateral BHR thresholds....121
4.4.4.1. Sensitivity analysis: bilateral BHRs with cross-sectional imaging....123
4.4.5. Derivation cohort: BHR thresholds vs. regulatory authority thresholds....123
4.4.6. Validation cohort: applying unilateral and bilateral BHR thresholds....125
4.4.6.1. Sensitivity analysis: validation cohort with cross-sectional imaging....126
4.4.7. Validation cohort: BHR thresholds vs. regulatory authority thresholds....126
4.5. Study strengths and limitations....127
4.6. Discussion....129
4.6.1. Deriving blood metal ion thresholds....129
4.6.2. Validating blood metal ion thresholds....132
4.6.3. Clinical implications of findings....133
4.7. Summary....137
5. INVESTIGATION – HIP IMAGING....139
5.1. Introduction....139
5.2. Objectives....141
5.3. Patients and methods: radiographic study....142

5.3.1. Study design 142
5.3.2. Pseudotumour group 142
5.3.3. Non-pseudotumour group 143
5.3.4. Radiographic assessments 143
5.3.4.1. Radiographic abnormalities 144
5.3.4.2. Measuring acetabular component position 145
5.3.4.3. Measuring femoral neck narrowing 146
5.3.4.4. Observer assessments 147
5.3.5. Statistical analysis 148
5.4. Results: radiographic study 148
5.4.1. Patient factors 148
5.4.2. Radiographic factors 149
5.4.3. Logistic regression modelling 152
5.4.4. Internal validation of the final multivariable model 154
5.5. Study strengths and limitations: radiographic study 155
5.6. Patients and methods: ultrasound vs. MRI study 156
5.6.1. Study design 156
5.6.2. Study inclusion and exclusion criteria 157
5.6.3. Cross-sectional imaging performed 158
5.6.4. Assessment of cross-sectional imaging 159
5.6.5. Assessment of revision procedures 160
5.6.6. Statistical analysis 160
5.7. Results: ultrasound vs. MRI study 161
5.7.1. Intra-operative findings 161
5.7.2. Cross-sectional imaging findings 161

5.7.3. Diagnostic test characteristics and kappa agreement: imaging vs. intra-operative findings....166
5.8. Study strengths and limitations: ultrasound vs. MRI study....167
5.9. Patients and methods: new pseudotumour study....169
5.9.1. Study design....169
5.9.2. Description of initial cohort and assessment....169
5.9.3. Description of cohort undergoing repeat assessment....170
5.9.3.1. Cohort assessed....170
5.9.3.2. Ultrasound assessment and classification....173
5.9.3.3. Outcomes....173
5.9.4. Cohort not undergoing repeat assessment....174
5.9.5. Statistical analysis....174
5.10. Results: new pseudotumour study....175
5.10.1. Changes between repeat imaging....175
5.10.2. Development of new pseudotumours....176
5.10.3. Diagnostic test characteristics for clinical scenarios....177
5.11. Study strengths and limitations: new pseudotumour study....178
5.12. Discussion....180
5.12.1. The role of radiographs....180
5.12.2. The role of cross-sectional imaging....182
5.12.3. The role of repeat cross-sectional imaging....185
5.12.4. Clinical implications of findings....186
5.13. Summary....190
6. MANAGEMENT – REVISION SURGERY....193
6.1. Introduction....193

6.2. Objectives 194
6.3. Patients and methods: ARMD outcome study 195
6.3.1. Study design 195
6.3.2. Data source 195
6.3.2.1. Background 195
6.3.2.2. Data collection 196
6.3.2.3. Data extract 198
6.3.2.4. Data cleaning 198
6.3.3. Study inclusion and exclusion criteria 201
6.3.4. Exposure variables 202
6.3.5. Outcomes 203
6.3.6. Statistical analysis 204
6.3.6.1. Sensitivity analysis: missing data 205
6.4. Results: ARMD outcome study 206
6.4.1. Intra-operative complications 206
6.4.2. Mortality 206
6.4.3. Re-revision surgery 209
6.4.4. Predictors of re-revision surgery 211
6.4.4.1. Univariable models 211
6.4.4.2. Final multivariable model 213
6.4.4.3. Analysis of multilevel predictors in the final multivariable model 214
6.4.4.4. Sensitivity analysis 215
6.5. Study strengths and limitations: ARMD outcome study 216
6.6. Patients and methods: ARMD vs. non-ARMD outcome study 218
6.6.1. Study design and data source 218

6.6.2. Study inclusion and exclusion criteria 218
6.6.3. Exposure variables 219
6.6.4. Outcomes 220
6.6.5. Statistical analysis 220
6.7. Results: ARMD vs. non-ARMD outcome study 222
6.7.1. Propensity score matching 222
6.7.2. Intra-operative complications 226
6.7.3. Mortality 227
6.7.4. Re-revision surgery 228
6.7.5. Sensitivity analysis 231
6.8. Study strengths and limitations: ARMD vs. non-ARMD outcome study 231
6.9. Discussion 233
6.9.1. Outcomes following MoMHA revision surgery 233
6.9.1.1. Intra-operative complications 233
6.9.1.2. Mortality 233
6.9.1.3. Re-revision surgery: indications for failure 235
6.9.1.4. Re-revision rates following ARMD vs. non-ARMD revision 236
6.9.2. Predicting re-revision surgery following ARMD revision 239
6.9.3. Clinical implications of findings 242
6.10. Summary 244
7. CONCLUSIONS AND FUTURE WORK 246
7.1. Summary of the problem and thesis aims 246
7.2. Conclusions 247
7.3. Clinical recommendations 251
7.4. Limitations 256

7.5. Recommendations for future work 257

Appendix 1. Methods for blood metal ion sampling 262

Appendix 2. Outcomes following ARMD revision surgery: review methodology and study quality 263

Appendix 3. Methodology and limitations for financial analysis of regulatory authority patient surveillance 265

Appendix 4. Data collection form used at the time of hip resurfacing arthroplasty surgery…….....268

Appendix 5. Postal questionnaire sent to hip resurfacing arthroplasty patients .269

Appendix 6. Oxford Hip Score questionnaire completed by hip resurfacing arthroplasty patients 271

Appendix 7. Patient and implant factors for non-revised hip resurfacing arthroplasty patients completing the follow-up questionnaire compared with those not completing the questionnaire 273

Appendix 8. NJR data collection form for primary hip arthroplasty 274

Appendix 9. NJR data collection form for revision hip arthroplasty.....277

Appendix 10. Details of cross-sectional imaging protocols at different centres 280

Appendix 11. Summary of the receiver operating characteristic analysis for optimal Birmingham Hip Resurfacing blood metal ion thresholds compared with fixed authority thresholds in the validation cohort 283

Appendix 12. Detailed statistical methodology and results for the clinical risk scoring tool 284

Appendix 13. Data collection form used by radiologist when assessing cross-sectional imaging 287

Appendix 14. Detailed review of diagnostic test characteristics for cross-sectional imaging vs. intra-operative findings 288
Appendix 15. University of California, Los Angeles (UCLA) activity score questionnaire completed by hip resurfacing arthroplasty patients 290
Appendix 16. Patient demographics for the cohort undergoing initial assessment that were subsequently excluded from the analysis 291
Appendix 17. Detailed review of diagnostic test characteristics for initial assessment findings vs. the development of new pseudotumours 293
Appendix 18. Letter of study approval from the National Joint Registry Research Sub-committee 295
Appendix 19. National Joint Registry patient consent form 297
Appendix 20. Methodology for multiple imputation of missing data.....299
Appendix 21. Competing-risk regression analysis to identify predictors of re-revision surgery when using the imputed dataset 301
Appendix 22. Indications for re-revision surgery in matched metal-on-metal hip arthroplasty revision patients 303
Appendix 23. Sensitivity analysis using the complete dataset with revision body mass index excluded 304
Appendix 24. Analysis of outcomes following revision surgery in the unmatched cohort…… 306
Appendix 25. Publications associated with this thesis.....307
BIBLIOGRAPHY 314

TABLE OF FIGURES

Figure 2.1 Histopathological specimen from a metal-on-metal hip arthroplasty revised for adverse reactions to metal debris. 21

Figure 2.2 Anteroposterior hip radiograph illustrating a metal-on-metal hip resurfacing arthroplasty with the acetabular component in excessive inclination. 34

Figure 2.3 Pseudotumour in association with a metal-on-metal hip resurfacing arthroplasty identified on ultrasound examination. 36

Figure 3.1 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years. 81

Figure 3.2 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years by gender. 83

Figure 3.3 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years by implant design.....84

Figure 3.4 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to eight years by year of primary surgery. 87

Figure 4.1 Inclusion and exclusion criteria for Birmingham Hip Resurfacing (BHR) patients in the derivation and validation cohorts. ARMD = adverse reactions to metal debris.....111

Figure 4.2 An explanation of receiver operating characteristic curve analysis with example lines for varying diagnostic test performance. 114

Figure 4.3 Receiver operating characteristic curve illustrating the ability of four blood metal ion parameters to distinguish between unilateral Birmingham Hip Resurfacing patients with and without adverse reactions to metal debris in the derivation cohort. 121

Figure 4.4 Receiver operating characteristic curve illustrating the ability of four blood metal ion parameters to distinguish between bilateral Birmingham Hip Resurfacing patients with and without adverse reactions to metal debris in the derivation cohort. 122

Figure 5.1 Anteroposterior hip radiograph illustrating the three zones on the acetabular side (blue) and the three zones on the femoral side (red), which were used when assessing radiolucent lines. 144

Figure 5.2 Method for measuring acetabular component inclination and version on an anteroposterior pelvic radiograph using ImageJ software. 145

Figure 5.3 Method for measuring femoral neck narrowing on an anteroposterior hip radiograph. 146

Figure 5.4 Plot to assess calibration of the final multivariable model. 154

Figure 5.5 Disagreement of ultrasound with the presence of an intra-operative pseudotumour which was also confirmed on MARS-MRI. 163

Figure 5.6 Disagreement of both ultrasound and MARS-MRI with the absence of a pseudotumour intra-operatively. 164

Figure 6.1 Histogram of revision body mass index values before data cleaning. 200

Figure 6.2 Histogram of revision body mass index values after data cleaning. 200

Figure 6.3 Kaplan-Meier cumulative implant survival rate following revision surgery performed for adverse reactions to metal debris at up to five years. 211

Figure 6.4 Kernel density plot illustrating the distribution of the propensity scores before matching. 225

Figure 6.5 Kernel density plot illustrating the distribution of the propensity scores after matching. 226

Figure 6.6 Kaplan-Meier cumulative implant survival rate following revision surgery at up to five years by revision indication.....229

Figure 7.1 Algorithm for the investigation and management of hip resurfacing arthroplasty patients.....252

TABLE OF TABLES

Table 2.1 Implant survival rates following Birmingham Hip Resurfacing at 10 to 15 years follow-up. 14

Table 2.2 Summary of published MARS-MRI pseudotumour classification systems. ..39

Table 2.3 Summary of studies comparing ultrasound and MARS-MRI for detecting adverse reactions to metal debris. 42

Table 2.4 Studies reporting the outcomes following metal-on-metal hip arthroplasty revision surgery performed for adverse reactions to metal debris. 51

Table 2.5 Follow-up guidance for metal-on-metal hip arthroplasty patients published by worldwide regulatory authorities. 60

Table 2.6 Differences in annual follow-up costs for all metal-on-metal hip arthroplasties recorded in the National Joint Registry when using various worldwide surveillance protocols. 68

Table 3.1 Indications for hip resurfacing arthroplasty revision surgery. 80

Table 3.2 Patient and implant factors for the whole cohort, and for hip resurfacing arthroplasties not revised for pseudotumour compared with those revised for pseudotumour. 82

Table 3.3 Multivariable Cox proportional hazards models for identifying hip resurfacing arthroplasties at risk of pseudotumour revision for the whole cohort, and subdivided by gender. 85

Table 4.1 Patient demographics and outcomes for the derivation and validation cohorts. 117

Table 4.2 Median (interquartile range) blood metal ion parameters for the derivation and validation cohorts. 118

Table 4.3 Summary of the receiver operating characteristic analysis for the derivation cohort. 120

Table 4.4 Summary of the receiver operating characteristic analysis for optimal Birmingham Hip Resurfacing blood metal ion thresholds compared with fixed authority thresholds in the derivation cohort. 124

Table 4.5 Summary of the receiver operating characteristic analysis for the validation cohort. 125

Table 5.1 Patient demographics for all hip resurfacing arthroplasties, and subdivided by study group. 149

Table 5.2 Radiographic features for all hip resurfacing arthroplasties, and subdivided by study group. 151

Table 5.3 Logistic regression models identifying patient and radiographic factors predictive of pseudotumour revision. 153

Table 5.4 Summary of patient demographics for revised hip resurfacing arthroplasties. 158

Table 5.5 Findings identified on cross-sectional imaging performed prior to revision surgery. 162

Table 5.6 Cases where either or both cross-sectional imaging modalities disagreed with the presence or absence of an intra-operative pseudotumour. 165

Table 5.7 Diagnostic test characteristics for cross-sectional imaging identifying intra-operative pseudotumours in revised hip resurfacing arthroplasties.....166

Table 5.8 Patient demographics for the cohort undergoing repeat assessment. 172

Table 5.9 Change in ultrasound grade between initial and repeat assessments. 176

Table 5.10 Diagnostic test characteristics for asymptomatic hip resurfacing arthroplasties not developing new pseudotumours following repeat ultrasound for three clinical scenarios based on the initial assessment. 178

Table 6.1 Patient and surgical factors for all metal-on-metal hip arthroplasties revised for adverse reactions to metal debris. 209

Table 6.2 Indications for hips undergoing re-revision surgery following revision surgery performed for adverse reactions to metal debris. 210

Table 6.3 Competing-risk regression analysis to identify predictors of re-revision surgery following revision surgery performed for adverse reactions to metal debris. 213

Table 6.4 Patient and surgical factors in the unmatched cohort and in the matched cohort. 224

Table 6.5 Outcomes following metal-on-metal hip arthroplasty revision surgery in the matched cohort. 227

Table 6.6 Implant survival rates following metal-on-metal hip arthroplasty revision surgery by specific revision indication. 230

1. INTRODUCTION

1.1. Overview

Total hip arthroplasty (THA) represents the most successful surgical treatment for the long-term alleviation of painful hip joint arthritis [1]. Almost 90,000 THAs are implanted annually in England, Wales, Northern Ireland, and the Isle of Man [2] with these numbers expected to increase markedly in future years [3]. However young (under 60 years of age) and active patients undergoing THA have traditionally experienced poor outcomes and high implant failure rates [4-6].

Metal-on-metal hip arthroplasty (MoMHA) was introduced in the form of stemmed MoM THA and hip resurfacing arthroplasty (HRA: bone preserving procedure with retention of the femoral head and neck) as a treatment option for young and active patients with hip arthritis. It was thought that these large-diameter MoM bearing surfaces would reduce implant wear and the risk of dislocation [7], thereby decreasing failure rates in a group of patients who had previously experienced unsatisfactory outcomes following conventional THA. Approximately 1.5 million MoMHAs have been implanted worldwide, with 35% of all primary THAs implanted annually in the United States (US) having a MoM bearing at one stage [8].

In recent years unexpectedly high short-term failure rates have been reported for most MoMHA designs, which has led to an almost worldwide cessation of these particular implants [2, 9-11]. Revision surgery for MoMHAs can be required for traditional modes of implant failure such as loosening, infection, and fracture, or for an

unanticipated failure mechanism termed adverse reactions to metal debris (ARMD) [2, 12, 13]. The latter includes painful MoMHA failures with one or more of the following features in the absence of another plausible diagnosis [14, 15]: large sterile effusions (including pseudotumour), macroscopic tissue necrosis, metallosis, and histopathological evidence of aseptic lymphocytic vasculitis associated lesions (ALVAL). Pseudotumour represents a specific type of ARMD, and can be defined as a cystic, solid, or mixed mass communicating with the hip joint [16].

In MoMHAs ARMD is the sequelae of metal debris released from the bearing surface and/or other THA modular implant junctions due to wear and corrosion [17-21]. This debris can result in destruction of the local bone and soft-tissues, as well as large invasive pseudotumours, which often require revision surgery [12, 14, 16]. Many early MoMHA revisions have now been performed for ARMD with the prevalence of ARMD revision surgery continuing to increase [2, 11]. Initial reports suggested the short-term outcomes following ARMD revision surgery were poor and inferior when compared with non-ARMD MoMHA revision surgery [12].

To identify ARMD early, regulatory authorities worldwide have published follow-up guidance for patients with MoMHAs [22-24]. Patients with ARMD may be asymptomatic which can make diagnosis difficult [25, 26], therefore currently most MoMHA patients require some form of annual surveillance for the implant life-time. This surveillance can include clinical examination, hip radiographs, blood metal ion sampling (surrogate marker of in-vivo MoM bearing wear), and cross-sectional imaging (with ultrasound or metal artefact reduction sequence magnetic resonance imaging (MARS-MRI)) [22-24]. However a recent review of MoMHA follow-up protocols

issued by five worldwide authorities demonstrated that the guidance was extremely variable between authorities, not evidence-based and very costly [27]. Furthermore the specific roles of the main investigative modalities are not yet established.

Although the prevalence of ARMD revision surgery is increasing, little remains known about outcomes following revision [28]. The poor outcomes initially reported following ARMD revision are likely to be related to the invasive and destructive nature of these lesions [12, 16]. This led regulatory authorities and orthopaedic surgeons to widely recommend performing early ARMD revision surgery, especially in patients with imaging abnormalities and/or rising blood metal ion concentrations [12, 22-24, 29]. However these recommendations are not robust and reflect a lack of evidence, and the impact of performing surgery early for ARMD is not yet known. Furthermore no guidance is available regarding the type of revision surgery (which components require revision, and which bearing surface/fixation methods to use) to perform for ARMD. It therefore remains unclear whether and when MoMHAs should be revised and, in those undergoing revision, the type of surgery required is uncertain. Therefore surgeons currently experience difficulty when managing MoMHA patients.

Despite the high failure rates of MoMHAs [9, 10], it is estimated that at least 80% of these implants remain in-situ worldwide [2, 11]. Over future years many MoMHA patients will be subjected to regular, but potentially unnecessary, investigations. A significant proportion of patients will ultimately undergo revision surgery and experience the morbidity associated with these procedures. Delivering this care will be hugely expensive for health services. It is therefore important to develop evidence-based guidance for the investigation and management of MoMHA patients.

To provide effective follow-up for MoMHA patients it is important to determine specifically which patients need surveillance, identify the role of each investigative modality, and establish the regularity and duration of follow-up for different patient subgroups. For managing MoMHA patients, detailed knowledge of outcomes following revision would allow surgeons to counsel patients more informatively about the risks of undergoing further surgery. Identifying prognostic factors of outcome following ARMD revision would provide surgeons with thresholds for performing revision, and guide decisions regarding the type of surgery required.

1.2. Hip arthroplasty terminology and cohorts assessed

In this thesis different terminology has been used to refer to specific types of hip arthroplasty. Each term has been defined below.

- THA = All types of THA regardless of the bearing surface, including MoMHA.
- Conventional THA = All THAs with non-MoM bearing surfaces.
- MoMHA = HRA and MoM THA.

Due to the cohorts available, this thesis has predominantly been based on HRA patients. However the National Joint Registry (NJR) cohort assessed also included MoM THA patients undergoing revision surgery. Therefore the NJR work was based on MoMHA patients.

1.3. Broad aims

The broad aims of this thesis were:

1. To determine the clinical impact of pseudotumour in HRA patients.
2. To develop evidence-based guidance for the investigation of HRA patients.
3. To develop evidence-based guidance for the management of MoMHA patients.

1.4. Specific objectives

The specific objectives of this thesis were:

1. To determine the 15 year revision rate for pseudotumour in HRA patients, and identify risk factors for pseudotumour revision.
2. To determine whether blood metal ion concentrations could effectively identify patients with unilateral and bilateral HRAs who had ARMD.
3. To determine whether a combination of patient and radiographic factors predicted HRA patients with evidence of pseudotumour.
4. To compare the diagnostic test characteristics of ultrasound, MARS-MRI, and combined imaging (ultrasound and MARS-MRI) for identifying HRA patients with intra-operative pseudotumours.
5. To identify risk factors associated with developing new pseudotumours on ultrasound in asymptomatic HRA patients.
6. To determine the outcomes following MoMHA revision surgery performed for ARMD, and to identify pre-operative and intra-operative factors predictive of re-revision surgery.
7. To compare the outcomes following MoMHA revision surgery performed for ARMD with the outcomes following MoMHA revision surgery performed for non-ARMD indications.

1.5. Structure of the thesis

This thesis has been structured as follows:

Chapter 2 provides a literature review about what is currently known regarding the clinical impact of ARMD and pseudotumour, and also describes how MoMHA patients are presently investigated and managed.

Chapter 3 reports the 15 year pseudotumour revision rate in a large prospective single centre cohort of HRA patients and explored risk factors for pseudotumour revision.

Chapter 4 presents a large prospective single centre cohort study assessing the effectiveness of blood metal ion concentrations for identifying both unilateral and bilateral HRA patients with ARMD. The optimal blood metal ion thresholds established from this study were subsequently validated in a large external cohort of HRA patients.

Chapter 5 describes two retrospective single centre studies exploring the roles of hip radiographs and cross-sectional imaging (ultrasound and MARS-MRI) for identifying HRA patients with pseudotumours. In addition a prospective longitudinal single centre study explored risk factors for developing new pseudotumours on ultrasound in patients with asymptomatic HRAs.

Chapter 6 reports two retrospective observational studies involving data from the NJR on MoMHA patients undergoing revision surgery. This included a study to identify factors predictive of re-revision surgery following ARMD revision, and a propensity

score matched study that compared the outcomes following ARMD and non-ARMD revision procedures.

Chapter 7 summarises the important findings and conclusions from this thesis, with recommendations subsequently made for relevant future work.

1.6. Data conventions

To avoid repetition within this thesis the statistical analyses sections for each experimental chapter only include the specific analyses relevant to the respective chapter. Unless otherwise stated all statistical analyses described in this thesis were performed using Stata Version 13.1 (Lakeway Drive, Texas, US). For descriptive statistics of normally distributed continuous variables the mean and either the range or standard deviation (SD) were used. For descriptive statistics of continuous variables without a normal distribution the median and interquartile range (IQR) were used. For categorical variables the number (percentage) were used. Unless otherwise stated the significance level for all analyses was $p<0.05$, with 95% confidence intervals (CI) also used.

1.7. Declaration

All work presented in this thesis was performed by Gulraj Matharu, unless otherwise stated. This work included, but was not limited to, study design, applications for data release from the NJR and other institutions, data collection (including searching clinical databases, analysing radiographs, and reviewing operative records), data analysis, interpretation of the results, and writing of the thesis. If more than one observer has

been mentioned Gulraj Matharu was always the first observer. The specific role of any other individuals performing work presented in this thesis has been clearly stated.

2. LITERATURE REVIEW*

2.1. Metal-on-metal hip arthroplasty

2.1.1. Rationale, development, and usage

Young and active patients undergoing conventional THA (metal with conventional polyethylene bearings) have traditionally experienced poor outcomes and high implant failure rates [4-6]. The Finnish Arthroplasty Register reported unsatisfactory 15 year all-cause implant survival rates of between 58% and 71% (depending on implant fixation) in patients under the age of 55 years following primary conventional THA [5]. In this young and active group excessive polyethylene wear debris over time resulted in osteolysis and aseptic component loosening, ultimately requiring THA revision surgery [30]. Therefore the surgical management of end-stage hip arthritis in young and active patients has been challenging.

In an attempt to improve outcomes for young and active patients with hip arthritis, modern 3rd generation HRAs with MoM bearing surfaces were developed during the 1990s. However neither the concept of HRA nor MoM bearings were new for treating patients with hip arthritis. The first attempt at any form of femoral head surface replacement was made by Smith-Petersen in the 1940s [31]. During the 1950s Sir John Charnley developed the first HRA as would be recognised today, which was an uncemented all Teflon surface arthroplasty [32]. Initial attempts at using HRA with MoM bearings were made in small patient cohorts by both Edward Haboush in 1953

* Work presented in this chapter has been published in:
Matharu GS, Mellon SJ, Murray DW, Pandit HG. Follow-up of metal-on-metal hip arthroplasty patients is currently not evidence based or cost effective. *J Arthroplasty* 2015; **30**(8): 1317-1323.

[33], and Maurice Muller in 1965 [34], though these were not successful and the concept was shortly abandoned. During the 1970s, metal-on-polyethylene HRAs were implanted by Michael Freeman [35], Heinz Wagner [36], and Harlan Amstutz [37] (1st generation devices). However these implants experienced issues with implant fixation and/or the bearing surface, ultimately leading to failure from osteolysis, component loosening, and/or accelerated wear [38].

It was subsequently recognised that some MoM HRAs originally implanted by Maurice Muller had only needed revision 25 years later [39]. Furthermore good long-term outcomes were also reported for three large-diameter MoM THA designs that had been implanted during the 1950s and 1960s, namely the McKee-Farrar, Ring, and Stanmore prostheses [40-42]. On this basis Derek McMinn, a surgeon from Birmingham in the United Kingdom (UK), concluded that the poor initial results of HRA were related to failure of materials rather than failure in concept [43]. Derek McMinn and Heinz Wagner each subsequently developed an all uncemented MoM HRA (2nd generation devices), however these implants also experienced early failures of both components for aseptic loosening [44, 45]. Refinements were made by Derek McMinn to develop a hybrid MoM HRA (cemented femoral component and uncemented acetabular component), which was the Birmingham Hip Resurfacing (BHR: initially manufactured by Midland Medical Technologies, Birmingham, UK) [43]. During a similar time period Harlan Amstutz, a surgeon from the US, also developed a hybrid MoM HRA (Conserve Plus, Wright Medical Technology, Memphis, Tennessee, US) [46]. The BHR and the Conserve Plus represented 3rd generation HRA devices and were both used clinically from 1997.

Third generation MoM HRA devices had numerous perceived advantages. Using MoM bearing surfaces resulted in significantly lower wear rates compared with conventional metal-on-polyethylene THA articulations [7]. Larger MoM femoral head sizes were theoretically associated with lower bearing wear rates compared with smaller sizes [43]. Furthermore large-diameter MoM bearings increased the femoral head-neck ratio, which had advantages such as increased range of hip motion, and a reduced risk of dislocation and impingement of the femoral neck on the acetabulum [47, 48]. Perceived advantages of HRA included restoration of native hip joint anatomy and biomechanics, with near normal physiological femoral loading leading to reduced stress-shielding [49-52]. The HRA procedure was also femoral bone conserving (head and neck retained), which suggested that revision surgery for a failed HRA to a THA would be relatively straightforward [53].

Promising early outcomes following modern MoM HRA were published initially by the designing surgeons [54-56], and subsequently by independent centres [57-61]. The use of HRA gradually gained acceptance from the profession with other companies soon introducing HRA implants, such as the Articular Surface Replacement (ASR: DePuy, Leeds, UK) which was introduced in 2003. The development of large-diameter MoM THAs closely followed. These devices were largely developed to treat isolated femoral component failures requiring revision (such as femoral neck fractures and avascular necrosis of the femoral head); however surgeons also used large-diameter MoM THAs as a primary arthroplasty in patients otherwise eligible for HRA but with abnormal femoral anatomy precluding HRA, and when wishing to utilise the benefits of large-diameter MoM bearings without undertaking technically demanding HRA [43, 59, 62-64].

Consequently in the mid-2000s the use of large-diameter MoMHA increased substantially. Approximately 1.5 million MoMHAs have been implanted worldwide, predominantly in young and active patients. Between 2005 and 2006, 35% of all primary THAs in the US had a MoM articulation [8]. The peak of MoMHA use in the NJR for England and Wales was recorded in 2007, during which 22% of all primary THAs had a MoM bearing surface [2].

2.1.2. Implant revision rates for all-causes

In recent years MoMHAs have been associated with unanticipated high short-term revision rates. The National Institute for Health and Care Excellence (NICE) currently recommend that prostheses used for primary THA should have an all-cause revision rate of 5% or less at ten years (equivalent to 0.5% or less per year) [65]. The major worldwide joint registries have demonstrated that as a collective MoMHAs do not meet these NICE criteria.

The NJR has reported high all-cause revision rates at 12 years for both HRA (13.6%) and uncemented MoM THA (22.1%) [2]. The Australian Joint Replacement Registry observed a 15 year all-cause revision rate of 13.3% for HRA, and 21.8% for MoM THAs [11]. The two commonest indications for revision surgery in HRAs were pain followed by ARMD in the NJR [2], and component loosening/lysis (33% of all revisions) and ARMD (28% of all revisions) in the Australian Joint Replacement Registry [11]. The two most frequent indications for revision surgery in MoM THAs were ARMD followed by pain in the NJR [2], and ARMD (42% of all revisions) and component loosening/lysis (27% of all revisions) in the Australian Joint Replacement

Registry [11]. Similarly recent data from the Finnish Arthroplasty Register demonstrated a high all-cause revision rate for all HRAs of 14% at ten years, with the commonest indications for revision being aseptic component loosening (53% of all revisions) followed by other indications that included ARMD (24% of all revisions) [66].

The revision rates for large-diameter MoM THAs have consistently been higher compared with HRAs, even when bearing surfaces are identical [2, 11, 15], with no large-diameter MoM THA implant design meeting the NICE criteria for continued usage [65]. The ASR and ASR XL were the worst performing HRA and MoM THA devices respectively, which were subject to a voluntary manufacturer recall in 2010 and subsequently were no longer implanted [67]. However the revision rates for these two devices with the same MoM bearing design were considerably different at 6 years (ASR 25% and ASR XL 49%) [15], suggesting different mechanisms were responsible for implant failure (Chapter 2.2.2.).

These unexpected outcomes have resulted in an almost worldwide cessation of the use of MoMHA. In 2015 the NJR reported that only 0.9% of all primary THAs implanted were HRAs (was 10.8% in 2006), with almost no MoM THAs recorded [2]. However the outcomes of HRA are heavily influenced by implant design, with the BHR consistently reported as the best performing HRA device worldwide [2, 11, 66]. Therefore it is important to acknowledge that there is evidence supporting the continued role of the BHR for treating carefully selected young and active male patients with hip arthritis [13, 68-75]. Both the designing surgeons and some independent reports have achieved the NICE criteria for continued implant usage in this select group of patients,

although none of these studies can satisfy the NICE criteria in female patients (Table 2.1.) [65].

Study and year	Number of hips	Overall survival rate (95% CI)	Survival rate in males (95% CI)	Survival rate in females (95% CI)
Designing centres				
Matharu 2013 [68]	447	94.1% (84.9-97.3) at 14 years ^	100% (100-100) at 14 years *^	91.2% (68.6-98.7) at 14 years *
Daniel 2014 [13]	1000	95.8% (95.1-96.5) at 15 years ^	98.0% (97.4-98.6) at 15 years ^	91.5% (89.9-93.2) at 15 years
Independent centres				
Murray 2012 [69]	646	87.1% (83.0-91.2) at 10 years	94.7% (92.0-97.4) at 10 years	73.9% (63.6-84.2) at 10 years
Coulter 2012 [70]	230	94.5% (90.1-96.9) at 10 years	97.5% (92.4-99.2) at 10 years ^	89.1% (79.2-94.4) at 10 years
Holland 2012 [71]	100	92.0% (86.7-97.3) at 10-years	94.6% (89.4-100) at 10 years	84.6% (70.7-98.5) at 10 years
Van Der Straeten 2013 [72]	250	92.4% (90.8-94.0) at 13 years	98.7% (97.5-99.8) at 13 years ^	79.8% (76.4-83.2) at 12 years
Reito 2014 [73]	261	91.0% (89.0-93.0) at 10 years	93.1% (91.0-95.2) at 10 years	86.0% (82.1-89.9) at 10 years
Mehra 2015 [74]	120	94.2% (88.8-98.7) at 10 years	96.0% (90.7-100) at 10 years ^	92.5% (85.8-99.9) at 10 years
Azam 2016 [75]	244	93.7% (NS) at 10 years	95.4% (NS) at 10 years ^	89.9% (NS) at 10 years

Table 2.1 Implant survival rates following Birmingham Hip Resurfacing at 10 to 15 years follow-up.
CI = confidence interval; NS = not stated.
*** Patients with primary hip osteoarthritis only, with aseptic revision used as the endpoint for survival analysis.**
^ Implant survival rate meets the criteria proposed by NICE for continued usage [65].

2.1.3. Modes of implant failure

Revision surgery for MoMHAs can be required for all of the traditional modes of conventional THA implant failure, such as aseptic component loosening, infection, peri-prosthetic fracture, osteolysis, and dislocation [2, 11, 13, 66, 76]. Less common modes of conventional THA implant failure have also been observed with MoMHAs,

including implant fracture and femoral head/acetabular component size mismatch [2, 66].

Two unique modes of HRA failure were observed due to retention of the femoral head and neck, namely femoral neck fractures and avascular necrosis of the femoral head [2, 11, 13, 66, 76]. In the initial HRA studies femoral neck fractures were reported as the main indication for revision within the first post-operative year, with the overall prevalence of fractures observed to be 1.5% or less in large cohorts [13, 63, 76-78]. Both patient (female gender and poor bone quality) [77] and surgical (femoral neck notching and disruption of the femoral head blood supply during HRA) [79, 80] risk factors have been implicated for femoral neck fractures following HRA. Compromise of the femoral head blood supply during HRA may also result in avascular necrosis and femoral head collapse, with a reported prevalence of up to 1.3% [13, 76]. It is worth acknowledging that many of these unique HRA femoral failures were revised to large-diameter MoM THAs with retention of the HRA acetabular component [12, 53, 63, 81].

From the mid-2000s onwards there were reports of an unexpected mode of failure associated with MoMHAs, which usually required revision within five years of primary arthroplasty [16, 82-84]. This clinical problem was later collectively defined by the term ARMD (Chapter 2.2.1.) [14]. This mode of MoMHA failure was initially considered to be related to problems at the MoM bearing surface [14, 17, 18]. However it subsequently became evident that ARMD could also arise from the femoral head and taper junction in MoM THAs [15, 19, 62, 85], with more recent reports of this problem occurring with conventional THAs without a MoM articulation [86, 87]. Over the years

ARMD has become increasingly recognised as a cause of MoMHA failure, with this being largely responsible for the almost worldwide cessation of MoMHA [2, 11, 66, 88].

2.2. ARMD and pseudotumour

2.2.1. Terminology

Numerous terms have been used to describe this unexpected mode of failure associated with MoMHAs, which are considered below.

Adverse reactions to metal debris (ARMD)

This term was introduced by Langton *et al*. in 2010 as an umbrella term for painful MoMHA failures with one or more of the following features in the absence of another plausible diagnosis [14, 15]: large sterile effusions (including pseudotumour), macroscopic tissue necrosis, metallosis and ALVAL. The term ARMD has generally been accepted as the most inclusive for describing this heterogeneous clinical problem.

Pseudotumour

In 2008 Pandit *et al.* introduced the term pseudotumour, which was defined as a cystic, solid, or mixed mass communicating with the hip joint [16]. These pseudotumour cases had presented clinically from 2006 onwards, with the initial cases discussed at the British Orthopaedic Oncology Society meeting prior to publication (Stanmore, UK, 2007). The peri-prosthetic masses were aseptic and non-neoplastic with uncertain aetiology, therefore were labelled pseudotumour. By definition pseudotumour represents a specific type of ARMD.

Metallosis

Metallosis has been defined as macroscopic black-staining of the soft-tissues associated with suboptimally functioning MoMHAs [89, 90].

Aseptic lymphocytic vasculitis associated lesions (ALVAL)

This term was described by Willert *et al*. in 2005 and referred to a specific lymphocyte-dominated histopathological appearance observed with failed second generation MoMHAs requiring revision surgery [82]. The characteristic features of ALVAL were diffuse and perivascular infiltrates of T and B lymphocytes and plasma cells, high endothelial venules, massive fibrin exudation, accumulation of macrophages, infiltrates of eosinophilic granulocytes, and tissue necrosis [82].

Adverse local tissue reaction (ALTR)

Some reports have used ALTR rather than ARMD when describing problems related to failing MoMHAs [91, 92]. The term ALTR has been more frequently used recently given the emerging reports of a similar clinical picture in conventional THAs [87, 93]. This is presumably because the aetiology and pathogenesis of this condition when associated with non-MoM bearing surfaces remain incompletely understood as it is possible that these reactions are not related to metal debris.

Adverse soft tissue reaction to particulate debris

This term was introduced by the NJR in June 2008 as a revision indication on data capture forms [2]. Although it was added prior to ARMD being formally described, adverse soft tissue reaction to particulate debris can be considered synonymous with ARMD in failing MoMHAs [14, 15].

Potential problems associated with all of the outlined terminology include considerable overlap and incorrect usage. For example all pseudotumours by definition have ARMD, however the converse is not necessarily true. Similarly ALVAL cannot be used to describe the response in all ARMD cases given that ALVAL refers to a specific histopathological appearance, which may not always be present [94, 95]. Furthermore diagnosis can be subjective, with different surgeons using variable criteria to diagnose any entity. These limitations have important implications, namely when making comparisons between different studies and when estimating the magnitude of the problem. This thesis subsequently uses the nomenclature of ARMD and pseudotumour, which are both defined as described above [14-16].

2.2.2. Aetiology and pathogenesis

Implant retrieval and histopathological analysis following revision surgery have both been key in establishing the aetiology and pathogenesis of ARMD associated with MoMHAs. The two largest implant retrieval centres in the UK have significantly contributed to our understating about MoMHA failure, namely the London Implant Retrieval Centre (the world's largest retrieval centre for failed orthopaedic implants) [96] and the Northern Retrieval Registry [97].

In summary ARMD in MoMHAs is caused by metal debris released from the bearing surface and/or other THA modular implant junctions, with the metal debris produced by wear and corrosion [17-21]. This metal debris subsequently elicits a response in the local peri-prosthetic tissues [82, 98, 99], which leads to the features frequently observed

at clinical presentation (Chapter 2.2.3.), investigation (Chapter 2.3.), and revision (Chapter 2.4.) in MoMHA patients with ARMD.

Early observations suggested ARMD was the result of a delayed hypersensitivity reaction (type IV) to metal debris given the histopathological features described in ALVAL (Chapter 2.2.1.) [16, 82]. Such reactions did not necessarily need to be associated with excess metal debris but instead could be caused by previous exposure to cobalt and chromium, such as wearing jewellery or staged implantation of bilateral MoMHAs [100]. However subsequent research in revised ARMD patients demonstrated that high bearing surface wear due to edge loading was the primary mechanism for developing this problem in HRAs, and even in the early MoM THA failures that were analysed [14, 17, 18]. Lymphocyte proliferation responses to cobalt and chromium also suggested that a delayed hypersensitivity reaction to metal debris might not represent the dominant biological process for developing ARMD [100]. However further work observed a small proportion of HRA patients revised for ARMD had low bearing surface wear with histopathological features of an exacerbated adaptive immune response, suggesting that some cases may actually be related to a delayed hypersensitivity reaction [20, 98, 99, 101]. Therefore it can broadly be concluded that most ARMD cases are a consequence of a normal immune response to excessive amounts of metal debris (estimated as 80% of ARMD cases in one study) [99], whilst the minority of ARMD cases develop from an exaggerated immune response to small amounts of metal debris.

It has subsequently become apparent that ARMD in association with MoM THAs is more commonly the result of metal debris released from mechanical wear and/or

corrosion at modular implant junctions [15, 19, 21, 62, 85, 97, 102] rather than bearing surface wear [14, 18]. Evidence suggests that the metal ion/particulate release in MoM THAs arises from micro-motion leading to opening of the modular junctions (most commonly the femoral head-taper and femoral taper-stem junctions) which allows the ingress of bodily fluids, causing re-passivation of the metal alloys within the modular junctions, followed by dissolving of the titanium and/or cobalt-chrome alloys [19, 103]. The outlined process commonly occurs at modular junctions with mixed articulating alloys, such as a titanium alloy stem with a cobalt-chrome femoral head [21, 104]. However cases of ARMD have been reported in modular junctions that comprise of cobalt-chrome alone [87, 93].

The histopathological appearance of peri-prosthetic tissues around MoMHAs revised for ARMD represents a broad spectrum with a number of scoring systems proposed to improve the categorisation of these responses, namely by Campbell *et al*. [98] and Grammatopoulos *et al*. [99]. These scoring systems aim to categorise the varying degrees of both the innate immune response (non-specific response causing macrophage accumulation and tissue necrosis) and adaptive immune response (characterised by significant lymphocytic infiltrate) to wear debris, which are observed with this heterogeneous condition [105]. However the interplay between the innate and adaptive immune responses are extremely complex [105], with more recent evidence demonstrating that a small proportion of patients exhibit features of tertiary lymphoid organs (TLOs) within the peri-prosthetic tissues around MoMHAs revised for ARMD (Figure 2.1.) [94, 95]. These TLOs have been reported in certain autoimmune conditions, such as rheumatoid arthritis [106], with the locally generated T and B cell lymphocytes having the potential to cause substantial tissue destruction [107].

Furthermore a recent study observed that the metal debris released from femoral head-taper junctions was of a different composition and morphology compared with that released from the bearing surface, which may explain the higher revision rates observed in MoM THAs compared with HRAs [15, 108]. These more recent observations require further investigation.

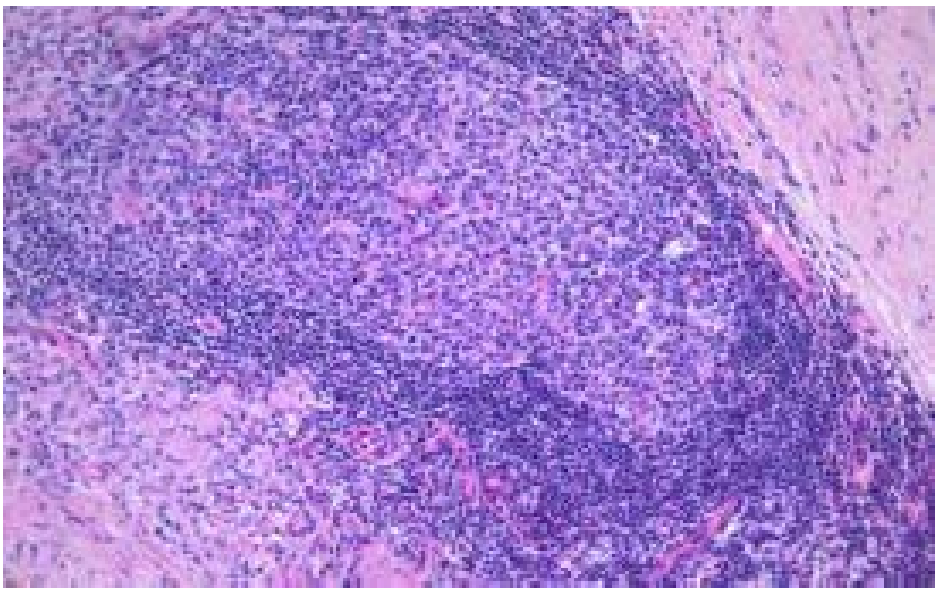


Figure 2.1 Histopathological specimen from a metal-on-metal hip arthroplasty revised for adverse reactions to metal debris.
Formalin fixed sections presented with image taken at ten times magnification. The specimen demonstrates an organised T cell and B cell lymphoid aggregate with a central follicular area. The appearances are characteristic of a tertiary lymphoid organ. This figure has been reproduced from the open-access publication by Mittal *et al*. [94].

2.2.3. Clinical presentation

Patients with ARMD can present with both local and systemic features.

2.2.3.1. Local presentation

Patients with ARMD are frequently symptomatic, however asymptomatic cases where revision surgery has either been recommended or performed are not uncommon [109, 110]. Symptomatic patients complain of pain in the groin, buttock, lateral aspect of the hip, or thigh [14, 16, 111]. Patients may also describe swellings or masses around the

hip joint and mechanical symptoms, such as clicking, clunking, grinding, squeaking, catching, or instability [14, 16, 112, 113]. In some cases patients may present with nerve palsy (femoral or sciatic) or vascular symptoms (claudication or venous thrombosis) due to compression from ARMD lesions, dislocation related to destructive lesions, or localised skin rashes [16, 114, 115]. Any of these outlined features may also be identified during a comprehensive clinical examination, during which it is important to assess gait for limping [112].

2.2.3.2. Systemic presentation

Patients with ARMD may develop systemic symptoms due to exposure to high cobalt and chromium concentrations. These systemic features are extremely rare but can broadly be divided into neurological (hearing and visual impairment/loss, peripheral neuropathy, and cognitive impairment), cardiovascular (cardiomyopathy, breathlessness), and endocrine (hypothyroidism, fatigue, malaise, depression) [116-119]. A review reported that MoMHA patients with systemic features had median serum cobalt concentrations of 35 μg/l (range 14 μg/l to 288 μg/l), with symptoms often resolving after revision to a non-MoM articulation [116].

There have been some case reports of death due to cardiac failure secondary to cobalt toxicity in MoMHA patients, even following revision arthroplasty [120, 121]. A recent study from the Australian Joint Replacement Registry reported an association between ASR XL MoM THAs and hospital admission for heart failure in male patients [122]. However this study had numerous methodological limitations. A larger unpublished study using NJR data demonstrated that MoMHA patients were not at increased risk of heart failure, with the same authors reporting that MoMHA patients with high blood

metal ion concentrations undergoing comprehensive cardiac investigations had no detectable heart pathology [123].

In the occupational setting an established relationship exists between high metal ion exposure and the development of certain cancers [124]. This has led to concerns regarding the biological effects of metals released into the body following MoMHA. Currently analysis of large cohorts from the NJR and Finnish Arthroplasty Register have not demonstrated any increased risk of cancer or mortality in MoMHA patients compared with conventional THA patients at short-term follow-up [125-129]. However these analyses should be repeated at extended follow-up given the long latent period required for most cancers to develop.

2.2.4. Differential diagnosis

Alternative diagnoses in patients with problematic MoMHAs can be divided into pathology which is either intrinsic or extrinsic to the hip [112, 130]. Intrinsic causes can be subdivided into intra-capsular (such as infection, loosening, instability, fracture) and extra-capsular (such as iliopsoas tendonitis or trochanteric bursitis), whilst extrinsic causes can include a wide range of conditions such as hernia, malignancy, peripheral vascular disease, and spinal pathology [112, 130].

When considering the diagnosis of ARMD, the most difficult differential diagnosis to exclude is infection. This is because the clinical picture of infection can be very similar to ARMD. Both conditions can have raised blood inflammatory markers and peri-prosthetic collections, which resulted in surgeons performing staged revisions when initially treating ARMD [16, 131]. Patients with MoMHAs where ARMD may be

suspected must therefore be thoroughly investigated to exclude any other potential causes (Chapter 2.3.).

2.2.5. Revision rates for ARMD and pseudotumour

A number of single-centre studies have reported cumulative MoMHA revision rates specifically for ARMD and pseudotumour within the first decade of implantation. A large cohort consisting of non-ASR HRAs reported a pseudotumour revision rate of 4% at eight years [132], with a 3% ARMD revision rate at five years reported in a cohort of BHRs [89]. Another study observed high six year ARMD revision rates for ASR HRAs and ASR XL MoM THAs of 25% and 49% respectively [15]. Reito *et al*. reported an eight year ARMD revision rate for large-head (>50 mm) ASR XL MoM THAs of 43% [133], and six year ARMD revision rates for small-head (<50 mm) ASR XL MoM THAs and ASR HRAs of 39% and 27% respectively [134]. The 36 mm Pinnacle system (DePuy, Leeds, UK) was one of the most commonly used MoM THAs worldwide, with ARMD revision rates of 16% at both nine years [102] and ten years [135] reported recently. No study has presently reported ARMD or pseudotumour revision rates more than ten years following MoMHA.

Numerous other studies have reported the prevalence of revision surgery performed for ARMD largely within ten years of primary MoMHA, however revision rates specifically for this endpoint were not reported. A recent systematic review of 122 eligible studies observed the prevalence of ARMD revision surgery to be variable for HRAs (0.13% to 9.5%) and MoM THAs (0.29% to 21.3%), with the prevalence closely related to the intensity of the patient surveillance protocols used in the different studies [136].

The large cohorts recorded in national arthroplasty registries would provide an ideal source for determining ARMD revision rates, with evidence that the prevalence of ARMD revision surgery is continuing to increase [2, 11]. However most registries only report the prevalence of ARMD revision rather than cumulative ARMD revision rates [11, 66]. The NJR does provide revision rates for ARMD in the form of patient-time incidence-rates (the total number of revisions for ARMD divided by the total of the individual patient-years at risk), which were high at 3.89 revisions per 1,000 patient-years at risk for HRA, and 9.68 revisions per 1,000 patient-years at risk for uncemented MoM THAs [2]. To put this into context, the patient-time incidence-rates for every failure indication for conventional MoM THAs were all under 1.5 revisions per 1,000 patient-years at risk [2]. However registries have potential limitations that can underestimate ARMD revision rates. In the NJR these limitations include the underreporting of MoMHA failures [96, 137], and the delayed introduction of ARMD as a revision indication in June 2008 [2]. Given the latter it is suspected that a substantial number of revisions truly performed for ARMD prior to 2008 were coded in the registry using other indications, such as unexplained pain [138]. Furthermore registries do not collect histopathological data, which is required to confirm the diagnosis of ARMD [82, 98, 99].

2.2.6. Risk factors for ARMD and pseudotumour revision

Large cohort studies have identified risk factors for all-cause MoMHA revision, which include female gender, small femoral components in HRAs, large femoral components in MoM THAs, implant design, and hip diagnoses other than primary osteoarthritis [9-11, 139, 140]. Some risk factors for ARMD and pseudotumour revision are similar to

those for all-cause revision, and can be classified into patient, implant, and surgical risk factors. However it is important to consider risk factors specifically for ARMD and pseudotumour revision given this information will help inform current follow-up regimens and assist in stratifying patients for surveillance [22-24, 141, 142].

2.2.6.1. Patient risk factors

Females with MoMHAs are at increased risk of ARMD and pseudotumour revision [69, 132, 134]. One study also identified young age as a predictor, reporting a high pseudotumour revision rate of 13% at 6 years in females under 40 years at HRA [132]. The increased risk of failure in females is likely to be multifactorial. Females require smaller components due to anatomical differences. Females also more frequently have abnormal acetabular anatomy compromising cup placement [132], as well as metal allergies, which may contribute to ARMD and pseudotumour [16, 143]. Gender and femoral head size have complex interactions with mixed conclusions reported when considering the risk of all-cause HRA revision surgery [10, 76, 139, 144]. However current evidence suggests that being female is the strongest predictor of pseudotumour revision rather than small femoral HRA component size [69, 132].

Although patients with a primary diagnosis of hip dysplasia have higher all-cause revision rates following HRA [13, 139], hip dysplasia was not predictive of pseudotumour revision when other variables, such as gender and femoral component size, were controlled for [69, 132]. Patients with bilateral MoMHAs have an increased risk of ARMD revision [102, 135]. This may be related to the increased metal ion and particulate generation from two implants compared with one [145], or delayed hypersensitivity responses to staged MoMHA implantation [146].

2.2.6.2. Implant risk factors

A higher risk of ARMD revision surgery has been observed with MoM THA compared with HRA, even when bearing surfaces were identical [15]. Furthermore the specific HRA and MoM THA implant design influences ARMD revision rates, with the ASR design performing poorest in both categories. In HRA, the ASR had a higher revision rate for ARMD at five years (10%) compared with both the Conserve Plus and BHR [20]. In MoM THAs, the ASR XL had a higher revision rate for ARMD [15] compared with Pinnacle MoM THAs [102, 135].

Specific design features have been implicated for the high ASR failure rates. The lower coverage arc of the acetabular component (angle subtended by the articular surface of the acetabular component: 146°-152°) and reduced diametrical clearance (calculated by subtracting the head diameter from the cup or liner diameter: 49.47 μm) can both increase the risk of edge loading [14, 15, 102, 134, 147, 148]. Furthermore the ASR XL femoral head-taper junction appears increasingly prone to releasing metal debris [15], with evidence that taper failure can be exacerbated by MoM THA systems with larger femoral head diameters, larger femoral head offsets, and varus femoral stems [19]. It has also been suggested that shorter and rougher femoral stem tapers can contribute to taper failure [149].

Small femoral component size in HRA has not been identified as predictive of ARMD or pseudotumour revision when gender was controlled for [69, 132]. In MoM THAs an increased risk of ARMD revision has been associated with smaller liner sizes in one study [102], with other authors observing larger femoral head sizes [133] and the use of

Corail femoral components (DePuy, Leeds, UK) with ASR acetabular systems both predict ARMD revision [133, 134].

2.2.6.3. Surgical risk factors

Acetabular component malposition has been associated with an increased risk of revision for ARMD and pseudotumour in HRAs [20, 113, 150]. Grammatopoulos *et al*. reported that the risk of pseudotumour requiring revision was higher in HRAs with acetabular components positioned outside an optimal zone (inclination of 35°-55° and anteversion 10°-30°) [150]. Other reports have confirmed excessive inclination, and either excessive or insufficient anteversion of the acetabular component to be associated with ARMD revision in HRAs [20, 113].

Hard-on-hard bearings have a low tolerance for acetabular component positioning outside an optimal zone, with such malposition associated with edge-loading and high bearing surface wear leading to failure [150]. However ARMD and pseudotumours can develop even when acetabular components are within the optimal zone [20, 150, 151], and well-functioning HRAs can have malpositioned acetabular components [13, 29, 70, 73]. Therefore the relationship between acetabular component position and HRA failure is complex, and may be better explained by more refined radiological parameters, such as the contact patch to rim distance (distance from the contact area between the femoral and acetabular components to the acetabular rim) [148, 152]. Substantial reduction in the femoral head-neck ratio following HRA can increase the risk of pseudotumour by increasing the risk of impingement and edge loading [153]. By contrast acetabular component malposition does not appear to predict ARMD revision in MoM THAs,

presumably because failure primarily occurs at the femoral head-taper junction [102, 154].

2.3. Investigative modalities

2.3.1. Blood metal ion sampling

2.3.1.1. Rationale, methods, and interpretation

Blood metal ion concentrations are a surrogate marker of in-vivo wear in MoMHAs, and therefore have been used to investigate patients [155]. Cobalt and chromium ions are measured, which constitute the primary elements of the MoM alloy. Wear (normal and excessive) and corrosion of MoMHAs causes release of both insoluble metal particles and soluble metal ions [43]. Metal particles are present in the synovial fluid and peri-prosthetic tissues. Metal ions enter the bloodstream, therefore allowing measurement, and are eventually excreted in the urine [156].

Whole blood or serum samples are used for measuring metal ion concentrations. Sample collection and analysis, and units of measurement have been described (Appendix 1.). Serial samples should be sent to the same laboratory given that clinically significant variations can occur in blood samples from the same patients analysed at different laboratories [157]. To prevent incorrect interpretation of serial results the same blood fraction must be used, though the preference has been whole blood over serum [158]. Serum cobalt and chromium are up to 1.4 times greater than their respective values in whole blood, with stronger correlation of cobalt in serum and whole blood compared with that of chromium [159, 160].

Advantages of blood metal ion sampling include being relatively inexpensive (£30 per test) [161] and simple if protocols are followed. Disadvantages relate to potential patient (renal function, diet/medication/supplements containing trace metals, occupational exposure, other metal eluting implants) and laboratory (contamination, inter-laboratory and intra-laboratory measurement variability) factors that can influence test interpretation [24, 156, 157]. Therefore management decisions should never be made based solely on one blood test.

2.3.1.2. Blood metal ions in healthy controls and well-functioning hips

In 3042 healthy subjects, 97% had both blood cobalt and chromium concentrations of 2 µg/l or less, with very little variation observed between individuals [162]. In well-functioning MoMHAs, metal ions can enter the bloodstream by day five following implantation [158]. Blood metal ion concentrations then increase significantly over the first year during the bearing running-in phase [163]. Concentrations then either stabilise or slowly decrease from one year postoperatively during medium-term follow-up [158, 164, 165]. However some authors have also observed significantly decreased blood metal ion concentrations by ten years [72, 145, 166].

2.3.1.3. Factors influencing blood metal ion concentrations

High blood metal ion concentrations in MoMHA patients have been associated with poorly functioning implants [145, 167, 168], including those specifically with ARMD requiring revision [14, 102, 154, 162]. Given that ARMD is caused by metal debris generated from wear and corrosion (Chapter 2.2.2.) and that blood metal ions are a

marker of these processes, many factors that influence blood metal ion concentrations are also risk factors for ARMD revision (Chapter 2.2.6.). Predictors of raised blood metal ion concentrations previously identified in MoMHAs can also be divided into patient (female gender, young age, increased time from MoMHA to blood test), implant (specific MoMHA design, small femoral HRA components, large femoral MoM THA components, bilateral MoMHAs), and surgical factors (acetabular component malposition outside the optimal zone [150], reduced contact patch to rim distance) [145, 148, 152, 169-178].

A recent large study involving 1,748 patients with numerous different MoMHA implant designs observed that MoM THAs had significantly higher blood metal ion concentrations compared with HRA patients [172]. This confirms the findings from two randomised controlled trials of MoM THAs and HRAs with identical bearing surfaces [179, 180], with the higher metal ion concentrations in MoM THAs likely to be related to the additional wear and corrosion occurring at the modular implant junction(s) (Chapter 2.2.2.). These mechanisms may also explain the observation that in HRAs the ratio of blood cobalt to chromium concentration is usually equal (one-to-one ratio), but in MoM THAs cobalt is often raised disproportionately compared with chromium [168, 172]. Although the cobalt-chromium ratio may be a useful parameter in MoM THAs the only study formally assessing its utility [181] had numerous important limitations [182, 183], therefore further studies are needed.

2.3.1.4. Blood metal ion thresholds for clinical use

Since 2010, blood cobalt and/or chromium concentrations above 7 μg/l have been a cause for concern by some authorities [23, 67, 141, 184]. Initial reports suggested that

this arbitrarily selected threshold had good specificity (89%) but poor sensitivity (52%) for detecting unilateral failed MoMHAs, with the optimal threshold identified as a cobalt or chromium of 4.97 μg/l (86% specificity and 63% sensitivity) [167]. Similarly another study including 190 unilateral MoM THAs demonstrated 7 μg/l had low specificity (65%) and sensitivity (57%) for identifying ARMD on MARS-MRI [185]. By contrast a subsequent study involving 597 unilateral ASR MoMHAs identified 7 μg/l as optimal for identifying implant failure (specificity 93% and sensitivity 51%), although the authors suggested blood metal ions should not be used alone for screening for failed implants [168].

Thresholds for poorly functioning HRAs have been proposed by Van Der Straeten *et al*. (cobalt 4.0 μg/l and chromium 4.6 μg/l for unilateral HRA) [145] and others (cobalt 4.5 μg/l) [162] with all cut-offs having higher specificity than sensitivity. Only one study has devised thresholds for poorly functioning bilateral HRAs, which were 5.0 μg/l for cobalt (91% specificity and 39% sensitivity) and 7.4 μg/l for chromium (93% specificity and 43% sensitivity) [145]. However these thresholds have not been validated. Blood metal ion concentrations above 20 μg/l have been associated with osteolysis at ARMD revision, therefore it was recommended that revision should be considered in patients above this threshold [186]. More recently a US consensus statement proposed risk-stratifying MoMHA patients to assist management, with blood cobalt and chromium concentrations arbitrarily grouped into low risk (<3 μg/l), moderate risk (3 μg/l to 10 μg/l), and high risk (>10 μg/l) [112]. However the clinical effectiveness of these new thresholds has yet to be investigated.

2.3.1.5. Repeat blood metal ion sampling

After repeating blood metal ions within one year in 254 unilateral ASR patients Reito *et al*. observed that most HRAs had concentrations below published thresholds [145] on both tests, whilst 32% of MoM THAs with normal initial cobalt concentrations experienced significant increases on repeat testing [187]. The authors reported similar observations in 76 bilateral ASR patients undergoing repeat blood testing [188]. Another study repeating blood metal ions at a mean of 27 months in 205 patients with HRAs observed cobalt concentrations did not significantly change with time since implantation [186]. Following MoMHA revision most patients have significantly decreased or normal blood concentrations within six months of bearing exchange, although chromium decreases at a slower rate than cobalt [189-191].

2.3.2. Radiographs

Radiographs provide important information on component position, bone quality, and implant fixation, therefore they can identify signs suggestive of MoMHA failure early [192]. However radiographs cannot directly diagnose ARMD or pseudotumours, given that these are predominantly soft-tissue lesions [193, 194]; but they can identify abnormalities associated with these lesions. Radiological signs reported in MoMHAs revised for ARMD include component loosening, osteolysis, femoral neck narrowing, component malposition (Figure 2.2.), fracture, subluxation/dislocation, and in HRAs femoral neck erosion due to impingement [20, 29, 112, 150, 153, 192, 195]. Furthermore radiographs can identify other causes of hip pain, such as heterotopic ossification (HO) [192].



Figure 2.2 Anteroposterior hip radiograph illustrating a metal-on-metal hip resurfacing arthroplasty with the acetabular component in excessive inclination.

When assessing radiographs it is important to compare serial images, including the immediate post-primary radiograph, for subtle signs of MoMHA failure. Surgeons are advised to assess the retro-acetabular, ischial, and pubic regions carefully [112]. Medial calcar erosion represents a radiological sign suggestive of ARMD in MoM THAs [196]. However it remains unclear whether femoral neck narrowing represents a normal physiological process following HRA or a clinically significant finding given that neck narrowing has been reported in both well-functioning patients [13, 70, 197] and ARMD revisions [192, 195].

All radiological signs suggestive of MoMHA failure can be identified from serial anteroposterior pelvic radiographs, apart from acetabular component version. The latter requires a lateral hip radiograph, though version can be calculated from an anteroposterior pelvic radiograph using specialised software (Chapter 5.3.4.2.).

Advantages of radiographs when assessing MoMHAs include their wide availability, low cost, and ease of interpretation, with the main disadvantage being radiation exposure to the pelvis.

2.3.3. Cross-sectional imaging

2.3.3.1. Ultrasound

Ultrasound has frequently been used as the initial cross-sectional imaging modality for investigating MoMHA patients [198-200]. Compared with MARS-MRI, the main advantages of ultrasound include being cheaper (ultrasound £49 vs. MARS-MRI £216) [161], faster to perform, not being affected by metal artefact from the prosthesis, and being more accessible with fewer patient contraindications (can use in patients with pacemakers or those with claustrophobia) [201]. Ultrasound also permits dynamic imaging, hip aspirations, and biopsies [201, 202]. The main disadvantages of ultrasound include the technique being operator dependent which makes images difficult to assess retrospectively, with potential difficulties when assessing deeper structures or examining larger patients [112, 201]. Ultrasound is also not as useful as MARS-MRI when planning revision surgery [203].

Ultrasound can be sensitive for identifying numerous abnormalities around MoMHAs [199], provided a systematic examination is performed [204]. These abnormalities include pseudotumours of variable sizes, which are often cystic or solid and located anterior or posterolateral to MoMHAs (Figure 2.3.) [198, 199, 202, 205]. A recent study of 82 MoMHAs undergoing revision concluded that ultrasound was effective for detecting intra-operative pseudotumours [206]. Ultrasound has also been useful for detecting other ARMD pathology include joint effusions, iliopsoas and trochanteric

bursal collections, capsular and bursal thickening, and synovitis [199, 200]. Other abnormalities that ultrasound can identify, which may not always be related to ARMD, include muscle atrophy and tendon avulsions, namely of the hip abductors and short external rotators [207, 208].

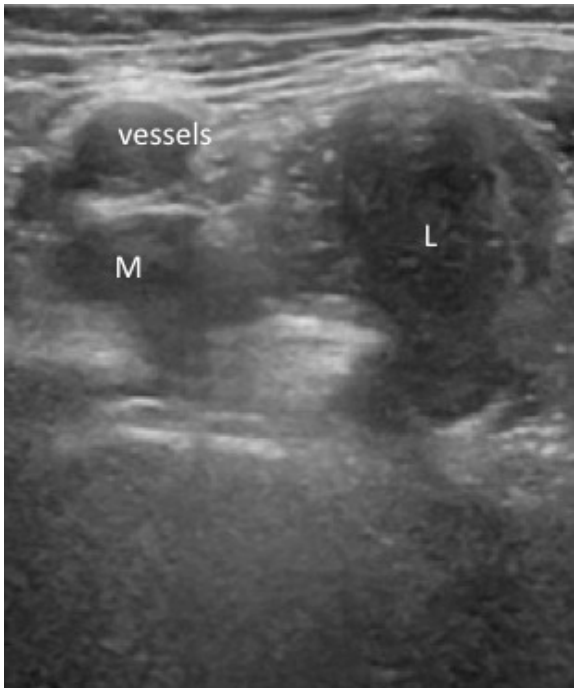


Figure 2.3 Pseudotumour in association with a metal-on-metal hip resurfacing arthroplasty identified on ultrasound examination.
Anterior view of the hip demonstrating an enlarged mixed pseudotumour lying in the psoas bursa region. The pseudotumour lies in close proximity to the femoral vessels (marked) and has a smaller medial (M) and a larger lateral (L) component.

Systems classifying ARMD abnormalities around MoMHAs on ultrasound have been proposed by Nishii *et al*. [199] (normal, joint-expansion pattern, cystic pattern, and mass pattern) and Matharu *et al*. [209] (normal, bursal collection/thickening, pathological effusion, and pseudotumour). The ultrasound grading systems proposed by Siddiqui *et al*. [210] were specifically designed for classifying pseudotumours (cystic thin-wall, cystic thick irregular wall with atypical fluid, and solid complex lesion) and muscle atrophy, and were devised from MARS-MRI classifications [25, 211]. It must

be acknowledged that the prevalence of ARMD in MoMHAs on both ultrasound and MARS-MRI has been much higher (up to 61%) [25, 26, 202, 205] than that confirmed at revision surgery (Chapter 2.2.5.). As ARMD on imaging has now frequently been observed around well-functioning conventional THAs [212-214], it is possible that current cross-sectional imaging classifications have over-diagnosed imaging ARMD in MoMHA patients [25, 199, 209, 210, 215, 216].

Ultrasound has a role in identifying both the development and progression of ARMD associated with MoMHAs. Repeat ultrasound in a small MoMHA cohort with asymptomatic ARMD on initial imaging demonstrated that most lesions increased in size in the short-term with occasional remission of small masses [109]. Similarly a study involving 48 patients undergoing repeat ultrasound at short-term follow-up observed 44% of MoMHAs with initially normal ultrasounds developed abnormalities by repeat assessment, with no factors predictive of progression [209]. By contrast, a recent study by Matharu *et al*. repeated ultrasounds in ten HRAs with asymptomatic pseudotumours between 4.0 to 6.5 years later, with no changes occurring in pseudotumour volume or patient reported outcomes between assessments [110]. This may suggest that some asymptomatic ARMD lesions can be safely monitored with interval scanning.

2.3.3.2. MARS-MRI

Most centres use MARS-MRI as the initial cross-sectional imaging modality for investigating MoMHAs, with MARS-MRI likely to have an increasing role in managing patients as techniques become refined [112]. Compared with ultrasound, the main advantages of MARS-MRI include not being operator dependent, providing

excellent visualisation of soft-tissue structures including deeper tissues, and being able to assess images retrospectively which is helpful when obtaining further opinions or planning revision [201, 203]. Comparison can be made with the contralateral hip, and it is also easier to compare serial MARS-MRI images rather than serial ultrasounds [201]. The main disadvantages of MARS-MRI include the high cost, peri-prosthetic metal artefact despite suppression techniques, longer examination times, reduced availability, limited intervention potential, and being contraindicated in patients with pacemakers [201].

Using an optimised MARS-MRI protocol allows detection of numerous abnormalities around MoMHAs. Pseudotumours have frequently been reported which can be of variable sizes, located anterior or posterolateral to MoMHAs, with a cystic, solid, or mixed consistency [25, 26, 216, 217]. The three main MARS-MRI pseudotumour classification systems have been summarised (Table 2.2.) [25, 215, 216]. Comparison of these systems suggested the Anderson *et al*. classification was most reliable [218]. The Modified Oxford Classification has been proposed more recently [219], which classifies MARS-MRI scans initially into four groups (normal, trochanteric fluid, effusion, ARMD), with the ARMD group subsequently classified according to Hauptfleisch *et al*. [216]. However small peri-prosthetic effusions may still be missed on MARS-MRI because of prosthetic artefact [201].

Anderson *et al*. [215]		**Hart *et al*. [25, 210]**		**Hauptfleisch *et al*. [216]**	
Grade A	*Normal* Normal post-operative appearances including seromas and small haematomas	Type 1	*Pseudotumour* Flat, thin-walled (≤2 mm); fluid-like content	Type 1	*Pseudotumour* Thin-walled cystic mass (cyst wall <3 mm)
Grade B	*Infection* Fluid-filled cavity with high signal T2 wall; inflammatory changes in soft tissues, ± bone marrow oedema	Type 2a	*Pseudotumour* Thick-walled (>2 mm) or irregular walls; fluid-like content	Type 2	*Pseudotumour* Thick-walled cystic mass (cyst wall >3 mm, but less than the diameter of the cystic component)
Grade C1	*Mild MoM disease* Peri-prosthetic soft-tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either <5 cm maximum diameter	Type 2b	*Pseudotumour* Thick-walled (>2 mm) or irregular walls; atypical fluid	Type 3	*Pseudotumour* Mainly solid mass
Grade C2	*Moderate MoM disease* Peri-prosthetic soft-tissue mass/fluid-filled cavity >5 cm diameter or C1 lesion with either of following: 1) muscle atrophy or oedema in any muscle other than short external rotator; or 2) bone marrow oedema: hyperintense on STIR	Type 3	*Pseudotumour* Solid		
Grade C3	*Severe MoM disease* Any of the following: 1) fluid-filled cavity extending through deep fascia; 2) tendon avulsion; 3) intermediate T1W soft-tissue cortical or marrow signal; 4) fracture				

Table 2.2 Summary of published MARS-MRI pseudotumour classification systems.
MoM = metal-on-metal; STIR = short tau inversion recovery; T1W = T1 weighted image; T2W = T2 weighted image.

When pseudotumours have been detected, MARS-MRI can provide detailed information regarding any soft-tissue or muscle invasion and destruction as well as neurovascular involvement, which assists when planning revision surgery [203]. Furthermore studies by Liddle *et al*. [203] (85% sensitivity and 59% specificity; n=39) and Lainiala *et al*. [220] (71% sensitivity and 87% specificity; n=167) have demonstrated that MARS-MRI was effective for diagnosing intra-operative pseudotumours. The latter study also reported that these results varied with MARS-MRI timing, with scans performed over one year before revision surgery having lower sensitivity (29%) for detecting intra-operative pseudotumours compared with scans performed within one year before revision surgery (77% to 88% sensitivity) [220].

Other ARMD abnormalities also detected on MARS-MRI include muscle atrophy, tendon avulsions, and osteolysis [201, 209, 221, 222]. Hip abductor atrophy has frequently been reported in MoMHAs and can progress over one year, with some authors suggesting atrophy may be a precursor to irreversible muscle destruction [222].

Like ultrasound, MARS-MRI can identify both the development and progression of ARMD associated with MoMHAs. Most studies repeating MARS-MRI in MoMHA patients with normal initial scans have observed that few changes occur within two years of repeat imaging [223, 224], with any development of ARMD from normal scans occurring over several years [225]. Reito *et al*. reported that of 119 MoMHAs (all ASRs) with normal initial MARS-MRI scans, 8% developed new pseudotumours when examination was repeated at a median time of 19 months although no factors predicted progression [224]. By contrast, a study of 155 MoM THAs undergoing repeat MARS-MRI within a mean time of 15 months observed that the risk of disease progression was

high in patients with initially normal scans (30%), and in those with isolated trochanteric fluid (47%) and effusions (58%) on initial MARS-MRI [219]. Disease progression on imaging was associated with high blood cobalt concentrations, or an irregular pseudocapsule lining at initial MARS-MRI [219].

Of the four studies repeating MARS-MRI in MoMHA patients with asymptomatic pseudotumours, one suggested most cystic lesions decreased in size by one year [226] whilst two reported little to no change with serial imaging at one year [223] and at a minimum of four years [227]. In the latter study pseudotumour progression on MARS-MRI was associated with increased cystic wall thickness and atypical mixed fluid signal, but was not related to blood metal ion concentrations [227]. The fourth study involving 24 asymptomatic MoMHAs reported that pseudotumours frequently changed in size when MARS-MRI was repeated at a mean time of 20 months, with larger lesions most likely to increase [228].

2.3.3.3. Ultrasound vs. MARS-MRI

Five studies have compared ultrasound and MARS-MRI for detecting ARMD (Table 2.3.) [201, 229-232]. Four studies demonstrated that ultrasound was effective and recommended ultrasound for initial screening, with MARS-MRI reserved if ultrasound abnormalities required further assessment [229-232]. Garbuz *et al*. concluded that a negative ultrasound excluded pseudotumour in asymptomatic patients given ultrasound was 100% sensitive [229]. Furthermore Kwon *et al*. repeated both the ultrasound and MARS-MRI in 42 MoMHAs at a mean time of 14 months, and demonstrated that ultrasound remained effective for detecting ARMD at longitudinal assessment [232]. By contrast, the authors of the smallest study reported that ultrasound was inferior to

MARS-MRI for detecting pseudotumours and muscle atrophy, however ultrasound was superior to MARS-MRI for diagnosing effusions and tendon pathology [201].

Study (participants)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Garbuz *et al*. [229] (40 hips) Ultrasound ***vs. MARS-MRI***	100	96	92	100
MARS-MRI ***vs. ultrasound***	92	100	100	96
Siddiqui *et al*. [201] (19 hips) Ultrasound ***vs. MARS-MRI***	69	83	90	56
Nishii *et al*. [230] (64 hips) Ultrasound ***vs. MARS-MRI***	74	92	NS	NS
Muraoka *et al*. [231] (83 hips) Ultrasound ***vs. MARS-MRI***	72	88	65	91
Kwon *et al*. [232] (42 hips) Initial: Ultrasound ***vs. MARS-MRI***	81	92	86	88
Repeat: Ultrasound ***vs. MARS-MRI***	86	88	75	94

Table 2.3 Summary of studies comparing ultrasound and MARS-MRI for detecting adverse reactions to metal debris.
NS = not stated. Gold standard imaging modality for diagnostic test characteristics has been highlighted in bold italics.

The major limitation of all studies was comparing ultrasound and MARS-MRI either by using MARS-MRI as the gold standard or by assuming agreement between modalities was acceptable [201, 229-232]. The benchmark for diagnostic accuracy in such studies should be correlation with the findings at revision. One study performing both

ultrasound and MARS-MRI validated imaging findings in five patients undergoing revision with good correlation between both modalities and intra-operative findings [201]. This parallels other observations when either ultrasound or MARS-MRI alone was performed [203, 206, 220]. Other important study limitations making interpretation difficult include delays of up to one year between performing ultrasound and MARS-MRI in the same patient, and including both HRAs and MoM THAs in the same cohort given the different failure mechanisms (Chapter 2.2.2.).

2.3.3.4. Computed tomography

Computed tomography (CT) scanning provides useful information regarding MoMHA component positioning and fixation, and can also detect osteolysis or fractures which may be radiographically occult [203, 233, 234]. Some authors suggest modern CT techniques allow the evaluation of pseudotumours [234]. One study concluded that CT was unsuitable for routine MoMHA imaging with a sensitivity of 44% for diagnosing pseudotumours compared with MARS-MRI, although CT was superior for detecting osteolysis [235].

Given the advantages of both ultrasound (Chapter 2.3.3.1.) and MARS-MRI (Chapter 2.3.3.2.), CT should not be recommended as a first-line imaging modality when assessing MoMHA patients. The main disadvantages of CT include pelvic radiation, image artefact from prostheses, and high cost.

2.3.4. Other investigations

Infection and ARMD can have similar presentations, therefore a comprehensive assessment is required. This includes blood tests for the erythrocyte sedimentation rate

(ESR) and C-reactive protein (CRP). Infected MoMHA revision patients have significantly higher ESR and CRP values compared with non-infected revision cases (which included ARMD failures) [236]. However the ESR and CRP may also be raised in aseptic ARMD patients [16, 131], and where ARMD and infection coexist [195, 237, 238].

Hip aspirations can assess for both infection (synovial fluid culture, sensitivity, white blood cell count, and cell differential i.e. percentage of polymorphonuclear cells) [130, 236] and ARMD (synovial cobalt and chromium concentrations) [14, 155, 202, 239]. Synovial metal ion concentrations can assist in cases where blood metal ions and cross-sectional imaging have been equivocal despite clinical concerns, and may be informative when patients have problematic bilateral MoMHAs. In a recent study leukocyte esterase strip testing of synovial fluid samples was reliable at ruling out infection in MoMHAs revised for ARMD [240]. Aspirations should be performed in an aseptic environment, however the main disadvantages include the potential risk of introducing infection [241], and the exposure to pelvic radiation if imagine intensification is employed.

Other useful investigations for unexplained symptoms include single-photon emission CT [242], diagnostic hip injections, and investigating other causes, for example using targeted imaging for spinal or non-orthopaedic pathology.

2.4. Revision surgery for ARMD

2.4.1. Threshold for revision surgery

The poor outcomes initially reported following ARMD revision surgery [12, 243] led regulatory authorities and surgeons to widely recommend performing early revision in MoMHAs with ARMD [12, 23, 24, 29, 48]. However there are currently no robust thresholds for performing ARMD revision surgery, which stems from a lack of evidence. Therefore surgeons currently have difficulty when managing patients with ARMD, namely which patients to revise and which to keep under surveillance.

The limited evidence-base and problems clinicians experience is highlighted by expert opinion (level 5 evidence) [244] commonly being used to manage patients [22, 245]. In 2012, one UK centre introduced an internet-based multidisciplinary team to assist other centres in managing their MoMHA patients with cases referred electronically [245]. The multidisciplinary team consists of experts (specialist surgeons and radiologists), who meet weekly to recommend management decisions based on their experience, regulatory guidance [23, 24], and the latest evidence. The proposed management recommendations have been implemented by referring institutions in 92% of cases [245].

For some cases management is clear. Timely revision surgery is needed for symptomatic patients with pseudotumours that may be solid, large, invasive and destructive to the soft-tissues and bone, sometimes with neurovascular damage [12, 14, 16, 186, 203]. Similarly asymptomatic patients with normal investigations are also straightforward and do not require revision. This leaves a large cohort where management remains uncertain. Such patients include those with moderate to no

symptoms with abnormal investigations, which can include non-destructive cystic pseudotumours and/or moderately raised blood metal ion concentrations [109, 110, 226-228]. This uncertainty is highlighted by a recent study where ten MoMHA clinical scenarios were used to assess the management decisions made by international experts from six centres [246]. Inconsistency in the agreement between centres was observed specifically when managing raised or rising blood metal ion concentrations, cystic pseudotumours, and peri-acetabular osteolysis [246]. Although studies have attempted to identify predictors of cross-sectional imaging progression [209, 219, 224, 227], and predictors of poor outcomes following ARMD revision surgery (Chapter 2.4.3.4.), these have been limited by short-term follow-up and were small and underpowered for identifying predictors. Therefore further good quality evidence is required to appropriately guide the management of MoMHA patients, and specifically to develop robust thresholds for performing ARMD revision surgery.

2.4.2. Intra-operative findings and surgical strategy

The surgical management of ARMD has evolved over time [12, 29, 203], with the heterogeneity of ARMD not being appreciated initially (Chapter 2.2.). Therefore MoMHA revisions performed for ARMD can range from simple to more complex cases [203].

Simple cases to manage surgically include synovitis with minimal tissue damage and metallosis, with a satisfactory reconstruction usually achieved with primary conventional THA implants. Furthermore when revising MoM THAs some surgeons have advised retaining acetabular and/or femoral components when these are well-fixed and well-positioned, with taper adapters recommended if femoral tapers are not

severely damaged [93, 191, 247]. In such instances procedures may be performed where either one component is revised (acetabular or femoral), or both components are retained with exchange of the modular components only (femoral head and acetabular liner, with or without the use of a taper adapter). Both strategies are simpler surgically compared with revising the entire MoM THA construct, and they reduce any potential risks associated with removing the acetabular and/or femoral components.

Complex cases can include large pseudotumours that are destructive to the peri-prosthetic bone and soft-tissues, and also involve vital neurovascular structures (such as the femoral vessels or the sciatic nerve). Such complex cases are likely to require extensive reconstruction with revision THA implants with or without augments. Occasionally staged revision procedures for ARMD may be considered for severe cases or when infection is also a concern. Furthermore surgical expertise from other specialties (plastics, vascular, orthopaedic oncology, or pelvic surgeons) may also be required depending on the lesion anatomy, nature of the destruction, and the neurovascular involvement. This highlights the importance of performing detailed pre-operative investigations, and planning for surgeons with other expertise to be available at revision as well as the necessary reconstructive implants [203].

Dealing with large soft-tissue lesions around MoMHAs has represented a new challenge for arthroplasty surgeons. Given the potential risk of ARMD recurrence [12, 28], the aim of revision surgery is to excise all metal debris within the soft-tissues, including pseudotumours, similar to an oncological resection. However in reality this aim can become compromised depending on the lesion anatomy, namely if it involves vital neurovascular structures and/or the primary soft-tissues contributing to hip

stability. In such cases some metal debris and/or pseudotumour may be intentionally left at revision, with the aim to maintain limb and hip joint function over the risk of ARMD recurrence.

Currently there is little guidance available regarding the specific type of reconstruction to perform for surgeons undertaking ARMD revision. The only universally accepted recommendation is revising to a non-MoM bearing surface, given the potential for high wear from MoM bearings (Chapter 2.2.2.) and the poor outcomes subsequently reported when MoM bearings were used in early ARMD revisions [195, 203, 248]. Apart from the bearing, individual surgeons decide on the most appropriate reconstruction when performing ARMD revisions based on the individual scenario and the surgeons own preferences and reconstructive experience.

2.4.3. Outcomes following ARMD revision surgery

2.4.3.1. Evidence available and patient demographics

A systematic review undertaken in 2013 by Matharu *et al*. identified six studies reporting the outcomes following 216 MoMHA revisions performed for ARMD [28]. This review was updated as described (Appendix 2.). Including the six initial studies the updated review identified a total of 15 unique studies eligible for inclusion with 803 MoMHA revisions for ARMD [12, 29, 90, 191, 195, 203, 247-255]. One further study reported the outcomes following 16 ARMD revisions at extended follow-up (median of 10.3 years) [256], with these cases previously reported on [12].

The patient demographics and outcomes following ARMD revision surgery for all included studies have been summarised (Table 2.4.). Most revisions were performed in

females, in young patients (under 60 years), and within five years of primary MoMHA (Chapter 2.2.6.). Revisions were almost equally split between HRAs and MoM THAs. The initial systematic review [28] concluded that there was limited evidence on outcomes following MoM THA revisions performed for ARMD [247]. Therefore a number of studies have subsequently reported on the outcomes following MoM THA revisions.

Study author and year	Hips revised for ARMD	Female hips (n)	Mean age (range) at revision in years	Mean time to revision (range) in years	Mean follow-up time after revision (range) in years	Risk of complications (n)	Risk of re-revision (n)	Main reasons for re-revision surgery (n)	Functional outcome **
Grammatopoulos 2009 [12]	16 HRA	100% (16)	51.3 (20-71)	1.6 (0.01-6.7) *	3.0 (0.8-7.2) *	50% (8)	38% (6)	Dislocation +/- ARMD recurrence (4) Loose cup (2)	Mean OHS 20.9
Rajpura 2011 [249]	11 HRA	36% (4)	53.5 (22-67)	3.8 (1.3-7.3)	1.8 (1.0-3.3)	18% (2)	18% (2)	ARMD recurrence (2)	Mean OHS 35.3
De Smet 2011 [29]	48 HRA	61% (NS) *	52.5 (18-71)	2.7 (0.3-8.4)	3.3 (0.3-10.1) *	≤ 23% (11) *	≤ 13% (6) *	Loose cup or stem (2) Infection (2) ARMD recurrence (1)	Mean HHS 93.1 *
Ebreo 2011 [90]	42 total HRA & THA	55% (23)	Median 61 (NS)	4.7 (1.3-7.8) *	2.2 (1.2-4.0) *	≤ 10% (4) *	≤ 2% (1) *	Infection (1)	Mean OHS 23.7 *
Liddle 2013 [203]	32 HRA	81% (26)	57.7 (25-74)	4.3 (0.9-10.9)	Median 2.5 (1.0-4.5) *	≤ 6% (2) *	≤ 6% (2) *	Dislocation (1) ARMD recurrence with loose cup (1)	Median OHS 36.5 *
Su 2013 [250]	13 HRA	85% (11)	NS	NS	2.3 (0.7-6.7) *	≤ 15% (2) *	≤ 15% (2) *	Infection (2)	Mean HHS 96.4
Munro 2014 [247]	19 THA	37% (NS) *	57.5 (46-76)	2.8 (0.6-4.9) *	2.1 (0.8-4.0) *	68% (13)	21% (4)	Dislocation and/or loose cup (3) ARMD recurrence (1)	Mean WOMAC (pain) 78 & (function) 83
Pritchett 2014 [248]	90 HRA	48% (43)	49.8 (32-71)	2.8 (1.3-4.9)	5.1 (3.0-9.8)	4% (4)	3% (3)	ARMD recurrence (1) Infection (1) Loose cup (1)	Mean HHS 93.2
Matharu 2014 [195]	46 HRA 18 THA	72% (46)	57.8 (31-79)	5.5 (1.1-13.8)	4.5 (1.0-14.6)	20% (13)	13% (8)	Dislocation (2) ARMD recurrence (2)	Median OHS 39

Continued /...

Study author and year	Hips revised for ARMD	Female hips (n)	Mean age (range) at revision in years	Mean time to revision (range) in years	Mean follow-up time after revision (range) in years	Risk of complications (n)	Risk of re-revision (n)	Main reasons for re-revision surgery (n)	Functional outcome **
Norris 2014 [251]	35 HRA	71% (25)	58.0 (30-76)	4.3 (1.5-9.6)	NS	NS	NS	NS	Mean OHS 33
Cip 2015 [252]	20 THA	47% (NS) *	49.6 (21-61) *	4.6 (2.7-6.7) *	2.3 (1.5-3.1)	10% (2)	5% (1)	Infection (1)	Mean HHS 85.1
Stryker 2015 [253]	58 THA	65% (NS) *	60.0 (17-84) *	3.9 (0.1-9.5) *	1.2 (0-10.2) *	20% (23)	16% (18)	Infection (7) Loose cup or stem (6) Dislocation (4)	NS
Lainiala 2015 [191]	49 HRA 166 THA	60% (130)	62.1 (SD 10.1)	4.7 (SD 1.3)	2.3 (1.0-NS)	5% (11)	3% (6)	Dislocation (4) Infection (1)	Median OHS 40
Van Lingen 2015 [254]	38 THA	69% (NS) *	63.0 (44-75) *	Median 3.7 (1.0-6.5) *	3.1 (2.1-4.7) *	24% (9)	8% (3)	Dislocation (3)	Mean HOOS 61.9 *
Liow 2016 [255]	25 HRA 77 THA	36% (35)	62.0 (41-85)	5.1 (1.4-18.3)	2.5 (2.2-4.3)	14% (14)	7% (7)	ARMD recurrence (3) Dislocation (2) Loose cup (2)	Mean HSS 75.6
Matharu 2017 [256] *Updated report on Grammatopoulos 2009* [12]	16 HRA	100% (16)	51.3 (20-71)	1.6 (0.01-6.7) *	Median 10.3 (7-15) *	69% (11)	44% (7)	Dislocation +/- ARMD recurrence (5) Loose cup (2)	Median OHS 21

Table 2.4 Studies reporting the outcomes following metal-on-metal hip arthroplasty revision surgery performed for adverse reactions to metal debris.
ARMD = adverse reactions to metal debris; HHS = Harris Hip Score; HOOS = hip disability and osteoarthritis outcome score; HRA = hip resurfacing arthroplasty; NS = not stated; OHS = Oxford Hip Score; SD = standard deviation; THA = total hip arthroplasty; WOMAC = Western Ontario and McMaster Universities Arthritis Index.
*** Studies did not provide the relevant data specifically for the cohort of patients undergoing revision for adverse reactions to metal debris (but rather for the whole cohort of metal-on-metal hip arthroplasty revisions that they reported on).**
**** Functional outcome scoring systems: OHS = 0-48 (48 best outcome) [257, 258]; HHS = 0-100 (100 best outcome) [259]; HOOS = 0-100 (100 best outcome) [260]; WOMAC = 0-100 (0 best outcome) [261].**

2.4.3.2. Complications and re-revision surgery

The risk of complications (up to 68%) and re-revision surgery (up to 38%) were variable amongst the short-term studies but generally high (Table 2.4.). In the first 16 ARMD revision procedures performed at one centre there was an increase in the risk of complications (50% to 69%) and re-revision (38% to 44%) from the initial study at a mean of three years follow-up [12] to the report at a median of 10.3 years [256]. Some studies did report implant survival rates following ARMD revision, which were 88% at both three years [255] and five years [195], and 56% at ten years [256].

Early observations suggested short-term outcomes following ARMD revision were poor, with half of patients sustaining major complications and over one-third requiring further operations [12]. Similar observations were reported in subsequent small cohorts reporting their initial experience following ARMD revision surgery [247, 249]. Furthermore the risk of complications and re-revision following the first ARMD revision cases were higher compared with MoMHA revisions for non-ARMD indications (fracture, loosening, infection), and compared with matched patients undergoing primary conventional THA [12]. The poor outcomes following ARMD revision were thought to relate to the invasive and destructive nature of these lesions [12, 14, 16, 48]. In addition to regular MoMHA patient surveillance these initial observations led regulatory authorities and orthopaedic surgeons to widely recommend performing early revision in MoMHAs with ARMD [12, 23, 24, 29, 48]. It was thought that this strategy would improve outcomes following ARMD revision. Therefore over time surgeons subsequently adopted a lower threshold for performing revision for ARMD.

The impact of performing early surgery for ARMD has not yet been established in large patient cohorts. Although there is some evidence that this strategy can improve outcomes it is likely to be multifactorial. One single surgeon series reporting outcomes following 48 ARMD revisions observed that the risk of complications and re-revisions were significantly reduced in the latest 31 cases compared with the first 17 revisions [29]. In addition to lowering the threshold for performing ARMD revisions in the latter cases, the surgeon made numerous other changes to their practice. This included more intensive surveillance with routine blood metal ion sampling, revising both HRA components, implanting larger ceramic femoral head sizes, and post-operatively using anti-dislocation braces. Furthermore outcomes following ARMD revision were not significantly inferior compared with the 65 non-ARMD revisions performed [29]. Although this was contrary to earlier observations [12], it is of interest that when this initial cohort were reviewed at extended follow-up there was no longer a significant difference in the risk of complications and re-revision following ARMD revisions compared with non-ARMD revisions [256].

There appears to be a general trend of a reduced risk of complications and re-revision in more recent studies with short-term follow-up (Table 2.4.). These improved outcomes compared with the initial experience [12] are suspected to relate to a number of the factors described, but namely regular patient surveillance and lowering of the threshold for performing ARMD revision surgery. However surgeon experience and ARMD revision case volume are likely to be strongly associated with subsequent outcomes [29], given that two of the largest studies had the lowest risk of complications and re-revision [191, 248]. One of these studies was a single surgeon series reporting the outcomes of

90 ARMD revisions [248], and the other was the largest published study involving 215 ARMD revisions performed by six surgeons [191]. Therefore on this basis it is suggested that to obtain better outcomes ARMD revisions should be performed in specialist centres by experienced and high volume MoMHA revision surgeons [29, 203].

In line with the initial systematic review [28], the commonest indications for re-revision surgery following ARMD revision in the included studies were dislocation, ARMD recurrence, aseptic component loosening, and infection (Table 2.4.). Many of these failure mechanisms are suspected to be multifactorial.

Dislocation can be related to ARMD lesions that are destructive and/or require extensive soft-tissue debridement. Both can compromise the hip abductors and/or short external rotators and therefore lead to hip instability [29]. Other risk factors for dislocation include reducing the large-diameter MoM bearing surface when exchanging to a non-MoM articulation, and not correcting any residual component malposition at ARMD revision (for example when performing modular revisions in MoM THAs) due to the potential risks associated with removing well-fixed implants [253]. Using large femoral head sizes ($\geq$36 mm) at ARMD revision have not prevented dislocations [247]. This suggests that instability following ARMD revision is a complex problem, with some surgeons using anti-dislocation braces, dual-mobility acetabular components, and an overall lower threshold for performing ARMD revision procedures in an attempt to reduce the subsequent dislocation risk [29, 248].

Most studies have reported at least one case of ARMD recurrence following ARMD revision (Table 2.4.). Recurrence can occur due to incomplete excision of metal debris

at revision. Reasons for incomplete excision include surgical error, and being unable to excise all the metal debris because of its proximity to vital neurovascular structures [12, 203]. In such cases serial post-revision cross-sectional imaging can be useful for monitoring progression of any residual ARMD [12, 247]. Recurrence of ARMD can also occur if there is another potential source of metal wear debris and/or corrosion, including implantation of another MoM bearing [195, 203, 248] or using a cobalt-chrome femoral head on a titanium alloy or cobalt-chrome femoral stem [87, 93, 247]. In light of these observations most surgeons now use ceramic femoral heads when performing ARMD revisions [12, 29, 203], with a taper adapter used if the stemmed femoral component is retained given the theoretical risk of ceramic head fracture when impacted onto retained tapers [93].

Aseptic loosening, usually of the acetabular component, has also frequently been reported following ARMD revision. This may be due to invasive and destructive ARMD causing significant osteolysis and/or iatrogenic bone loss when removing a well-fixed component, which increases the risk of failed implant osseointegration. One study experienced a number of failures due to aseptic acetabular loosening despite the absence of bone defects, however this was thought to be due to the fibre metal-backed acetabular components implanted at revision [247].

The development of deep infection following ARMD revision can be multifactorial. Risk factors for infection may include multiple operations, incomplete excision of metal debris or necrotic tissue at revision, and the retention of primary MoMHA components [203, 247].

2.4.3.3. Functional outcomes

Patient-reported outcomes reported following the initial ARMD revisions were poor [12, 90, 247], with these also being inferior compared with MoMHA revisions for non-ARMD indications, and compared with primary conventional THAs [12]. The poor functional outcomes reported in an early report following 16 ARMD revisions [12] persisted at extended follow-up, and remained inferior compared with MoMHA revisions for non-ARMD indications [256]. These poor functional outcomes were also considered a consequence of the invasive and destructive nature of ARMD lesions, with a lower threshold for performing such revision procedures widely adopted [12, 23, 24, 29, 48].

The majority of subsequent studies have reported good or excellent functional outcomes following ARMD revision surgery [29, 191, 195, 203, 248-250, 252], defined as a mean Oxford Hip Score (OHS) of 34 or above or a Harris Hip Score (HHS) of 80 or above [258, 262]. However not all of the more recent studies achieved good functional outcomes following ARMD revision [251, 254, 255], which suggests other factors may influence outcomes after these procedures.

2.4.3.4. Predicting re-revision and guiding surgical decisions

Surgeons currently experience difficulty when managing MoMHA patients [245, 246]. Identifying prognostic factors of outcome following ARMD revision would provide surgeons with thresholds for performing revision, and help guide decisions regarding the type of reconstruction to perform. Large cohort studies would provide adequate statistical power to address these important clinical issues. However the only large cohorts reporting on outcomes following MoMHA revision surgery did not specifically

report outcomes following ARMD revision [243, 263, 264]. The evidence regarding factors predictive of poor outcomes following ARMD revision surgery is therefore limited to small studies that were not adequately powered to identify prognostic factors [29, 195, 203, 255].

In some of the earliest MoMHAs revised for ARMD a significantly higher risk of re-revision and inferior functional outcomes were reported when another MoM bearing was implanted at revision [195, 203, 248]. This is suspected to be due to the continued source of metal wear debris and/or corrosion, and has led to using non-MoM bearing surfaces at ARMD revision. Solid pseudotumours have been a risk factor for re-revision surgery following ARMD revision [203]. Similarly a recent study of 102 ARMD revisions identified MARS-MRI evidence of solid lesions with abductor deficiency to be predictive of post-revision complications, with pre-revision radiographic implant loosening also predicting poor outcomes [255]. However this study did not find other factors to be predictive of post-revision complications, such as pre-revision blood metal ions, type of revision articulation and femoral head size [255]. It is therefore important for these and other pre-revision and intra-operative factors to be assessed in large patient cohorts undergoing ARMD revision surgery, as this will allow identification of any important prognostic factors of outcome, which could be used to guide patient management.

2.5. Regulatory authority guidance for patient surveillance

2.5.1. Variation in regulatory authority guidance

Guidance for MoMHA patient follow-up has been issued by various worldwide regulatory authorities and has been summarised (Table 2.5.). These authorities include

the UK Medical and Healthcare products Regulatory Agency (MHRA) [23, 184], the European Federation of National Associations of Orthopaedics and Traumatology [22, 265], the US Food and Drug Administration (FDA) [24], Therapeutic Goods Administration of Australia [142], and Health Canada [141]. Major differences exist in follow-up guidance between authorities, which will now be considered. It is acknowledged that the guidance issued by the MHRA is currently being updated.

	MHRA UK	European	FDA US	TGA Australia	Health Canada
Distinguishes between HRA and THA	Yes	Yes	No	Yes	No
Follow-up protocol	All THA ≥36mm + symptomatic HRA – annually for implant life Asymptomatic HRA* - as per local protocol	All THA ≥36mm and HRA with risk factors**** – annually for implant life All HRA without risk factors - annually for first 5 years (then as per local protocol)	All MoM hips*** Asymptomatic = every 1 to 2 years Symptomatic = at least every 6 months	All MoM hips with symptoms, & asymptomatic THA ≥36mm or HRA≤45mm – at least annually Other MoM hips with no symptoms – as per practice for non-MoM hips	All MoM hips with symptoms – no guidance given on regularity of follow-up All MoM hips without symptoms - annually for first 5 years (then as per local protocol)***
Follow-up for symptomatic patients	Ions + imaging	X-ray + ions + imaging	X-ray + ions + imaging	X-ray + ions + imaging	X-ray + ions + imaging
Follow-up for asymptomatic patients	THA = ions** HRA = see above	All MoM hips = x-ray + ions Further imaging if x-ray abnormal or Co between 2-7µg/l	Clinical review	Asymptomatic THA ≥36mm or HRA≤45mm = x-ray + ions + imaging Other MoM hips with no symptoms (see above)	Clinical review
Blood metal ion sampling	Whole blood (Co and/or Cr)	Whole blood (Co only)	Whole blood (Co and/or Cr)	Whole blood or serum (Co and Cr)	Whole blood or serum (Co and Cr)
Blood metal ion thresholds of concern	>7 µg/l	2-7 µg/l	None stated	None stated	>7 µg/l
Plain radiographs recommended	Not stated	All patients	Symptomatic patients only	All patients	Symptomatic patients only

Continued /...

	MHRA UK	European	FDA US	TGA Australia	Health Canada
Cross-sectional imaging recommended	MARS-MRI or ultrasound	MARS-MRI or ultrasound or CT	MARS-MRI or ultrasound or CT	MARS-MRI or ultrasound	MARS-MRI or ultrasound
Consider revision surgery	If imaging abnormal and/or blood metal ion levels rising	(1) If imaging abnormal and/or blood metal ion levels raised or rising (2) If Co >20µg/l	Decide in response to overall clinical scenario and test results, but consider early revision in patients with progressive lesions	If persistent symptoms, imaging abnormalities and/or where blood metal ions are rising	If symptoms and positive MARS-MRI (soft-tissue mass) If positive MARS-MRI (soft-tissue mass), increasing in size

Table 2.5 Follow-up guidance for metal-on-metal hip arthroplasty patients published by worldwide regulatory authorities.
Cr = chromium; Co = cobalt; CT = computed tomography; FDA = Food and Drug Administration; HRA = hip resurfacing arthroplasty; MHRA = Medical and Healthcare products Regulatory Agency; MARS-MRI = metal artefact reduction sequence magnetic resonance imaging; MoM = metal-on-metal; TGA = Therapeutic Goods Administration; THA = total hip arthroplasty; UK = United Kingdom; US = United States.
*** Excludes all asymptomatic Articular Surface Replacement hip resurfacings, and Birmingham Hip Resurfacings implanted in all females and in males with femoral head sizes below 48 mm.**
**** Imaging recommended if blood metal ion concentrations rising (also all asymptomatic Articular Surface Replacement XL patients require cross-sectional imaging).**
***** Advises closer follow-up for patients at increased risk of device wear such as females, those with bilateral implants, suboptimal component alignment, or hip resurfacings with small femoral head sizes (less than or equal to 44 mm).**
****** Risk factors include small femoral head size (<50 mm), female gender, and low coverage arc.**

2.5.1.1. Type of metal-on-metal hip arthroplasty

The FDA and Canadian authorities do not distinguish between the follow-up of HRAs and MoM THAs. The other three authorities (MHRA, Australian, European) make this distinction and follow-up large-diameter (≥36 mm) MoM THAs regardless of symptoms [22, 23, 142]. Furthermore, these three authorities stratify HRA follow-up by symptoms and potential ARMD risk factors. Small HRA femoral head sizes are a high-risk group requiring at least annual surveillance in European (<50 mm) and Australian (≤45 mm) guidance. More recently the MHRA recommended that BHRs in all female patients and in male patients with small femoral head sizes (<48 mm) represent a high-risk group requiring continued surveillance regardless of symptoms [184]. Although FDA and Canadian guidance stratifies follow-up according to symptoms alone, both advise closer follow-up (review intervals not stated) for patients with ARMD risk factors, such as females and those with bilateral MoMHAs.

2.5.1.2. Follow-up regularity

Only the FDA advocate universal follow-up for all MoMHAs for the implant lifetime (six monthly reviews if symptomatic, and one to two yearly if asymptomatic). All other authorities recommend at least annual follow-up for most MoMHAs, including large-diameter MoM THAs, HRAs with ARMD risk factors, and all symptomatic patients. Asymptomatic HRA patients and those without ARMD risk factors are either reviewed according to local protocol, or annually for five years followed by local protocol.

2.5.1.3. Investigative modalities

All authorities stratify investigations according to patient symptoms. Symptoms are defined by three authorities as pain and abnormal gait (including limping) [23, 24, 141], whilst European and Australian guidance do not define symptoms. The FDA and Canadian guidance further define symptoms as noises from the hip, decreased range of motion, swelling, local nerve palsy, and dislocation. Although patient reported outcome measures such as the OHS [258] are reliable and responsive instruments, no guidance recommends their use.

For symptomatic MoMHA patients, blood metal ions and cross-sectional imaging are universally recommended. For asymptomatic MoMHAs recommendations include clinical review only [24, 141], radiographs with metal ions [22], metal ions alone for MoM THAs [23], and radiographs, metal ions, and cross-sectional imaging for all MoM THAs and for HRAs with small femoral head sizes [142].

All authorities recommend whole blood for determining metal ion concentrations, with serum also acceptable in Australia and Canada. European guidance requires measurement of cobalt only, however all other authorities recommend sampling both cobalt and chromium.

The MHRA make no recommendations regarding hip radiographs, whilst the FDA and Canada suggest radiographs in symptomatic patients only. All authorities advocate either MARS-MRI or ultrasound for cross-sectional imaging, however both the FDA and European guidance consider CT to also be acceptable as an initial cross-sectional imaging modality.

2.5.1.4. Blood metal ion thresholds

The UK MHRA was first to publish blood metal ion thresholds for concern in 2010 [67]. The blood cobalt and/or chromium threshold of above 7 µg/l continues to be recommended in more recent MHRA publications, with repeat testing recommended within three months of abnormal results [23, 184]. Canadian guidance also recommends blood metal ion concentrations above 7 µg/l as a cause for concern. European guidance suggests asymptomatic patients with blood metal ion concentrations between 2 µg/l and 7 µg/l require cross-sectional imaging. No blood metal ion thresholds for concern are presently recommended by the FDA and Australia.

2.5.1.5. Threshold for revision surgery

All authorities suggest revision should be considered in patients with abnormal cross-sectional imaging and/or raised blood metal ion concentrations, especially if imaging abnormalities are progressing and/or blood metal ions are rising. European guidance recommends blood cobalt concentrations above 20 µg/l as an indication for considering revision [22].

2.5.2. Regulatory authority guidance versus available evidence

As major differences exist in MoMHA patient follow-up guidance between regulatory authorities, it is important to consider the guidance issued in relation to the best available evidence (Chapters 2.1, 2.2, 2.3, and 2.4.).

2.5.2.1. Do all patients require follow-up?

Evidence clearly demonstrates a higher failure rate for MoM THAs over HRAs, so this distinction should be made in surveillance protocols. By contrast, young and active male patients with BHRs have a very low risk of failure and therefore need less regular surveillance than other HRA patients (Chapter 2.1.2.).

It is concerning that FDA and Canadian guidance fails to separate follow-up protocols for MoM THAs and HRAs. European and Australian guidance risk stratifies HRAs by patient and implant factors, which is advisable given current evidence (Chapter 2.2.6.). Although many BHR patients now require regular surveillance given recent changes to the MHRA guidance [184], the MHRA have not modified surveillance for all other non-ASR HRA devices with these patients continuing to be reviewed as per local protocol if asymptomatic. This contradicts the evidence given that the BHR has consistently performed better than all other HRA designs worldwide (Chapter 2.1.2.).

2.5.2.2. Should symptoms decide follow-up?

Although all authorities stratify follow-up according to symptoms, asymptomatic ARMD has been observed in MoMHA patients with a number of these eventually requiring revision (Chapter 2.3.3.). Furthermore most MoMHA patients remain asymptomatic [13, 68, 72, 102, 135, 154], even those with high-risk implants like the ASR [168, 186]. Stratifying follow-up according to symptoms may result in ARMD being undetected in asymptomatic patients, which could have serious consequences. Of particular note the FDA and Canadian guidance allocates limited resources for asymptomatic MoM THA patients who are known to be at increased risk of failure (Chapter 2.1.2.).

2.5.2.3. What is the role of blood metal ion sampling?

The role of blood metal ion sampling has been described (Chapter 2.3.1.). Variable information has been provided by regulatory authorities regarding blood metal ion analysis. However most do recommend serial blood samples undergo analysis at the same accredited laboratory, and using the same blood fraction.

Present evidence suggests that no single blood metal ion threshold will reliably identify ARMD. However the 7 µg/l threshold recommended by the MHRA and Canada is too high. Evidence suggests the threshold for concern in unilateral MoMHA patients should be below 5 µg/l (Chapter 2.3.1.4.). European guidance recognises the threshold should be between 2 µg/l and 7 µg/l with the exact level still to be determined.

2.5.2.4. What is the role of cross-sectional imaging?

The role of cross-sectional imaging has been described (Chapter 2.3.3.). All regulatory authorities recommend ultrasound or MARS-MRI can be used as the initial cross-sectional imaging modality with no preference stated. This is appropriate, as presently no consensus exists regarding which modality is most effective for identifying ARMD or monitoring progression. By contrast, CT is not suitable for routine MoMHA imaging. Therefore the FDA and European guidance should not recommend CT as a modality for initial cross-sectional imaging.

2.5.2.5. When should follow-up be repeated?

Decisions regarding follow-up regularity require well-designed longitudinal cohort studies with patients undergoing reassessment. Current longitudinal studies have largely involved small cohorts with short-term follow-up (Chapter 2.3.). Therefore it remains unclear how regularly patients with investigative abnormalities should be reviewed, and whether patients with normal investigations can be safely discharged, or whether they need repeat assessment.

Repeat blood metal ion sampling in MoM THA patients may be useful, but no more regularly than annually. In HRAs with low blood metal ion concentrations, repeated testing in the short-term appears unnecessary without clinical changes. Repeat blood sampling within three months for patients with high initial concentrations as recommended by some authorities [23, 141] appears unnecessary, unless there is suspicion high readings may be spurious or involve contaminated samples. Repeat cross-sectional imaging in MoMHA patients should not be performed more frequently than annually based on current evidence. The costly FDA surveillance for symptomatic patients (six monthly) is therefore not justified.

2.5.2.6. What is the threshold for revision surgery?

Robust thresholds for performing revision for ARMD in MoMHA patients have not been established due to a lack of evidence (Chapter 2.4.1.). The current recommendations from all regulatory authorities regarding the threshold for revision (Chapter 2.5.1.5.) appear reasonable until prognostic factors of outcome following revision surgery are identified. The European recommendations for considering revision based on blood cobalt concentrations above 20 µg/l alone were derived from

one single-centre study which observed such cases were frequently associated with osteolysis, even in asymptomatic patients [186].

2.5.3. Financial analysis of regulatory authority guidance

Using MoMHA follow-up guidance from worldwide authorities (Table 2.5.) a financial analysis was performed for four common clinical scenarios, with potential limitations relating to this analysis also detailed (Appendix 3.). For symptomatic MoM THAs and symptomatic HRAs, annual follow-up costs per MoMHA patient were identical within each authority. However costs varied considerably between authorities, with FDA follow-up costs (£988) three-times those of the MHRA (£330). For asymptomatic MoM THA patients cost variations were even greater, with Australian costs (£494) nearly six-times more expensive than those in FDA and Canada (£84). In asymptomatic HRA patients costs ranged from those for local protocols (MHRA and Australia) up to £278 for European guidance.

To more accurately estimate the financial impact of follow-up, NJR data was used which included 67,363 MoMHAs (35,470 HRA and 31,893 MoM THA) implanted in England and Wales [2], with a financial analysis performed using the methods described (Appendix 3.). Massive cost differences existed between authorities when considering national surveillance (Table 2.6.). MHRA guidance was cheapest for overall MoMHA annual follow-up (£8,264,064). European (£22,708,226) and the FDA (£22,321,020) protocols were the most expensive, both approaching three-times the MHRA costs. The FDA protocol was the most costly for surveillance of all symptomatic MoMHA patients (£18,210,816), Australian guidance was most expensive for asymptomatic MoM THAs (£9,453,184), and the European guidance was most

costly for asymptomatic HRAs (£8,283,010). The European follow-up protocol was most expensive overall given most HRAs (84%) were asymptomatic.

	MHRA UK	Europe	FDA US	TGA Australia	Health Canada
MoM THA (≥36mm) Symptomatic (n = 12,757)	£4,209,810	£6,301,958	£12,603,916	£6,301,958	£6,301,958
MoM THA (≥36mm) Asymptomatic (n = 19,136)	£2,181,504	£5,319,808	£1,607,424	£9,453,184	£1,607,424
HRA Symptomatic (n = 5,675)	£1,872,750	£2,803,450	£5,606,900	£2,803,450	£2,803,450
HRA Asymptomatic (n = 29,795)	£0	£8,283,010	£2,502,780	£0	£2,502,780
Total annual cost for follow-up of all MoM hip arthroplasties using each protocol	£8,264,064	£22,708,226	£22,321,020	£18,558,592	£13,215,612

Table 2.6 Differences in annual follow-up costs for all metal-on-metal hip arthroplasties recorded in the National Joint Registry when using various worldwide surveillance protocols.
FDA = Food and Drug Administration; HRA = hip resurfacing arthroplasty; MHRA = Medical and Healthcare products Regulatory Agency; MoM = metal-on-metal; TGA = Therapeutic Goods Administration; THA = total hip arthroplasty; UK = United Kingdom; US = United States.

2.5.4. Conclusions regarding currently proposed surveillance

Serious concerns were raised regarding patient safety when high MoMHA failure rates were publicised [14, 16, 67, 266]. Medico-legal pressures have compounded the problem leading worldwide regulatory authorities to issue guidance based on the available evidence. Generally follow-up guidance was conservative to reflect these

concerns, but also because ARMD was not well understood. Over time the understanding of ARMD has improved, although it remains incomplete.

Substantial differences exist between recommended surveillance protocols, which no longer reflect current evidence. These include not stratifying patients according to implant type (MoM THA vs. HRA) or ARMD risk factors, placing little emphasis on asymptomatic patient follow-up, and recommending suboptimal blood metal ion thresholds. Authorities must address these issues when updating their guidance, which should help ensure follow-up becomes more unified between authorities.

The inadequacy of current guidance is highlighted when considering MoMHA patients with asymptomatic ARMD. Four authorities would not identify HRA patients whilst asymptomatic, as further investigations would not be indicated [23, 24, 141, 142]. European guidance would identify patients if blood cobalt was above 2 µg/l [22]. A similar situation occurs with asymptomatic MoM THAs with ARMD, with Australian guidance being most likely to identify such patients [142]. Consideration should therefore be given to performing baseline investigations in most MoMHA patients. This would help risk stratify patients and make decisions regarding subsequent surveillance; however further research is required to inform this.

The financial implications of current MoMHA surveillance protocols are vast with some countries spending almost three times more on annual follow-up than others. Costs from the presented analysis are likely conservative given they do not consider the full logistical and resource implications to institutions (extra staff and clinics required, and time taken to perform and report investigations). Despite the almost global

cessation of MoMHA, current follow-up protocols are clearly not financially sustainable. What is known is that certain patients, namely male BHR patients with no ARMD risk factors, do not require intense and costly follow-up recommended by some authorities, such as the European guidance [22].

A number of questions still need to be addressed to formulate evidence-based guidance for the investigation and management of MoMHA patients. These specifically relate to which patients need surveillance, identifying the role of each investigative modality, establishing the regularity and duration of follow-up for different patient subgroups, and providing surgeons with information to guide decisions about the threshold and type of ARMD revision surgery required. The seven specific objectives described in this thesis cover these particular aspects of clinical care (Chapter 1.4.). It is anticipated that formulating such an evidence-based approach would reduce current anxieties and concerns in MoMHA patients who do not need regular review, and allow resources to be concentrated towards patients requiring surveillance or revision surgery. It is also expected developing more evidence-based management protocols would protect surgeons and institutions from a medico-legal perspective.

2.6. Summary

- In recent years MoMHAs have been associated with high short-term revision rates. Many revisions have been performed for ARMD and pseudotumour, which were unanticipated complications associated with MoMHA. However revision rates for these problems in the second decade are not yet known.

- To identify ARMD and pseudotumour early worldwide regulatory authorities have published follow-up guidance for patients with MoMHAs. Most patients require some form of annual surveillance for the implant lifetime, which includes blood metal ion sampling and hip imaging.

- Major differences exist in the follow-up guidance issued by five worldwide regulatory authorities for MoMHA patients. The proposed surveillance protocols are currently not evidence-based and very costly. Furthermore surgeons currently have little robust information to counsel patients about the risks of undergoing revision surgery for ARMD.

- To provide effective follow-up for MoMHA patients it is important to determine which patients need surveillance, identify the role of each investigative modality (including the blood metal ion thresholds for clinical concern), and establish the regularity and duration of follow-up for different patient subgroups.

- For managing MoMHA patients, detailed knowledge of outcomes following revision would allow surgeons to more informatively counsel patients about the risks of undergoing further surgery. Furthermore identifying prognostic factors of outcome following ARMD revision would provide surgeons with thresholds for performing revision, and help guide decisions regarding the type of surgery required.

3. CLINICAL IMPACT OF PSEUDOTUMOURS*

3.1. Introduction

Revision rates specifically for ARMD and pseudotumour within ten years of HRA have been high, with evidence that the prevalence of this complication is increasing (Chapter 2.2.5.). However revision rates in the second decade for these endpoints remain unknown.

Risk factors reported for both all-cause HRA revision and pseudotumour revision have previously been investigated (Chapter 2.2.6.). Gender and femoral head size have complex interactions, but it would appear that being female is the strongest risk factor for pseudotumour revision rather than small femoral HRA component size [69, 132]. The MHRA recently recommended against future BHR implantation in all females and in males requiring small femoral components (<48 mm), with regular surveillance advised for these subgroups [184]. However updated guidance was not issued for the surveillance of other non-ASR HRA designs [184].

Although large cohorts have assessed revision rates and predictors of HRA revision they are limited by short-term follow-up (mean 3.4-7.1 years), [10, 76, 132, 139, 144, 267] with only one study specifically investigating pseudotumour revision rates and risk factors [132]. Furthermore previous study limitations may render the reported predictors of pseudotumour revision inaccurate. Knowledge of the correct predictors of

* Work presented in this chapter has been published in:
Matharu GS, Judge A, Murray DW, Pandit HG. Prevalence of and risk factors for hip resurfacing revision: a cohort study into the second decade after the operation. *J Bone Joint Surg Am* 2016; **98**(17): 1444-1452.

pseudotumour revision are important to inform follow-up regimens and stratify patients for surveillance given the vast number of HRAs implanted worldwide [2, 11, 27]. Registry studies provide large numbers but have numerous limitations (Chapter 2.2.5.). Therefore, large cohort studies with detailed data and revision indications determined histopathologically would ideally be suited to determine pseudotumour revision rates.

3.2. Objectives

The objectives of this chapter were to determine the 15 year revision rate for pseudotumour in HRA patients at a single specialist arthroplasty centre, and identify risk factors for pseudotumour revision.

3.3. Patients and methods

3.3.1. Study design

This prospective single centre cohort study was performed at the Nuffield Orthopaedic Centre (NOC), which is a specialist tertiary referral teaching hospital in Oxford in the UK. All patients receiving primary HRAs between 1st June 1999 and 31st December 2009 were eligible (n=1,473). Only patients with a minimum follow-up of 1 year were included (44 hips excluded). This was to allow all implants to have the potential to fail given that HRA revisions within the first year have been reported for all indications [2, 11, 76]. The present study was based on 1,429 primary HRAs (1,216 patients). The local ethical committee had granted a waiver for the follow-up of NOC hip arthroplasty patients. Furthermore HRA patients were reviewed more recently according to national MHRA recommendations [23].

3.3.2. Data source

Clinical data has been collected and stored on a prospectively maintained electronic research database (MySQL, Version 5.5.53, Oracle Corporation, Redwood Shores, California, US) at the NOC since the first HRA implantation in June 1999. All surgeons performing HRA at the NOC completed a paper data collection form at the time of operation (Appendix 4.). Data entered on these forms included patient age, gender, date of surgery, hip laterality, hip diagnosis, implant design, and femoral head size. These data were subsequently entered onto the electronic database by trained research assistants. Outcome data collected during patient follow-up (Chapters 3.3.4. and 3.3.5.) were also entered onto the database, including postoperative OHSs, blood metal ion results (sample date, and blood cobalt and chromium concentrations), and details of any revision procedures performed (date, indication, findings). Supplementary data for determining updated outcomes following HRA were obtained by searching the hospital electronic clinical records database (Electronic Patient Record (EPR)) and the electronic radiology database (Picture Archiving and Communication System (PACS), GE Healthcare, Barrington, Illinois, US) for all 1,429 cases (Chapter 3.3.5.).

3.3.3. Surgery performed

Seven non-designer surgeons performed or supervised all operations using one of four HRA designs: BHR (Smith & Nephew, Warwick, UK), Conserve Plus (Wright Medical Technology, Memphis, Tennessee, US), Recap (Biomet, Swindon, UK), and Cormet (Corin, Cirencester, UK). Implants were used according to surgeon preference. The BHR was used throughout 1999 to 2009, but was almost exclusively used between 1999 and 2004. A small number of Cormet implants were used between 1999 and 2001. The Conserve Plus and Recap devices were used from 2003 and 2005 respectively.

Broad indications were initially used for selecting patients for HRA. The procedure was offered to young active patients with hip arthritis, with no exclusions specifically made based on gender, femoral head size, hip diagnosis, or history of metal allergy [132]. All procedures were performed in a clean-air laminar flow operating theatre with the patient in the lateral decubitus position. Surgical approach was based on surgeon preference. Most operations were performed using an extended posterior approach as described by the designing BHR surgeon [43], with the remainder performed using an anterolateral approach. Acetabular (uncemented fixation) and femoral (cemented fixation) resurfacing components were implanted according to the respective manufacturer's guidance.

3.3.4. Patient follow-up

Routine follow-up of HRA patients at the NOC was performed at the following intervals postoperatively: six weeks, one year, five years, and ten years. For the early reviews (six weeks and one year) patients underwent clinical assessment (history, examination, anteroposterior pelvic radiographs, OHS questionnaire). The latter reviews (five and ten years) consisted of postal questionnaires requesting information on any further surgery performed (Appendix 5.) and completion of the OHS questionnaire [58, 69]. Symptomatic patients identified from these assessments underwent further investigations, including cross-sectional imaging.

The OHS is a validated patient reported outcome questionnaire used to measure pain and disability following hip arthroplasty [257]. The OHS comprises a 12-item questionnaire with patients selecting one of five possible responses for each question

(Appendix 6.). The modified OHS scoring system proposed by Murray *et al.* was used, with each question scored from 0 (worst outcome) to 4 (best outcome) therefore giving an overall possible OHS ranging from 0 (worst outcome) to 48 (best outcome) [258].

In light of the two MHRA alerts the NOC adapted its patient follow-up regimen for HRAs [23, 67]. In 2010 and 2012 all patients received a postal questionnaire enquiring about hip problems (pain, swelling, and noises), and were requested to complete the OHS questionnaire. Non-responders were sent another postal questionnaire, and if this was not returned they received a telephone call to complete the necessary data collection. Patients with hip problems and/or suboptimal OHSs ($\leq$41/48) [262] were reviewed as per MHRA recommendations [23] and underwent an updated clinical assessment (detailed above) in addition to cross-sectional imaging and blood metal ions. Subsequent follow-up was dependent on the findings from the postal questionnaire and any investigations performed. Patients were either discharged, kept under continued regular surveillance with or without further investigations, or underwent revision surgery. In addition, a number of asymptomatic HRA patients (201 hips in 158 patients) all underwent ultrasound and blood metal ion sampling as part of a previous ethically approved study [202].

3.3.4.1. Cross-sectional imaging

At the NOC ultrasound was used as the first line cross-sectional imaging modality with all scans performed by an experienced fellowship trained musculoskeletal radiologist. MARS-MRI was performed in patients whom: (1) ultrasound was equivocal but a pseudotumour or other hip pathology was suspected; and (2) ultrasound confirmed a

pseudotumour or other hip pathology, but more detailed anatomical information was required before revision.

3.3.4.2. Blood metal ion sampling

Whole venous blood was collected from the antecubital vein of patients for metal ion analysis. Blood was collected in trace element tubes containing sodium EDTA with methods employed to reduce the contamination risk [268]. All blood samples were analysed in an MHRA approved laboratory (Guilford Royal Surrey County Hospital Trace Element Laboratory, UK). This laboratory regularly participates in the Trace Elements External Quality Assessment Scheme, which has reported excellent accuracy of measurement and reproducibility [269]. Cobalt and chromium concentrations were measured using a quadrupole inductively coupled plasma mass spectrometer (Thermo Scientific XSeries 2, Thermo Fisher Scientific, Waltham, Massachusetts, US) (limit of detection 0.05 μg/l and reporting limit 0.35 μg/l).

3.3.5. Outcomes

The outcome of interest was revision surgery (defined as removal or exchange of any primary HRA component) performed for pseudotumour. The definition used for diagnosing an intraoperative pseudotumour was a cystic, solid, or mixed mass communicating with the hip joint [16, 216]. The diagnosis of pseudotumour was subsequently confirmed if there was also histopathological evidence of lymphocytic infiltrates (including ALVAL) and a phagocytic macrophage response to metal wear debris, with or without tissue necrosis [82, 98, 99].

All HRAs undergoing revision surgery were identified using the electronic HRA research database and from searching the hospital databases (EPR and PACS). When revisions had occurred at the NOC the patient records were reviewed to collect relevant details (including date, indication, surgical findings, components removed and implanted). Revision diagnoses were determined using pre-operative investigations (blood inflammatory markers and metal ion concentrations, pelvic radiographs, and cross-sectional imaging), operative records, and the results from histopathological and microbiological analyses performed on tissues excised at revision. If revision surgery was performed elsewhere, institutions were contacted to complete data collection. Revisions were known to have occurred elsewhere when: (1) patients informed us by telephone/post; (2) patients were reviewed for other reasons at the NOC and provided revision details; and (3) the primary surgeon identified revisions from their registry profile [2]. For assessing implant survival patients were censored on the date of revision surgery (regardless of indication), or, if not revised, on the date of their most recent review (clinical review or questionnaire completion). The last potential patient follow-up date for this study was 24th August 2015. No patients were awaiting revision surgery when this work was completed.

3.3.6. Statistical analysis

Potential predictors of HRAs undergoing pseudotumour revision surgery that were assessed included: age at primary, gender, unilateral/bilateral HRA(s), primary diagnosis (primary osteoarthritis, dysplasia, or other), HRA design (one of four designs) and femoral head size. Differences between covariates amongst revision and non-revision groups were assessed using two-sided unpaired t-tests (numerical data), and for

categorical data either the Chi-squared test or Fisher's exact test (latter used when the expected frequency in any cell was under five).

Survival analysis was performed using the Kaplan–Meier method. The endpoint for survival analysis was revision surgery performed for pseudotumour. Cox proportional hazards models (univariable and multivariable) were used to assess the association of predictor variables on time to pseudotumour revision. For continuous predictors (age and femoral head size), fractional polynomial regression modelling assessed the assumption of linearity with outcome, with data categorised if the assumption was not satisfied. The proportional hazards assumption was assessed for each predictor using Schoenfeld's residuals, with data grouped into discrete time periods if the assumption was violated. Likelihood ratio tests assessed evidence of two-way interactions between gender and other predictors. Due to the colinearity between gender and femoral head size, all subsequent analyses were stratified by gender. A univariable Cox regression model assessing the effect of year of primary HRA surgery on pseudotumour revision rates was also performed.

3.4. Results

3.4.1. Outcomes

There were 180 HRAs (156 patients) revised for all causes at a mean time of 4.8 years after arthroplasty (range 0.01-15.5 years) (Table 3.1.). The prevalence of all cause revision surgery was 12.6% (180/1,429).

Indication for revision surgery	Number of hips revised in whole cohort (%)	Number of hips revised in females only (%)	Number of hips revised in males only (%)
Pseudotumour	111 (61.7)	72 (67.3)	39 (53.4)
Femoral neck fracture	29 (16.1)	9 (8.4)	20 (27.4)
Aseptic loosening – acetabular	8 (4.4)	7 (6.5)	1 (1.4)
Unexplained pain	7 (3.9)	5 (4.7)	2 (2.7)
Deep infection	5 (2.8)	2 (1.9)	3 (4.1)
Aseptic loosening - femoral	4 (2.2)	3 (2.8)	1 (1.4)
Acetabular component malposition with or without femoral malposition	4 (2.2)	3 (2.8)	1 (1.4)
Dislocation	4 (2.2)	2 (1.9)	2 (2.7)
Avascular necrosis of the femoral head	4 (2.2)	3 (2.8)	1 (1.4)
Aseptic loosening - both components	2 (1.1)	1 (0.9)	1 (1.4)
Pain from impingement	1 (0.6)	0 (0.0)	1 (1.4)
Heterotopic ossification	1 (0.6)	0 (0.0)	1 (1.4)
TOTAL	180 (100)	107 (100)	73 (100)

Table 3.1 Indications for hip resurfacing arthroplasty revision surgery.

There were 111 HRAs (95 patients) revised for pseudotumour at a mean time of 6.1 years after arthroplasty (range 0.5-15.5 years). The prevalence of pseudotumour revision surgery was 7.8% (111/1,429). Pseudotumour was the commonest indication for revision surgery accounting for 61.7% of all revisions (Table 3.1.).

The remaining 1,249 HRAs (1,060 patients) remained in-situ at a mean follow-up time of 8.4 years after arthroplasty (range 1.0-15.4 years). The median OHS at this time was

46 (IQR 42-48). A follow-up questionnaire was completed for 94.9% (n=1,185) of the non-revised hips. Non-responders tended to be female and more commonly had Cormet implants (Appendix 7.).

3.4.2. 15 year revision rates

The cumulative all-cause HRA revision rate at 15 years was 19.5% (95% CI 16.2%-23.2%; 21 hips remained at risk at 15 years). The cumulative HRA revision rate for pseudotumour at 15 years was 14.0% (95% CI 11.0%-17.7%) (Figure 3.1.).

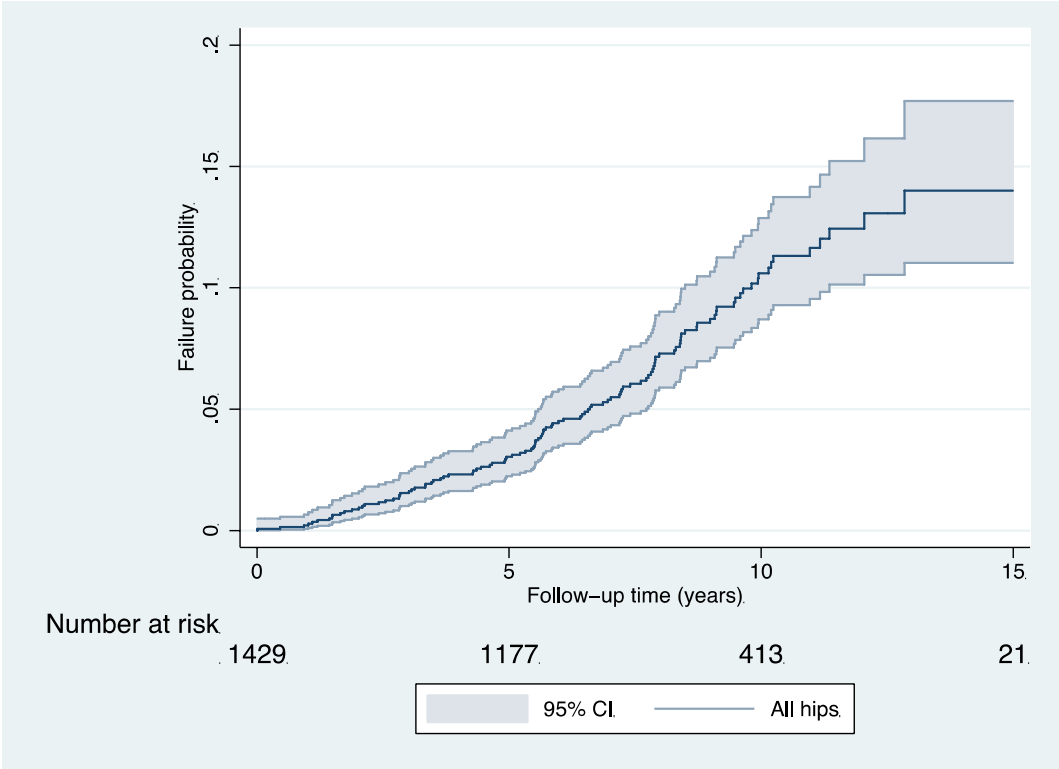


Figure 3.1 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years.
The number of hips at risk at five year intervals are illustrated.

3.4.3. Predicting pseudotumour revision: whole cohort

Differences between hips undergoing pseudotumour revision and those not undergoing pseudotumour revision are summarised (Table 3.2.). Female gender, young age, and small femoral head size were significant risk factors for pseudotumour revision surgery.

	Whole cohort (n=1429)		**No pseudo-tumour revision (n=1,318)**	**Pseudo-tumour revision (n=111)**	**p-value**
Gender					
Male	853 (59.7)		814 (61.8)	39 (35.1)	**<0.001**
Female	576 (40.3)		504 (38.2)	72 (64.9)	
Age (years)	53.8		54.0	51.5	**0.011**
Mean (range)	(16.5-85.5)		(16.5-85.5)	(28.7-69.9)	
Laterality					
Unilateral	1003 (70.2)		924 (70.1)	79 (71.2)	0.814
Bilateral	426 (29.8)		394 (29.9)	32 (28.8)	
Hip diagnosis					
Primary OA	1253 (87.7)		1155 (87.6)	98 (88.3)	0.439
Dysplasia	122 (8.5)		115 (8.7)	7 (6.3)	
Other	54 (3.8)		48 (3.6)	6 (5.4)	
Femoral head size					
(mm)	48.4		48.6	46.6	**<0.0001**
Mean (range)	(38-58)		(38-58)	(38-54)	
Implant					
BHR	641 (44.9)		591 (44.8)	50 (45.1)	0.481
Conserve Plus	631 (44.2)		584 (44.3)	47 (42.3)	
Recap	139 (9.7)		128 (9.7)	11 (9.9)	
Cormet	18 (1.3)		15 (1.1)	3 (2.7)	

Table 3.2 Patient and implant factors for the whole cohort, and for hip resurfacing arthroplasties not revised for pseudotumour compared with those revised for pseudotumour.
BHR = Birmingham Hip Resurfacing; OA = Osteoarthritis. P-values <0.05 in bold text. Values in brackets are percentages unless stated otherwise in the left hand column.

Univariable regression analysis demonstrated that factors significantly increasing the risk of pseudotumour revision surgery were female sex (hazard ratio (HR) = 2.86; 95% CI=1.93-4.22; p<0.001; Figure 3.2.), small femoral head size (HR per 2 mm = 0.88; 95% CI=0.84-0.92; p<0.001), young age at primary surgery (HR per year = 0.98; 95% CI=0.96-0.99; p=0.011), and Conserve Plus (HR=1.74; 95% CI=1.15-2.63; p=0.009) and Recap (HR=2.54; 95% CI=1.29-5.00; p=0.007) implants (both compared to BHR implants; Figure 3.3.).

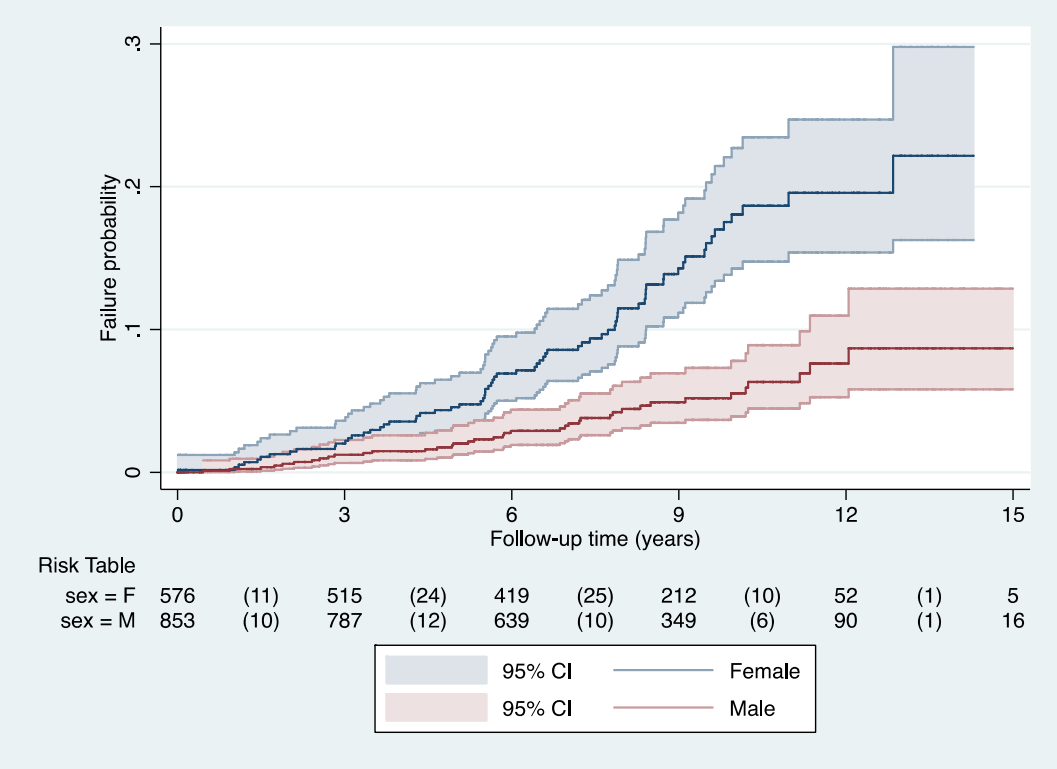


Figure 3.2 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years by gender.
Risk table indicates the number of hips at risk at three year intervals, with the corresponding number in brackets detailing the number of hips revised during the three year interval. The pseudotumour revision rate in females was 22.2% at 13 years (95% CI 16.3%-29.8%; 27 hips remained at risk) compared with 8.7% in males (95% CI 5.8%-12.9%; 57 hips remained at risk).

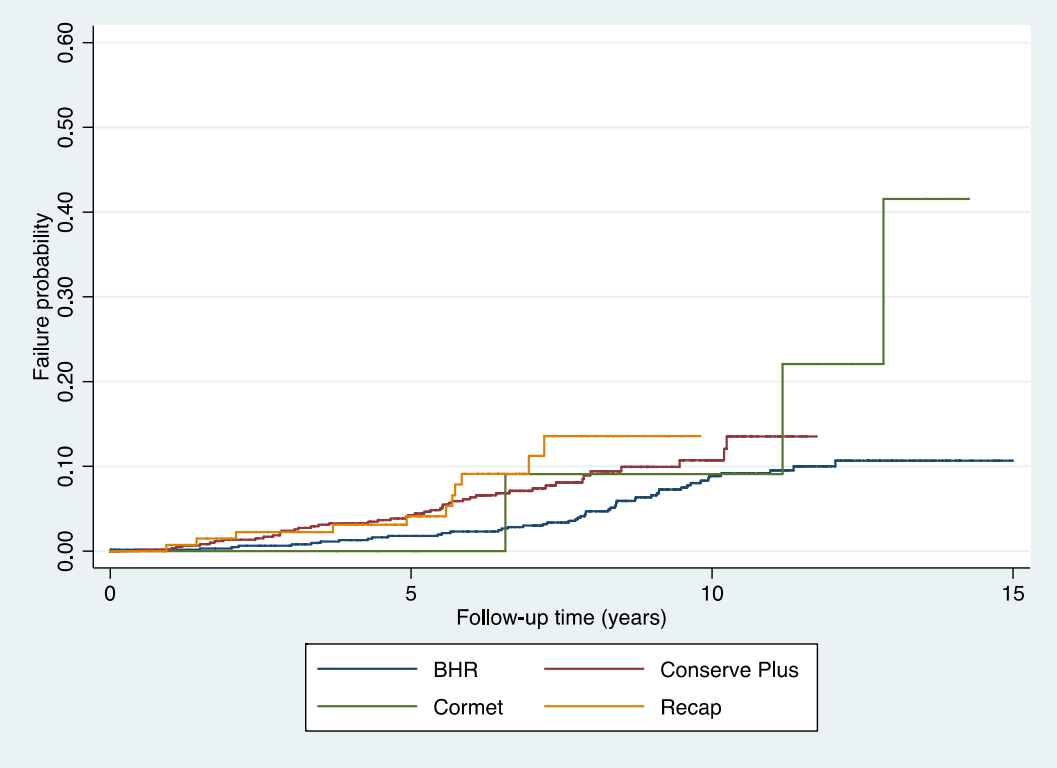


Figure 3.3 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to 15 years by implant design.
Confidence intervals have not been included for clarity. The nine year revision rates for pseudotumour by implant design were as follows: Birmingham Hip Resurfacing (BHR)=6.6% (95% CI=4.8%-9.0%; 414 hips remained at risk); Conserve Plus=9.9% (95% CI=7.3%-13.4%; 132 hips remained at risk); Recap=13.6% (95% CI=7.3%-24.4%; 9 hips remained at risk); Cormet=9.1% (95% CI=1.3%-49.2%; 10 hips remained at risk).

All factors identified from the univariable analysis remained significant predictors of pseudotumour revision risk in the multivariable regression analysis (Table 3.3.).

Covariate	Revision for pseudotumour							
	Whole cohort HR (95% CI)	**p-value**		**Female only HR (95% CI)**	**p-value**		**Male only HR (95% CI)**	**p-value**
Gender								
Female vs. Male	2.03 (1.19-3.44)	**0.009**		NA	NA		NA	NA
Age (per year)	0.98 (0.96-0.99)	**0.020**		0.97 (0.95-0.99)	**0.019**		0.98 (0.95-1.02)	0.370
Laterality								
Bilateral vs. unilateral	1.05 (0.69-1.61)	0.813		1.24 (0.74-2.09)	0.410		0.81 (0.38-1.73)	0.581
Hip diagnosis								
Primary OA	1.00	Ref		1.00	Ref		1.00	Ref
vs.								
Dysplasia	0.53 (0.24-1.18)	0.120		0.57 (0.25-1.30)	0.180		*	1.00
Other	1.51 (0.64-3.59)	0.348		1.68 (0.57-4.97)	0.351		1.23 (0.29-5.34)	0.778
Femoral head size (per 2 mm)	0.94 (0.88-0.99)	**0.045**		0.95 (0.87-1.03)	0.189		0.94 (0.84-1.04)	0.235
Implant								
BHR vs.	1.00	Ref		1.00	Ref		1.00	Ref
Conserve Plus	1.77 (1.14-2.73)	**0.011**		1.19 (0.70-2.05)	0.518		4.37 (1.86-10.26)	**0.001**
Recap	2.79 (1.40-5.55)	**0.003**		1.89 (0.72-4.94)	0.196		6.54 (2.14-19.97)	**0.001**
Cormet	2.40 (0.74-7.77)	0.143		1.86 (0.44-7.83)	0.400		4.12 (0.52-32.91)	0.181

Table 3.3 Multivariable Cox proportional hazards models for identifying hip resurfacing arthroplasties at risk of pseudotumour revision for the whole cohort, and subdivided by gender. BHR = Birmingham Hip Resurfacing; CI = Confidence interval; HR = Hazard ratio; NA = Not applicable; OA = Osteoarthritis; Ref = reference group. P-values <0.05 in bold text. *No hips in this particular subgroup.

3.4.4. Predicting pseudotumour revision: stratified by gender

Due to the colinearity between gender and femoral head size, and also because of significant interactions between gender and the other predictors the multivariable

regression analysis was repeated with females and males assessed separately. This demonstrated that the predictors of pseudotumour revision risk were gender specific (Table 3.3.). In females, young age at primary surgery was the only covariate significantly increasing pseudotumour revision risk (HR=0.97 per year; 95% CI=0.95-0.99; p=0.019). The ten year pseudotumour revision rates by age group in females were as follows: under 50 years=24.5% (95% CI=17.7%-33.4%; 56 hips remained at risk); 50-60 years=13.5% (95% CI=8.8%-20.4%; 60 hips remained at risk); over 60 years=16.0% (95% CI=9.3%-26.8%; 36 hips remained at risk). When the cohort had been stratified by gender small femoral head size was no longer a significant predictor of pseudotumour revision risk (female HR=0.95 per 2 mm; 95% CI=0.87-1.03; p=0.189).

In males, Conserve Plus (p=0.001) and Recap (p=0.001) implant designs were the only factors significantly increasing pseudotumour revision risk (Table 3.3.). The nine year pseudotumour revision rates by implant design in males were as follows: BHR=1.8% (95% CI=0.8%-4.1%; 254 hips remained at risk); Conserve Plus=7.9% (95% CI=4.9%-12.4%; 86 hips remained at risk); Recap=12.1% (95% CI=5.2%-26.8%; 7 hips remained at risk); Cormet=0% (95% CI=0%-0%; 6 hips remained at risk). The low pseudotumour revision rate in males with the BHR implant persisted until 15 years (4.6%; 95% CI=2.3%-9.2%; 17 hips remained at risk).

3.4.5. Effect of year of primary surgery

The effect of year of primary HRA surgery (before 2007 versus 2007 onwards) on pseudotumour revision rates was examined. This time point was clinically important as from 2007 onwards all consultant hip arthroplasty surgeons at the NOC were aware of

pseudotumours complicating HRAs with a number of these cases revised by this time [16]. Surgeons therefore managed patients differently from 2007 onwards and employed a lower threshold for undertaking further investigations, such as cross-sectional imaging, even in asymptomatic patients [16, 202].

There was a significantly increased risk of undergoing pseudotumour revision in HRAs implanted from 2007 onwards compared with those implanted prior to 2007 (HR=2.44; 95% CI=1.51-3.95; p<0.001). The eight year pseudotumour revision rates were 11.4% (95% CI=7.3%-17.5%) when HRAs were implanted from 2007 onwards compared with 6.1% (95% CI=4.7%-7.9%) when implanted prior to 2007 (Figure 3.4.).

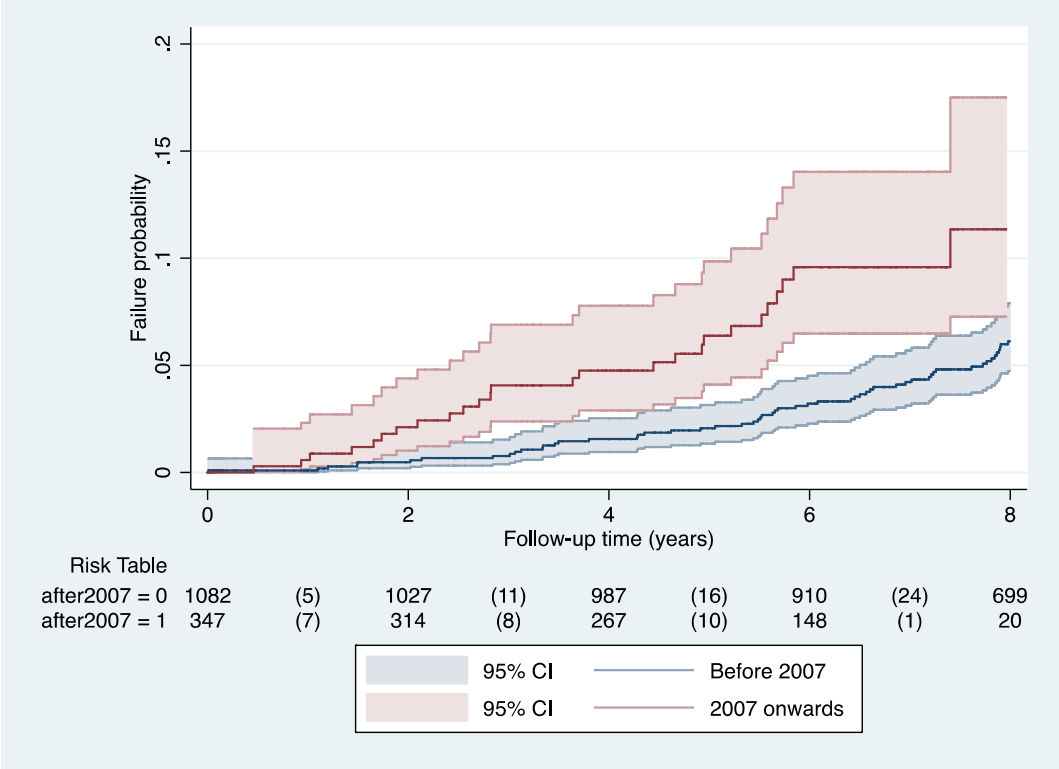


Figure 3.4 Kaplan-Meier cumulative pseudotumour revision rate for hip resurfacing arthroplasties at up to eight years by year of primary surgery.
Risk table indicates the number of hips at risk at two year intervals, with the corresponding number in brackets detailing the number of hips revised during the two year interval.

3.5. Study strengths and limitations

The study strengths include the large sample size and longer follow-up compared to previous reports [10, 76, 139, 144, 267], with this representing the first independent study providing information on pseudotumour revision risk into the second decade following HRA. Systematic methods were used to identify all revisions, including contacting other centres and assessing surgeons' registry profiles [2]. Revision indications were retrospectively confirmed using operative and histopathological findings. Therefore it was possible to identify pseudotumour revisions performed before this entity was formally established [14, 16]. Robust survival analysis methodology was used, including gender specific analysis. Furthermore the large cohort size ensured that the final regression models were not overfitted. Finally, the generalisability of the study findings is improved as a non-designing centre, which included many surgeons using a range of common implant designs.

Limitations include the potential for pseudotumour revision rates to be influenced by patient surveillance bias (NOC surgeons aware of problem from 2007 onwards, with MHRA patient recalls in 2010 and 2012) [23, 67] and a varying threshold for performing revision surgery over time (both considered further in Chapter 3.6.). Pseudotumour revision was used to define failure given this represents a problem serious enough to require intervention. However some non-revised asymptomatic patients may have had radiological pseudotumours [110, 223]. Although this would increase the reported pseudotumour failure rates it must be acknowledged that the natural history of radiological pseudotumours in both asymptomatic MoMHA and conventional THA patients remains unclear [212, 223]. The study findings could have

also been influenced by some patients not completing the postal questionnaire, and the possibility that some patients underwent revision elsewhere. Although survival analyses were robust there is potential for residual confounding. Finally, due to design differences the findings from this HRA cohort cannot be extrapolated to MoM THA patients.

3.6. Discussion

3.6.1. 15 year revision rates

This large prospective single centre cohort study observed a high 15 year revision rate for pseudotumour following HRA (14.0%), with pseudotumour representing the commonest indication for revision (61.7% of all revisions). Pseudotumour revision rates have increased significantly since an earlier report from this centre (4% at 8 years) [132]. Arthroplasty registries are ideally suited for assessing implant revision rates with their large numbers. Both the UK and Australian registries report ARMD to be the second commonest indication for HRA revision behind pain (UK) and loosening/lysis (Australia) [2, 11]. However registries have a number of significant limitations that can influence the reporting of pseudotumour revision rates. These include only adding this revision indication to registry databases in recent years, no histopathological data to confirm the surgeons intraoperative diagnosis of pseudotumour, and the potential to underreport future revisions [96, 137]. Hence there may be a considerable delay before registries can more accurately report pseudotumour revision rates.

Only the designer BHR surgeon has reported 15 year revision rates for his first 1,000 cases (4.2% for all causes), with pseudotumours accounting for 18% of all hips revised

[13]. The inferior outcomes reported here are likely to be related to patient, surgical, and implant factors. Broad patient selection criteria were used for HRA at the NOC (Chapter 3.3.3.). In hindsight this will have contributed to the high pseudotumour revision rates, especially in females. As HRA is technically demanding for the surgeon it is suspected the learning curves for the multiple surgeons contributing to this study will have influenced outcomes [64]. Finally although the BHR is well established as the best performing HRA design worldwide [2, 11], only 45% of the present cohort received this implant which also affected overall revision rates (Figure 3.3.).

The 15 year revision rate for pseudotumour in HRA patients reported here (14.0%) was higher than the 15 year all cause revision rate for young patients (under 55 years) following conventional THA reported in the Australian Joint Replacement Registry (9.9% for 30,956 THAs) [11]. As registries potentially underreport revisions [96, 137], it is important to also consider smaller cohorts reporting on young conventional THA patients. These reports suggest that patients under 50 receiving conventional THA can achieve 15 year all cause revision rates of 15% or less [270-272]. Therefore these all cause revision rates from single centres are similar to the revision rates for pseudotumour alone reported here, and are certainly lower than the all cause revision rate for the present HRA cohort (19.5%).

Pseudotumour revision rates continued to increase into the second decade after HRA. This was an unanticipated finding given one would suspect that revision rates would plateau once the poorly performing designs and the prostheses implanted with technical errors were revised early. The reasons for this observation are likely to be multifactorial. A major contributing factor is the increased awareness of surgeons to this complication

since 2007, after early reports of pseudotumours were published [16, 82]. At this centre all surgeons were not only aware of pseudotumours complicating MoM HRAs from 2007 onwards [16], but also began actively investigating asymptomatic well-functioning patients with blood metal ions and cross-sectional imaging [202]. Subsequent patient recalls in 2010 and 2012 triggered by MHRA alerts [23, 67] led to large numbers of patients being seen and investigated who would have otherwise been discharged. This major surveillance bias introduced at this centre, and at other centres [29, 73], since 2007 has clearly influenced pseudotumour revision rates (Figure 3.4.). Patients undergoing HRA from 2007 onwards had an almost 2.5 times increased risk of undergoing pseudotumour revision compared with patients having surgery before 2007 (a time when there was a lack of awareness and underreporting of pseudotumour). Although the ideal method for assessing the change in pseudotumour revision rates over time after an intervention period (introduction of regular surveillance) is by using interrupted time series analysis [273], very large numbers of annual implantations are required which this study was lacking. However performing such analysis and controlling the follow-up interval would only increase the effect size reported here following the commencement of regular surveillance. Therefore the current analysis actually underestimates the effect of surveillance bias on pseudotumour revision rates. This surveillance bias has also recently been illustrated in a systematic review, which observed that the highest prevalence of pseudotumour revision surgery occurred in centres with the most intensive HRA surveillance (universal blood metal ions and cross-sectional imaging) [136].

Another reason why pseudotumour revision rates have continued to increase over time relates to the lowering of the threshold for performing revision surgery in recent years.

Initially surgery was only performed in symptomatic patients with large lesions [16]. However in light of the poor short-term outcomes following revision surgery [12, 28], the indications at this and other centres have changed to now include mildly symptomatic patients with smaller pseudotumours [12, 29]. The HRA implants used at this centre also varied over time, which could influence pseudotumour revision rates (Chapter 3.3.3.). The BHR was almost exclusively used initially, with worse performing HRA implant designs used predominantly in more recent years. A final suggestion as to why pseudotumour revision rates continued to increase into the second decade relates to a potentially unrecognised mechanism of MoM bearing failure. Metal-on-polyethylene THAs may fail due to aseptic loosening both early (because of technical errors) and late (due to wear) [2, 11]. It is therefore possible that pseudotumours can develop both early and later for similar reasons, with the late pseudotumours representing the ultimate failure mechanism for previously well-functioning MoM HRAs.

3.6.2. Predicting pseudotumour revision

Female gender and small femoral head size were both independent predictors of pseudotumour revision in the whole cohort. Gender and femoral head size have complex interactions, with previous studies reporting mixed conclusions regarding their effects on HRA revision rates [10, 76, 132, 139, 144, 267]. This is why gender specific analyses were performed here. This analysis demonstrated that female gender was the more important risk factor for pseudotumour revision, with small femoral head size no longer a risk factor when the cohort had been stratified by gender. Therefore patients should be stratified by gender for clinical surveillance.

Females had twice the risk of undergoing pseudotumour revision compared with males. There are a number of reasons why females are at increased risk of pseudotumour revision. Females have larger native head-neck ratios, which can be considerably reduced following HRA, thereby increasing the risk of edge loading, high wear, and pseudotumour development [153]. Females also require smaller implants compared with males due to anatomical differences. These smaller implants are more sensitive to malposition, which can similarly cause edge loading and increased wear [20, 150]. Females have an increased risk of metal allergy [143], which may contribute to pseudotumours. Increased hip movement and flexibility as well as specific gait patterns in females [274, 275] may also cause impingement and/or edge loading with pseudotumour formation.

Predictors of pseudotumour revision surgery were found to be gender specific. In females the only risk factor was being young at surgery, whist in males the only predictors of pseudotumour revision were implantation with Conserve Plus and Recap implants. The highest risk of pseudotumour revision was observed in young females, which confirmed the results of an earlier short-term report from this centre [132]. Females under 50 years of age had a ten year pseudotumour revision rate of 24.5%. Reasons why younger females are at increased risk of pseudotumours compared with older females include the indication for surgery. Young females undergoing hip arthroplasty are more likely to have childhood hip disorders, such as dysplasia and other femoral and/or acetabular version abnormalities, and to have undergone previous hip surgery [13, 139, 267]. These abnormalities and previous operations increase the risk of component malposition, high wear, and pseudotumour formation [20, 150], with even the designing BHR surgeon reporting inferior outcomes in such patients [13].

Older females are more likely to undergo arthroplasty for primary hip osteoarthritis with component positioning in such cases suspected to be less challenging. Furthermore younger patients are more likely to be active. Given that females already have specific kinematics placing them at risk of impingement and/or edge loading [274, 275], such increased activity in younger females may contribute to the higher risk of pseudotumour.

The study findings suggest that non-BHR implants have higher failure rates in males, which supports observations from arthroplasty registries [2, 11]. Our data in males regarding good outcomes and low pseudotumour revision rates at extended follow-up in BHR patients confirm observations from designing surgeons [13, 68] and independent centers [20, 69, 74], which support the use of the BHR in highly selected patients. The present study findings do not support the recent MHRA recommendations of not implanting BHRs in males requiring small femoral components (<48 mm), but the findings do support not implanting BHRs in females [184]. Although the present study suggests surveillance of male HRA patients should be stratified by implant design this is not currently recommended. Apart from the recent BHR guidance already discussed [184], the MHRA only distinguishes HRA patients with recalled ASR implants from other designs [23]. This is concerning given the high pseudotumour revision rates with Conserve Plus and Recap implants in males observed in this chapter, with high all cause revision rates also confirmed for these particular designs in arthroplasty registries [2, 11].

3.6.3. Clinical implications of findings

In light of the study findings it is recommended that current worldwide follow-up protocols [22-24, 112] are adapted to ensure that the risk factors for pseudotumour revision surgery are appropriately weighted when stratifying HRA patients for surveillance. This stratification must be gender specific given that females are at increased risk of pseudotumour revision, regardless of implant design. Based on this study and the known performance of all HRA designs [2, 11] the risk stratification of HRA patients for surveillance should be as follows, starting with the highest risk of pseudotumour revision:

1. All patients with ASR HRAs
2. Young females (under 50 years at primary surgery)
3. All other females
4. Males with non-BHR implants (such as Conserve Plus and Recap)
5. Males with BHR implants

Patients with recalled implants, such as the ASR, should all undergo regular surveillance as is currently recommended [23]. The specific investigations needed and regularity of follow-up required for each of the other subgroups listed above will be considered in subsequent thesis chapters. However the low risk of pseudotumour revision in male patients with BHR implants at extended follow-up suggests that this is the only subgroup that may not require regular follow-up for the implant lifetime.

It is important for the orthopaedic community to learn lessons from the problems experienced with MoMHAs. After pre-clinical testing the introduction of new

technologies must undergo a rigorous and transparent process. This should include usage by small groups of experts, ideally in prospective randomised trials and independently controlled surveillance programmes. Surgeons using new technologies during these initial stages have a responsibility to report any problems to the manufacturers and authorities early [23, 24]. The results of this initial experience must be interpreted in combination with any registry data [2, 11] before widespread introduction. Adopting such an approach will allow innovation to continue, but also ensure patient safety.

Organisations such as Beyond Compliance and the Orthopaedic Data Evaluation Panel (ODEP) were set up to assist the safe introduction and usage of new implants in clinical practice [276, 277]. Beyond Compliance comprises a panel of independent experts who work with the implant manufacturers after they have received a CE mark for any new or modified implantable medical device. The expert panel subsequently works with the manufacturer to appraise the relative risks of these new devices, and decides on the rate at which the new implant should be introduced to the market. These decisions are made based on clinical data collected from patients who receive the devices, such as patient reported outcome measures and implant failure rates. ODEP was started in 2002 and also comprises of an independent panel of experts who regularly review evidence submitted from implant manufacturers about their products. ODEP provide a rating for each implant based on the length of patient follow-up, and quality and quantity of the data submitted. The minimum benchmark standard for implants used by ODEP is a revision rate equivalent to under 1% per year, with better performing implants distinguished more recently (starred rating if revision rate equivalent to under 0.5% per year) in line with the updated NICE guidelines [65].

Whilst the process described above may actually appear to delay new technologies reaching the market, which has partly been caused by the widespread problems with MoMHAs, and be considered by some to potentially stifle future innovation it is important to consider the overall objective. This is to provide patients with the best possible outcomes whilst maintaining their safety. In terms of hip arthroplasty it is important to remember that we already have one of the most successful surgical procedures available even with current implants and technologies on the market [1]. Therefore any future technologies must be designed to demonstrate a clear benefit to patients whilst not putting them at any additional risk.

3.7. Summary

- Revision rates for pseudotumour were high at up to 15 years following HRA (14.0%). These pseudotumour revision rates were also higher compared to previously reported 15 year all cause revision rates in young patients following primary conventional THA.

- Pseudotumour was the commonest indication for HRA revision surgery accounting for 61.7% of all revisions.

- Pseudotumour revision rates continued to increase into the second decade after HRA. The reasons for this are likely to be multifactorial and include: (1) surgeons becoming increasingly aware about pseudotumours, (2) more intense and regular patient surveillance, (3) lowering of the threshold for performing pseudotumour revision surgery in recent years, and (4) the potential for late

pseudotumour development which may be the ultimate failure mechanism for previously well-functioning HRAs.

- Risk factors for pseudotumour revision surgery were gender specific. Young females (under 50 years at primary surgery) had the highest risk of pseudotumour revision. In males the only predictor of pseudotumour revision was implantation with Conserve Plus and Recap designs.

- Current worldwide follow-up recommendations must ensure that the risk factors identified for pseudotumour revision surgery are appropriately weighted when stratifying HRA patients for surveillance.

- As many HRA patients worldwide will need regular clinical surveillance over future years it is important to establish the specific roles of the investigations used during follow-up, such as blood metal ions (Chapter 4.) and hip imaging (Chapter 5.). These commonly used investigative modalities are considered in the subsequent two chapters.

4. INVESTIGATION – BLOOD METAL IONS*

4.1. Introduction

Although blood metal ions are currently recommended in all follow-up protocols [22-24], there is presently no international consensus on the acceptable metal ion threshold(s) for concern in HRA patients. Thresholds for concern have been published by the MHRA (7 μg/l) [23, 67, 184]. However subsequent studies have reported variable optimal blood metal ion thresholds for identifying poorly functioning unilateral HRAs (4 μg/l to 7 μg/l), with different thresholds for concern also published in a recent US consensus statement [112] (Chapter 2.3.1.4.). Many of these proposed thresholds are used clinically, however most have not been validated.

Given that HRA revision rates vary considerably by implant design [2, 11, 66] it is possible that optimal blood metal ion thresholds for identifying patients with ARMD also differ between implant designs [172]. This could explain some of the variability observed between optimal blood metal ion thresholds [145, 162, 167, 168]. In an ASR HRA cohort, a threshold of 7 μg/l was considered optimal for identifying failing implants [168]. The ASR represents the worst performing HRA device worldwide [2, 11, 66]. Therefore it is not known whether the optimal blood metal ion threshold for

* Work presented in this chapter has been published in:

Matharu GS, Berryman F, Brash L, Pynsent PB, Treacy RB, Dunlop DJ. The effectiveness of blood metal ions in identifying patients with unilateral Birmingham Hip Resurfacing and Corail-Pinnacle metal-on-metal hip implants at risk of adverse reactions to metal debris. *J Bone Joint Surg Am* 2016; **98**(8): 617-626.

Matharu GS, Berryman F, Brash L, Pynsent PB, Dunlop DJ, Treacy RB. Can blood metal ion levels be used to identify patients with bilateral Birmingham Hip Resurfacings who are at risk of adverse reactions to metal debris? *Bone Joint J* 2016; **98-B**(11): 1455-1462.

ASR implants would apply to BHR implants, which represents the most commonly used HRA device worldwide [2, 11, 66].

Although the BHR also represents the best performing HRA device [2, 11, 66], studies have consistently reported inferior implant survivorship in females compared with males (Chapter 2.1.2.). Furthermore as most BHR patients are asymptomatic [13, 68-73] it is suspected that this large subgroup will be subjected to regular investigation for many years [22-24], especially following the recent MHRA surveillance modifications for many BHR patients [184]. Therefore it is important to establish clinically useful blood metal ion thresholds, which could informatively guide BHR patient follow-up. The lack of evidence-based guidance for blood metal ion thresholds in bilateral BHR patients is also concerning [145], given that this subgroup is considered to be at increased risk of ARMD by some authorities [24, 141] with up to 22% of BHR patients having bilateral implants (Chapter 4.3.1.). Any blood metal ion thresholds identified must subsequently be validated in external centres before being introduced clinically [278], despite this type of rigorous approach being absent in many clinical prediction studies [279].

4.2. Objectives

The objective of this chapter was to determine whether blood metal ion concentrations could effectively identify patients with unilateral and bilateral BHR implants who had ARMD.

4.3. Patients and methods

4.3.1. Study design

Two prospective cohort studies were performed. The first was a single centre study based on consecutive patients receiving primary BHR implants between July 1997 and September 2014 at the Royal Orthopaedic Hospital (ROH), Birmingham, UK, which is a specialist tertiary referral teaching hospital. This patient cohort was used to devise optimal blood metal ion thresholds for identifying patients who had ARMD and will subsequently be referred to as the derivation cohort. This study was registered with the hospital board, however ethical approval was not required as patients were assessed according to published MHRA guidance [23]. By February 2017, 3,749 primary BHRs had been implanted in 3,074 patients (22% of patients had bilateral BHRs).

The second study involved an external cohort of patients from two European centres receiving primary BHR implants between June 1999 and December 2012. This patient cohort was used to validate the optimal blood metal ion thresholds established in the derivation cohort and will subsequently be referred to as the validation cohort. Centre 1 was the NOC, Oxford, UK (Chapter 3.3.1.) and centre 2 was Coxa Hospital for Joint Replacement, Tampere, Finland. Coxa Hospital is the largest specialist tertiary referral joint replacement centre in the Nordic countries. Institutional review board approval was obtained to investigate patients at centre 2 (R11196: Pirkanmaa Hospital District ethics committee).

4.3.2. Data source

4.3.2.1. Derivation cohort

At the ROH clinical data have been collected and stored on a prospectively maintained electronic research database (MySQL, Version 5.6.14, Oracle Corporation, Redwood Shores, California, US) since the first HRA implantation in July 1997. Up until April 2003 surgeons would complete a paper data collection form at the time of operation, which included data on patient age, gender, date of surgery, hip laterality, hip diagnosis, implant design, and femoral head size. From April 2003 onwards surgeons would complete the NJR data capture form for primary hip arthroplasty (Appendix 8.). These data were subsequently entered onto the electronic database by trained research assistants.

Outcome data collected during patient follow-up were also entered onto the database, including postoperative OHSs (Appendix 6.), blood metal ion results, and details of any revision procedures performed (date, indication, findings). After laboratory analysis blood metal ion results (sample date, and blood cobalt and chromium concentrations) were electronically sent to a clinical scientist at the ROH through a secure portal. These results were then periodically (typically daily) uploaded to the hospital's electronic database, with unique patient identifiers (hospital number, name, date of birth, and gender) used to ensure accurate merger.

Revision procedures performed within the hospital were recorded by surgeons using either paper forms produced by the hospital before April 2003, or by using the NJR data capture form for revision hip arthroplasty (Appendix 9.). These procedures could be linked to the primary surgery when performed at the ROH using unique patient

identifiers. Revisions performed at other centres were also entered onto the database as applicable. Revisions were known to have occurred elsewhere when patients had informed the centre about such procedures (in person, by post, or by telephone), when surgeons had received correspondence from other clinicians about further procedures, and where the primary surgeon had identified revisions from their registry profile [280]. For completeness the lists of all patients awaiting surgery under hip surgeons at the ROH were searched to identify any pending procedures for BHR patients eligible for inclusion.

For this study the electronic radiology database (PACS) at the ROH was searched for all patients included in the derivation study to identify hips with any cross-sectional imaging evidence of ARMD (Chapter 4.3.6.).

4.3.2.2. Validation cohort

For centre 1, information about the data source has been detailed (Chapter 3.3.2.). At centre 2 clinical data has been collected and stored on a prospectively maintained electronic research database (Excel, Version 2011, Microsoft Corporation, Redmond, Washington, US) since the first HRA implantation in May 2001. Electronic data collection forms were completed by operating surgeons at the time of primary and revision surgery, which included data on patient age, gender, date of surgery, hip laterality, and hip diagnosis. In the case of revision surgery details of intra-operative findings were also recorded. For all procedures electronic bar code scanners were used intra-operatively to collect detailed information on implants used, including manufacturer and component sizes. Revisions were identified as being performed elsewhere by using methods described at the other centres (Chapters 3.3.5. and 4.3.2.1.),

with data from the Finnish Arthroplasty Registry confirming that very few primary hip arthroplasties implanted at centre 2 that are revised actually undergo revision at a different centre (3%). Other outcome data collected during patient follow-up were also entered onto the database, including postoperative HHSs [259] and blood metal ion results (sample date, and blood cobalt and chromium concentrations). All data were entered onto the electronic database by trained research nurses and research physiotherapists.

At both validation centres the electronic radiology databases were searched for all eligible patients to identify hips with any cross-sectional imaging evidence of ARMD (Chapter 4.3.6.). Furthermore lists of all patients awaiting surgery under hip surgeons at both centres were searched to identify any pending procedures for included patients. At centre 2, these particular tasks were performed by a second individual (orthopaedic surgical trainee).

4.3.3. Surgery performed

At validation centre 1, details about the BHR surgery performed have been provided (Chapter 3.3.3.). There were 20 surgeons performing BHR surgery at the ROH. One designing surgeon performed the majority of cases (52%). Much smaller volumes were performed by the other 19 non-designer surgeons, with the second highest volume BHR surgeon performing 10% of cases, and a further nine surgeons performing less than ten cases each. At validation centre 2, four non-designer surgeons implanted all BHRs.

The BHR is manufactured from an as-cast high-carbon cobalt chromium molybdenum alloy. The acetabular component is hemispherical and has a hydroxyapatite porous

coating allowing uncemented press-fit fixation, whilst the stemmed spherical femoral component requires cement for fixation. Acetabular component sizes range from 44 mm to 66 mm and femoral component sizes range from 38 mm to 58 mm, with both types of component available in 2 mm increments. The acetabular component should be 6 mm or 8 mm larger than the femoral component.

At the ROH and validation centre 2, the BHR implant was offered to young patients with hip arthritis. In general, the operation was offered to males (under 65 years) and females (under 60 years) with normal bone stock assessed on pelvic radiographs and an expectation that the patient would return to an active lifestyle, which may include sports. Indications for the procedure have subsequently evolved with increasing knowledge about the importance of correct component positioning in obtaining successful outcomes (specifically acetabular component inclination and combined acetabular and femoral component anteversion) [20, 150, 281]. This led surgeons to become more selective when offering BHR surgery to patients. Patient exclusions were made based on a combination of factors (including small femoral head sizes and hips with non-primary osteoarthritis) if it were thought that component positioning would be significantly compromised for BHR implantation. All procedures were performed in a clean-air laminar flow operating theatre with the patient in the lateral decubitus position using an extended posterior surgical approach as described by the designing surgeon [43].

4.3.4. Patient follow-up

Patient follow-up and investigation has been described for validation centre 1 (Chapter 3.3.4.). At the ROH, BHR patients routinely underwent a clinical assessment (history,

examination, anteroposterior pelvic radiographs, OHS questionnaire [257]) at the following intervals postoperatively: six weeks, one year, five years, ten years, and 15 years [170]. In addition some surgeons reviewed patients more regularly at their discretion, or used postal questionnaires to identify any further procedures and complete an OHS [56, 68, 282]. A similar follow-up protocol was routinely used at validation centre 2, with postal questionnaires used as an adjunct to clinical assessment at up to eight years from primary BHR [73, 283]. Patients identified as being symptomatic from these assessments at both the ROH and validation centre 2 underwent further investigations, including cross-sectional imaging.

Following the two MHRA alerts the routine patient follow-up regimen for BHR patients were adapted at both the ROH and validation centre 2 [23, 67]. In 2010, all BHR patients from the ROH were contacted by post and invited for an updated clinical assessment (detailed above). All symptomatic patients complaining of hip/groin pain, regardless of symptom severity and OHS, underwent blood metal ion sampling and cross-sectional imaging. All asymptomatic patients (typically with an OHS $\geq$34/48; representing good or excellent scores) [69] with ARMD risk factors also underwent blood metal ion sampling. These risk factors included small femoral components, malpositioned acetabular components (especially high acetabular component inclination and/or excessive anteversion), radiological features suggestive of failure (including osteolysis, radiolucent lines and narrowing of the femoral neck), hip instability, limping, and noises from the hip (including squeaking, clicking, grating, locking and clunking), either reported by the patient and/or elicited during examination [16, 20, 144, 150, 170]. Blood metal ion concentrations were also measured in patients with significant concerns as a consequence of media attention. Therefore blood sampling was

performed in a heterogeneous group of patients considered to be at risk of having ARMD. In line with MHRA guidance, cross-sectional imaging was performed in all asymptomatic patients with blood cobalt and/or chromium concentrations above 7 μg/l [23]. Selected asymptomatic patients with blood metal ion concentrations of 7μg/l or below also underwent cross-sectional imaging if there were clinical concerns about radiological BHR appearances, or if patients had a contralateral arthroplasty meeting the criteria for cross-sectional imaging.

In 2011, all BHR patients undergoing primary surgery at validation centre 2 were recalled for an updated clinical assessment (detailed above), regardless of symptoms. This assessment included universal blood metal ion sampling. All symptomatic patients, and asymptomatic patients with blood cobalt and/or chromium concentrations of 5 μg/l or above [167] underwent cross-sectional imaging. Selected asymptomatic patients with blood metal ions below 5 μg/l also underwent cross-sectional imaging at the surgeons' discretion, for example if there were concerns about radiological appearances, or if the contralateral arthroplasty required any form of cross-sectional imaging.

Subsequent patient follow-up at both the ROH and validation centre 2 was dependent on the findings from the clinical assessment and any investigations performed. Patients were either discharged, kept under continued regular surveillance with or without further investigations, or underwent revision surgery.

4.3.4.1. Cross-sectional imaging

At the ROH the cross-sectional imaging protocol recommends ultrasound for symptomatic patients and MARS-MRI for asymptomatic patients, with all ultrasounds performed by one of six fellowship trained musculoskeletal radiologists. As asymptomatic patients were suspected to need repeat imaging, serial MRIs were chosen to allow comparative analysis given serial ultrasounds are more difficult to compare, especially when different radiologists perform the examinations. Ultrasound was performed when MARS-MRI was contraindicated. This departmental protocol was based on advice from a group of expert musculoskeletal radiologists at the time of the patient recall [23, 67], with recent evidence confirming that both of these modalities have an important role in assessing HRAs (Chapter 2.3.3.3.).

Details regarding cross-sectional imaging performed at validation centre 1 have been described (Chapter 3.3.4.1.). At validation centre 2, MARS-MRI was used as the first line cross-sectional imaging modality. Ultrasound was used if MARS-MRI was contraindicated, for example in patients with pacemakers or those with claustrophobia. All ultrasounds were performed by one of three fellowship trained musculoskeletal radiologists. The detailed protocols for performing both MARS-MRI and ultrasound at each of the three centres have previously been described [204, 206, 209, 220, 284] and have also been summarised (Appendix 10.).

4.3.4.2. Blood metal ion sampling

Cobalt and chromium ion concentrations were measured in whole blood samples. The methods used at the ROH were similar to those described for validation centre 1 (Chapter 3.3.4.2.). Blood samples were analysed at a different MHRA approved

laboratory (City Hospital Trace Elements Laboratory, Birmingham, UK) using a different quadrupole inductively coupled plasma mass spectrometer (Agilent 7500cx, Agilent Technologies Inc., Santa Clara, California, US) (limit of detection 0.06 μg/l and reporting limit 0.6 μg/l).

At validation centre 2, whole venous blood was collected from the antecubital vein into a trace element tube containing sodium EDTA [73]. The samples were analysed at the Finnish Institute for Occupational Health with cobalt and chromium ion concentrations measured using a dynamic reaction cell quadrupole inductively coupled plasma mass spectrometer (Agilent 7500cx, Agilent Technologies Inc., Santa Clara, California, US) (limit of detection 0.06 μg/l and reporting limit 0.6 μg/l). This method has been validated in a previous blinded inter-laboratory study [285].

4.3.5. Study inclusion and exclusion criteria

The study selection criteria were identical for both the derivation and validation cohorts. All patients had to either have unilateral or bilateral BHR implants and have undergone at least one blood metal ion test following arthroplasty. Only unilateral BHR patients with a native contralateral hip at blood sampling were included, whilst bilateral patients were only included if both implants were BHRs at the time of sampling. Blood samples performed up until September 2014 were eligible for inclusion in the derivation cohort, and those performed up until October 2016 were eligible for inclusion in the validation cohort.

All blood samples had to be performed at least one year following primary BHR (or the most recent BHR in bilateral patients) to allow for the running-in phase of the bearing

surfaces [163]. The blood sample also had to be performed before any revision surgery. Patients undergoing blood sampling before revision surgery for non-ARMD indications (including infection, fracture, aseptic loosening, unexplained pain, dislocation and avascular necrosis) were excluded. This was to reduce the risk of confounding factors when devising specific blood metal ion thresholds for identifying patients with ARMD. In all non-ARMD cases the intraoperative revision findings, histopathology, and microbiology were assessed to confirm the absence of ARMD. After applying the study selection criteria the derivation cohort consisted of 309 patients with unilateral BHRs and 185 patients with bilateral BHRs, and the validation cohort consisted of 323 patients with unilateral BHRs and 93 patients with bilateral BHRs (Figure 4.1.).

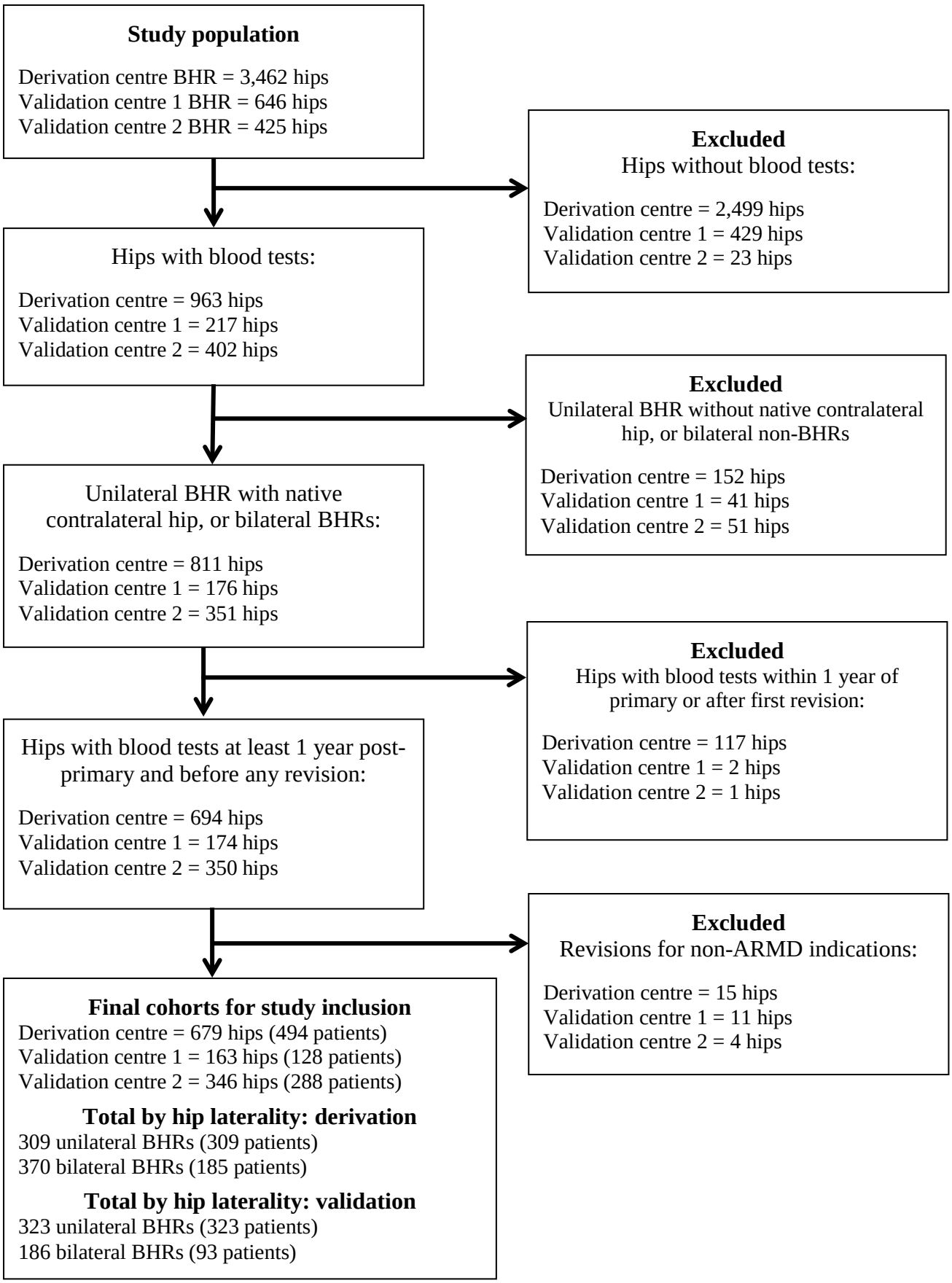


Figure 4.1 Inclusion and exclusion criteria for Birmingham Hip Resurfacing (BHR) patients in the derivation and validation cohorts. ARMD = adverse reactions to metal debris.

4.3.6. Outcomes

Eligible patients were assigned one of two possible outcomes (ARMD or non-ARMD) depending on their status at the study end date (September 2014 for the derivation cohort and October 2016 for the validation cohort). The ARMD group included patients revised or awaiting revision for ARMD (Chapter 2.2.1.), and those with cross-sectional imaging evidence of ARMD under surveillance because of clinician and/or patient preference. A diagnosis of ARMD was made if there was cross-sectional imaging and intra-operative evidence of a pseudotumour, or if there was no pseudotumour but abnormalities including peri-prosthetic effusions (on ultrasound this was defined as a depth of fluid exceeding 15 mm at the anterior joint line), significant metallosis, synovitis, tissue damage and/or necrosis (Chapter 3.3.5.) [25, 199, 202, 209]. The commonest reasons for patients with imaging evidence of ARMD not undergoing revision surgery were that they were either unfit, or were unwilling to undergo further procedures whilst asymptomatic. Revision surgery was recommended based on the outcome of the complete clinical assessment, which included hip radiographs and cross-sectional imaging. Blood metal ions alone were never used to make decisions regarding revision as previously recommended [168]. In hips undergoing revision surgery the diagnosis of ARMD was confirmed histopathologically (Chapter 3.3.5.). The non-ARMD group included all non-revised patients with primary BHRs in-situ, regardless of symptoms, with either no cross-sectional imaging evidence of ARMD or no such imaging performed.

Bilateral BHR patients with one or both hips in the ARMD group were considered to have ARMD, otherwise the patient was placed in the non-ARMD group. If patients had

multiple blood tests the most recent result was used in the non-ARMD group, with the result immediately prior to revision used in the ARMD group.

4.3.7. Statistical analysis

All statistical analysis was performed using R Version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria). Baseline differences between patients in the ARMD and non-ARMD groups were assessed using either the Chi-squared test or two-sided unpaired t-tests (age, femoral head size, and time to blood test). Unpaired t-tests were also used to compare the logarithms of the blood metal ion parameters between ARMD and non-ARMD groups. The logarithm was required to transform the asymmetric blood metal ion distributions to approximately normal distributions as recommended [239]. The four blood metal ion parameters of interest were cobalt, chromium, maximum cobalt or chromium (the higher value of the pair), and the cobalt-chromium ratio (cobalt divided by chromium, and non-dimensional).

Receiver operating characteristic (ROC) curve analysis is an established method of assessing the performance of a diagnostic test [286]. A ROC curve is drawn by plotting sensitivity (true positive rate) against 1 minus specificity (or 100 minus specificity if presented as a percentage; also called a false positive rate) for all possible test thresholds. A useful test produces a curve lying to the left of the 45° line (Figure 4.2.). The further the curve is towards the top left corner, the higher the area under the curve (AUC) and the better the discriminatory performance of the test (100% AUC represents a test with perfect discrimination; 50% AUC represents a non-discriminatory test). The analysis of ROC curves can also be used to define the optimal threshold to maximise the discriminatory ability for any given test.

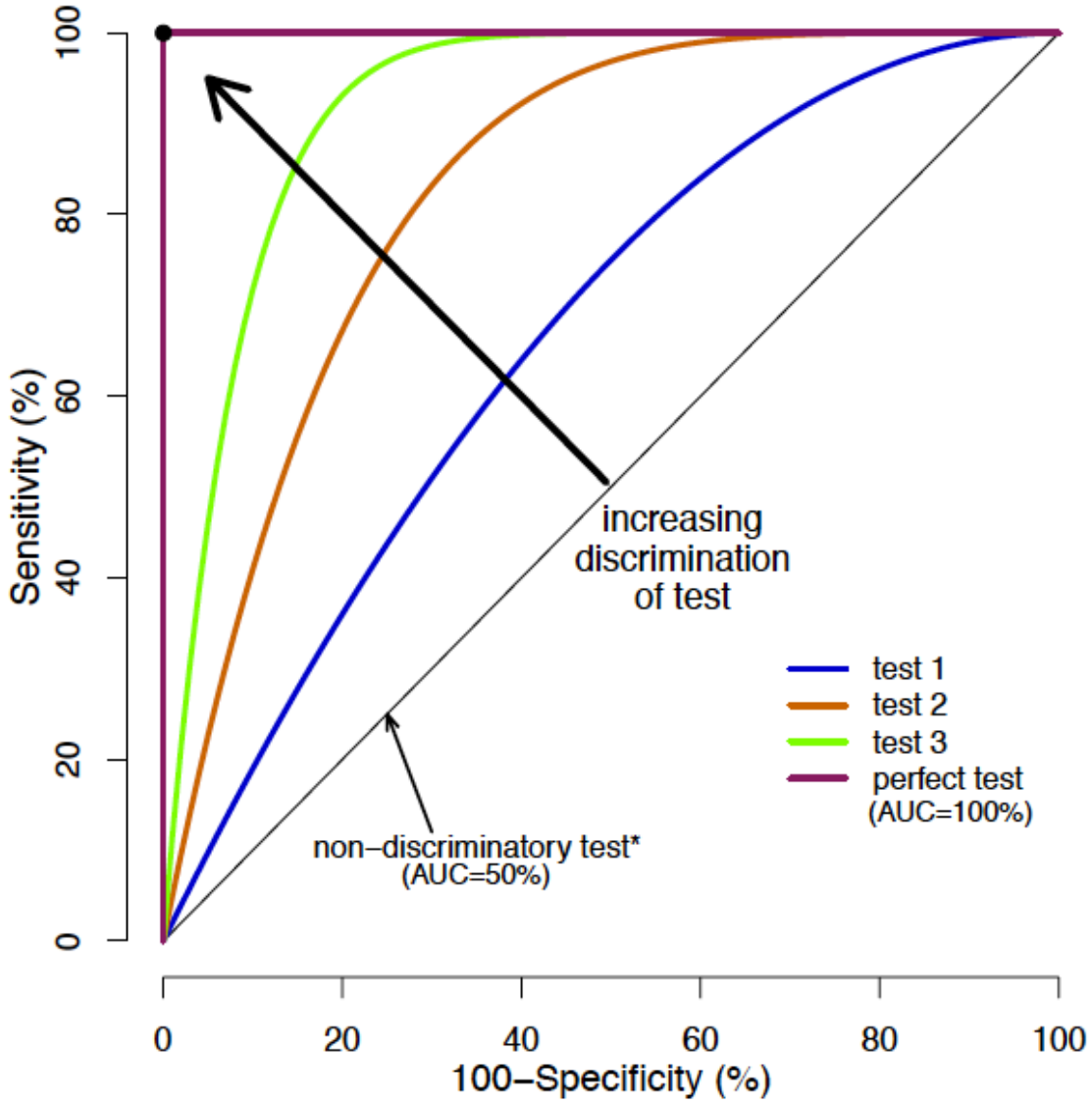


Figure 4.2 An explanation of receiver operating characteristic curve analysis with example lines for varying diagnostic test performance.
The black solid circle in the top left corner represents the perfect diagnostic threshold, with sensitivity and specificity both equal to 100%. The asterisk indicates a non-discriminatory test that is no better than a chance finding. Of the three tests shown, test 3 provides the best diagnostic performance, followed by test 2 and then test 1.

A ROC curve was used to determine the optimal blood metal ion thresholds for identifying unilateral and bilateral BHR patients with ARMD in the derivation cohort. This is a standard method for selecting the threshold maximising the sum of the specificity and sensitivity. The optimal blood metal ion threshold was defined as the threshold corresponding to the point on the curve nearest the top left corner. The

sensitivity (percentage of patients with ARMD above the threshold), specificity (percentage of patients without ARMD below the threshold), positive predictive value (PPV: percentage of patients above the threshold with ARMD), and negative predictive value (NPV: percentage of patients below the threshold without ARMD) were calculated for the optimal thresholds for each of the four blood metal ion parameters. Rates of misclassification were also calculated for all thresholds. The DeLong test was used to compare the AUCs between the different blood metal ion parameters in the derivation cohort [287]. The optimal thresholds identified for the blood metal ion parameters maximising the AUC in unilateral and bilateral BHRs in the derivation cohort were subsequently applied to the validation cohort. ROC curve analysis was similarly used to establish the discriminatory characteristics for the respective derivation cohort thresholds for identifying patients with ARMD in the validation cohort.

The McNemar test was used to compare the numbers of patients with ARMD who were not detected when using the optimal blood metal ion thresholds from this study with the thresholds proposed by the UK MHRA (7 μg/l) [23] and US surgeons (3 μg/l and 10 μg/l) [112]. This analysis was performed separately in both the derivation and validation cohorts.

4.4. Results

4.4.1. Patient demographics and outcomes

The derivation cohort consisted of 494 BHR patients (309 unilateral and 185 bilateral), and the validation cohort consisted of 416 BHR patients (323 unilateral and 93

bilateral). Patient demographics and outcomes have been summarised for both cohorts (Table 4.1.). The proportion of patients in the ARMD group in the derivation cohort was 11.3% (56/494), and 13.2% (55/416) in the validation cohort.

4.4.2. Blood metal ion concentrations

The four blood metal ion parameters (cobalt, chromium, maximum cobalt or chromium, and the cobalt-chromium ratio) have been summarised for both the derivation and validation cohorts (Table 4.2.). Apart from the cobalt-chromium ratio, all blood metal ion parameters were significantly higher in the ARMD group compared with the non-ARMD group for unilateral and bilateral BHRs in both the derivation cohort (all $p<0.0001$) and the validation cohort (all $p\leq0.0002$).

	Parameter	All patients	ARMD group	Non-ARMD group	p-value
Derivation cohort					
Unilateral BHR	Number of patients (%)	309 (100)	26 (8)	283 (92)	
	Gender (F/M) (%)	145 (46.9) / 164 (53.1)	22 (84.6) / 4 (15.4)	123 (43.5) / 160 (56.5)	**0.0001**
	Age at blood test (yr)	57.2 (19.6-78.7)	54.3 (32.8-74.0)	57.4 (19.6-78.7)	0.0641
	Femoral head size (mm)	48.5 (38-58)	46.0 (42-50)	48.7 (38-58)	**0.0006**
	Time from primary to blood test (yr)	7.8 (1.0-16.2)	8.9 (3.5-14.3)	7.6 (1.0-16.2)	0.0742
Bilateral BHR	Number of patients (%)	185 (100)	30 (16)	155 (84)	
	Gender (F/M) (%)	89 (48.1) / 96 (51.9)	19 (63.3) / 11 (36.7)	70 (45.2) / 85 (54.8)	0.104
	Age at blood test (yr)	60.4 (27.6-79.3)	60.3 (32.1-76.4)	60.4 (27.6-79.3)	0.949
	Femoral head size (mm)	48.5 (38-58)	46.1 (38-54)	49.0 (38-58)	**<0.0001**
	Time from latest primary to blood test (yr)	7.0 (1.0-16.8)	8.6 (2.4-15.2)	6.7 (1.0-16.8)	**0.014**
Validation cohort					
Unilateral BHR	Number of patients (%)	323 (100)	38 (12)	285 (88)	
	Gender (F/M) (%)	117 (36.2) / 206 (63.8)	27 (71.1) / 11 (28.9)	90 (31.6) / 195 (68.4)	**<0.0001**
	Age at blood test (yr)	63.0 (32.5-84.4)	59.8 (34.3-77.5)	63.4 (32.5-84.4)	0.0673
	Femoral head size (mm)	50.5 (38-58)	47.2 (38-58)	50.9 (38-58)	**0.0001**
	Time from primary to blood test (yr)	9.1 (1.9-15.2)	7.7 (1.9-14.9)	9.2 (2.4-15.2)	**0.0084**
Bilateral BHR	Number of patients (%)	93 (100)	17 (18)	76 (82)	
	Gender (F/M) (%)	33 (35.5) / 60 (64.5)	12 (70.6) / 5 (29.4)	21 (27.6) / 55 (72.4)	**0.0022**
	Age at blood test (yr)	63.5 (36.5-82.6)	59.3 (49.4-70.4)	64.4 (36.5-82.6)	**0.0069**
	Femoral head size (mm)	50.7 (42-58)	48.4 (42-54)	51.0 (42-58)	**0.0088**
	Time from latest primary to blood test (yr)	8.2 (1.2-14.3)	6.7 (3.2-10.9)	8.6 (1.2-14.3)	**0.0181**

Table 4.1 Patient demographics and outcomes for the derivation and validation cohorts.
ARMD = adverse reactions to metal debris; BHR = Birmingham Hip Resurfacing; F = female; M = male. Unless stated otherwise values presented are means, with ranges in brackets. P-values <0.05 in bold text (for comparisons between the ARMD and non-ARMD groups).

	Parameter	All patients	ARMD group	Non-ARMD group	p-value
Derivation cohort					
Unilateral BHR	Number of patients (%)	309 (100)	26 (8)	283 (92)	
	Co (µg/l)	1.06 (0.71-2.01)	11.95 (2.71-61.12)	1.00 (0.71-1.56)	**<0.0001**
	Cr (µg/l)	1.56 (1.09-2.50)	9.91 (3.12-35.36)	1.46 (1.04-2.16)	**<0.0001**
	Max Co or Cr (µg/l)	1.61 (1.12-2.76)	13.28 (3.12-61.12)	1.51 (1.09-2.29)	**<0.0001**
	Co-Cr ratio	0.71 (0.52-1.00)	1.16 (0.78-1.94)	0.69 (0.51-0.95)	**<0.0001**
Bilateral BHR	Number of patients (%)	185 (100)	30 (16)	155 (84)	
	Co (µg/l)	2.24 (1.47-4.01)	39.41 (6.83-86.18)	1.95 (1.39-2.98)	**<0.0001**
	Cr (µg/l)	2.70 (1.98-4.73)	22.41 (7.36-47.74)	2.39 (1.87-3.95)	**<0.0001**
	Max Co or Cr (µg/l)	2.81 (2.08-5.43)	39.41 (9.28-86.18)	2.60 (1.92-4.11)	**<0.0001**
	Co-Cr ratio	0.83 (0.67-1.12)	1.59 (0.94-2.00)	0.78 (0.66-1.01)	**<0.0001**
Validation cohort					
Unilateral BHR	Number of patients (%)	323 (100)	38 (12)	285 (88)	
	Co (µg/l)	1.40 (1.00-1.95)	9.75 (2.35-31.08)	1.30 (1.00-1.70)	**<0.0001**
	Cr (µg/l)	1.29 (0.90-2.40)	8.92 (3.50-33.41)	1.18 (0.80-1.80)	**<0.0001**
	Max Co or Cr (µg/l)	1.50 (1.10-2.55)	10.51 (4.16-44.92)	1.40 (1.00-2.00)	**<0.0001**
	Co-Cr ratio	1.06 (0.78-1.38)	0.72 (0.46-1.17)	1.08 (0.83-1.38)	**0.0046**
Bilateral BHR	Number of patients (%)	93 (100)	17 (18)	76 (82)	
	Co (µg/l)	2.10 (1.50-3.30)	6.30 (4.80-22.50)	1.80 (1.30-2.52)	**0.0002**
	Cr (µg/l)	2.30 (1.40-5.10)	8.50 (5.20-24.10)	2.04 (1.30-2.90)	**0.0002**
	Max Co or Cr (µg/l)	2.50 (1.70-5.10)	8.50 (5.36-24.10)	2.30 (1.60-2.92)	**0.0001**
	Co-Cr ratio	0.94 (0.68-1.20)	0.89 (0.50-1.26)	0.94 (0.73-1.18)	0.3890

Table 4.2 Median (interquartile range) blood metal ion parameters for the derivation and validation cohorts.
ARMD = adverse reactions to metal debris; BHR = Birmingham Hip Resurfacing; Co = cobalt; Cr = chromium. P-values <0.05 in bold text (for comparisons between the ARMD and non-ARMD groups).

4.4.3. Derivation cohort: formulating unilateral BHR thresholds

Optimal blood metal ion thresholds for discriminating between unilateral BHR patients with and without ARMD were dependent on the specific metal ion parameter used (Table 4.3.). Compared with the other three parameters, cobalt produced the maximum AUC of 90.5% (95% CI=82.8%-98.1%). The cobalt AUC was significantly greater than the cobalt-chromium ratio AUC (p=0.0005), but not significantly greater than that for chromium (p=0.8483) or the maximum cobalt or chromium (p=0.4213) (Figure 4.3.). The cobalt threshold for identifying unilateral BHR patients with ARMD providing the optimal diagnostic test characteristics was 2.15 µg/l. This threshold provided 88.5% sensitivity, 84.5% specificity, 34.3% PPV, and 98.8% NPV (Table 4.3.).

	Co	Cr	Max Co or Cr	Co-Cr ratio
Unilateral BHR (n=309)				
AUC (%)	90.5 (82.8-98.1)	90.3 (82.5-98.1)	89.9 (82.0-97.9)	76.3 (65.8-86.8)
Optimal threshold	2.15 µg/l	2.47 µg/l	2.45 µg/l	1.12
Sensitivity (%)	88.5 (76.9-100)	88.5 (76.9-100)	92.3 (80.8-100)	57.7 (38.5-76.9)
Specificity (%)	84.5 (80.2-88.3)	79.2 (74.2-83.8)	76.7 (71.4-82.0)	87.3 (83.4-90.8)
PPV (%)	34.3 (24.1-46.3)	28.0 (19.5-38.6)	26.7 (18.6-36.6)	29.4 (18.7-43.0)
NPV (%)	98.8 (96.4-99.6)	98.7 (96.2-99.5)	99.1 (96.7-99.7)	95.7 (92.5-97.6)
Misclassification of patients (%)	15.2	20.1	22.0	15.2
Bilateral BHR (n=185)				
AUC (%)	90.0 (83.5-96.4)	90.6 (83.0-98.1)	91.0 (84.5-97.4)	78.9 (68.4-89.5)
Optimal threshold	5.72 µg/l	5.51 µg/l	5.52 µg/l	1.15
Sensitivity (%)	76.7 (60.0-90.0)	83.3 (70.0-96.7)	83.3 (70.0-96.7)	73.3 (56.7-86.7)
Specificity (%)	92.9 (88.4-96.8)	91.0 (86.4-95.5)	88.4 (83.2-92.9)	85.2 (79.4-90.3)
PPV (%)	67.6 (51.9-83.4)	64.1 (49.0-79.2)	58.1 (43.4-72.9)	48.9 (34.3-63.5)
NPV (%)	95.4 (92.0-98.7)	96.6 (93.6-99.5)	96.5 (93.4-99.5)	94.3 (90.4-98.1)
Misclassification of patients (%)	9.7	10.3	12.4	16.8

Table 4.3 Summary of the receiver operating characteristic analysis for the derivation cohort. AUC = area under the curve; BHR = Birmingham Hip Resurfacing; Co = cobalt; Cr = chromium; NPV = negative predictive value; PPV = positive predictive value. 95% confidence intervals are in brackets.

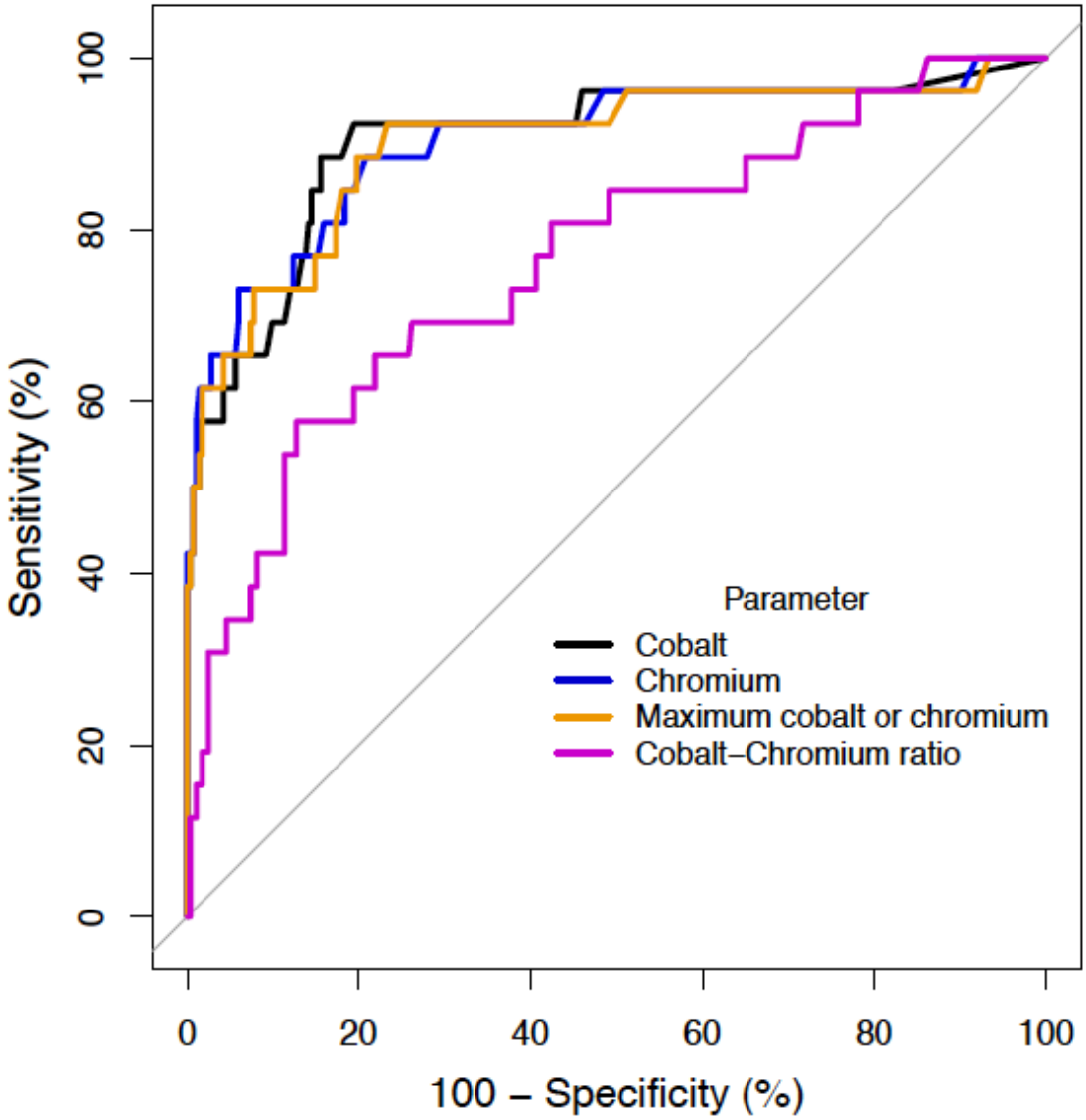


Figure 4.3 Receiver operating characteristic curve illustrating the ability of four blood metal ion parameters to distinguish between unilateral Birmingham Hip Resurfacing patients with and without adverse reactions to metal debris in the derivation cohort.

4.4.4. Derivation cohort: formulating bilateral BHR thresholds

Optimal blood metal ion thresholds for discriminating between bilateral BHR patients with and without ARMD were dependent on the specific metal ion parameter used (Table 4.3.). Compared with the other three parameters, maximum cobalt or chromium produced the highest AUC of 91.0% (95% CI=84.5%-97.4%). The maximum cobalt or chromium AUC was significantly greater than the cobalt-chromium ratio AUC

(p=0.019), but not significantly greater than that for cobalt (p=0.574) or chromium (p=0.721) (Figure 4.4.). The maximum cobalt or chromium threshold for identifying bilateral BHR patients with ARMD providing the optimal diagnostic test characteristics was 5.52 µg/l. This threshold provided 83.3% sensitivity, 88.4% specificity, 58.1% PPV, and 96.5% NPV (Table 4.3.).

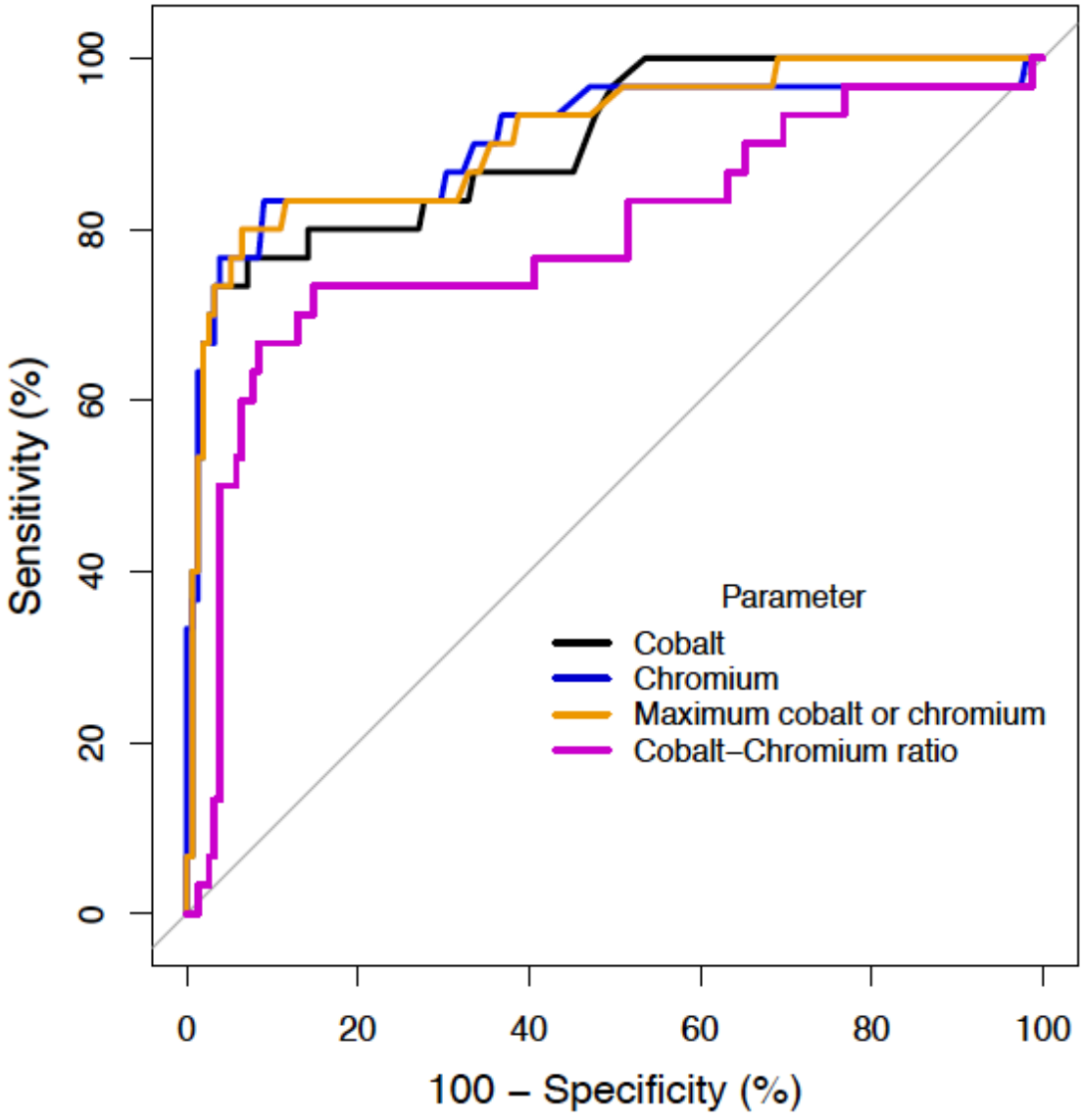


Figure 4.4 Receiver operating characteristic curve illustrating the ability of four blood metal ion parameters to distinguish between bilateral Birmingham Hip Resurfacing patients with and without adverse reactions to metal debris in the derivation cohort.

4.4.4.1. Sensitivity analysis: bilateral BHRs with cross-sectional imaging

Optimal blood metal ion thresholds were assessed in a subgroup of bilateral BHR patients who had all undergone cross-sectional imaging. This subgroup included 111 patients (60% of the bilateral BHR derivation cohort), of whom 30 (27%) were in the ARMD group and 81 (73%) were in the non-ARMD group. Similar results were obtained in this subgroup compared with all bilateral BHR patients. In bilateral BHR patients with cross-sectional imaging, the maximum cobalt or chromium threshold for identifying patients with ARMD providing the optimal diagnostic test characteristics was also 5.52 μg/l. The AUC was 88.7% (95% CI=81.2%-96.2%), and this threshold provided 83.3% sensitivity, 85.2% specificity, 67.6% PPV, and 93.2% NPV.

4.4.5. Derivation cohort: BHR thresholds vs. regulatory authority thresholds

Fixed blood metal ion thresholds for concern proposed by the UK MHRA (7 μg/l) [23] and US surgeons (3 μg/l and 10 μg/l) [112] were applied to the cohort (Table 4.4.). In general compared with fixed regulatory authority thresholds, the newly devised BHR thresholds provided the best balance between sensitivity and specificity, higher NPVs, and lower PPVs. Applying the BHR thresholds to the derivation cohort resulted in eight patients with ARMD being missed (1.6% of cohort). More patients with ARMD were missed when using fixed regulatory thresholds in the derivation cohort: 3 μg/l=10 patients missed (2.0%); 7 μg/l=17 patients missed (3.4%); 10 μg/l=21 patients missed (4.3%). Significantly more patients with ARMD were missed when applying the 7 μg/l (p=0.008) and 10 μg/l (p=0.001) fixed thresholds. The difference between the 3 μg/l fixed threshold and the BHR thresholds was not statistically significant (p=0.480).

	Blood metal ion threshold (µg/l)	AUC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Misclassification (%)	Patients with ARMD not found
Unilateral BHR (n= 309)	Co = 10	90.5 (82.8-98.1)	50.0 (30.8-69.2)	98.6 (97.2-99.7)	76.5 (56.3-96.6)	95.5 (93.2-97.9)	5.5	13
	Co = 7		57.7 (38.5-76.9)	98.2 (96.5-99.7)	75.0 (56.0-94.0)	96.2 (94.0-98.4)	5.2	11
	Co = 3		69.2 (50.0-88.5)	88.7 (84.8-92.2)	36.0 (22.7-49.3)	96.9 (94.8-99.0)	12.9	8
	Co = 2.15		88.5 (76.9-100)	84.5 (80.2-88.3)	34.3 (24.1-46.3)	98.8 (96.4-99.6)	15.2	3
Bilateral BHR (n=185)	Max Co or Cr = 10	91.0 (84.5-97.4)	73.3 (57.5-89.20	96.1 (93.1-99.2)	78.6 (63.4-93.8)	94.9 (91.5-98.3)	7.6	8
	Max Co or Cr = 7		80.0 (65.7-94.3)	93.6 (89.7-97.4)	70.6 (55.3-85.9)	96.0 (92.9-99.1)	8.6	6
	Max Co or Cr = 3		93.3 (76.5-98.8)	60.6 (52.4-68.3)	31.5 (22.3-42.3)	97.9 (92.0-99.6)	34.1	2
	Max Co or Cr = 5.52		83.3 (70.0-96.7)	88.4 (83.2-92.9)	58.1 (43.4-72.9)	96.5 (93.4-99.5)	12.4	5

Table 4.4 Summary of the receiver operating characteristic analysis for optimal Birmingham Hip Resurfacing blood metal ion thresholds compared with fixed authority thresholds in the derivation cohort.
ARMD = adverse reactions to metal debris; AUC = area under the curve; BHR = Birmingham Hip Resurfacing; Co = cobalt; Cr = chromium; NPV = negative predictive value; PPV = positive predictive value. 95% confidence intervals are in brackets.

4.4.6. Validation cohort: applying unilateral and bilateral BHR thresholds

In the unilateral BHR validation cohort the AUC for the cobalt parameter was 89.4% (95% CI=82.8%-96.0%) (Table 4.5.). When the unilateral BHR threshold from the derivation cohort was applied (cobalt=2.15 μg/l), the sensitivity, specificity, PPV, and NPV for distinguishing between patients with and without ARMD in the validation cohort were 78.9%, 86.7%, 44.1%, and 96.9% respectively.

In the bilateral BHR validation cohort the AUC for the maximum cobalt or chromium parameter was 89.2% (95% CI=81.3%-97.1%) (Table 4.5.). When the bilateral BHR threshold from the derivation cohort was applied (maximum cobalt or chromium=5.52 μg/l), the sensitivity, specificity, PPV, and NPV for distinguishing between patients with and without ARMD in the validation cohort were 70.6%, 86.8%, 54.5%, and 93.0% respectively.

	Unilateral BHR (n=323)	**Bilateral BHR (n=93)**
Optimal threshold	Co = 2.15 μg/l	Max Co or Cr = 5.52 μg/l
AUC (%)	89.4 (82.8-96.0)	89.2 (81.3-97.1)
Sensitivity (%)	78.9 (66.0-91.9)	70.6 (48.9-92.2)
Specificity (%)	86.7 (82.7-90.6)	86.8 (79.2-94.4)
PPV (%)	44.1 (32.3-55.9)	54.5 (33.7-75.4)
NPV (%)	96.9 (94.7-99.0)	93.0 (87.0-98.9)
Misclassification of patients (%)	14.2	16.1

Table 4.5 Summary of the receiver operating characteristic analysis for the validation cohort. AUC = area under the curve; BHR = Birmingham Hip Resurfacing; Co = cobalt; Cr = chromium; NPV = negative predictive value; PPV = positive predictive value. 95% confidence intervals are in brackets.

4.4.6.1. Sensitivity analysis: validation cohort with cross-sectional imaging

Optimal blood metal ion thresholds identified from the derivation cohort were applied to a subgroup of BHR patients in the validation cohort who had all undergone cross-sectional imaging. This subgroup included 156 unilateral BHR patients (48% of the unilateral validation cohort; 38 ARMD and 118 non-ARMD), and 56 bilateral BHR patients (60% of the bilateral validation cohort; 17 ARMD and 39 non-ARMD). Similar results were obtained in these subgroup analyses compared with analyses performed in the whole validation cohort.

In unilateral BHR patients from the validation cohort with cross-sectional imaging the AUC for the cobalt parameter was 86.5% (95% CI=79.0%-94.1%), which was not significantly different to the AUC in all unilateral BHR patients (p=0.579). Applying the optimal cobalt threshold of 2.15 μg/l provided 78.9% sensitivity, 82.2% specificity, 58.8% PPV, and 92.4% NPV. In bilateral BHR patients from the validation cohort with cross-sectional imaging the AUC for the maximum cobalt or chromium parameter was 82.4% (95% CI=70.7%-94.0%), which was not significantly different to the AUC in all bilateral BHR patients (p=0.343). Applying the optimal maximum cobalt or chromium threshold of 5.52 μg/l provided 70.6% sensitivity, 76.9% specificity, 54.5% PPV, and 93.0% NPV.

4.4.7. Validation cohort: BHR thresholds vs. regulatory authority thresholds

Applying the BHR thresholds to the validation cohort resulted in 13 patients with ARMD being missed (3.1% of the cohort) (Appendix 11.). More patients with ARMD were missed when using fixed regulatory thresholds in the validation cohort: 3 μg/l=14 patients missed (3.4%); 7 μg/l=22 patients missed (5.3%); 10 μg/l=30 patients missed

(7.2%). Significantly more patients with ARMD were missed when applying the 7 μg/l (p=0.0077) and 10 μg/l (p=0.0001) fixed thresholds. The difference between the 3 μg/l fixed threshold and the BHR thresholds was not statistically significant (p=1.0).

4.5. Study strengths and limitations

Study strengths include the large sample size of both the derivation and validation cohorts, and the longer follow-up time compared with previous studies assessing blood metal ion concentrations [145, 167, 168]. In particular the derivation cohort of bilateral BHRs represents the largest report of blood metal ion concentrations in bilateral HRA patients [145]. Reporting on only the BHR implant, which represents the most commonly used HRA device worldwide [2, 11, 66], is likely to reduce the risk of any potential confounding related to implant design as well as increase the generalisability of the findings. Contrary to previous studies [145, 167, 168] a robust definition for ARMD failure was used (Chapter 4.3.6.) which is likely to improve the diagnostic test characteristics obtained for the various ion thresholds. Although the cobalt-chromium ratio has been briefly considered in previous studies [168, 181], this represents the first study formally assessing this parameter in HRA patients in a robust manner. This study was strengthened considerably by including a validation cohort. The validation cohort was large, with a similar size compared with the derivation cohort which is contrary to most validation studies [288]. Furthermore the cohort was truly external as it involved patients from two geographically different centres, which is considered superior methodologically compared with other types of validation studies such as internal or temporal validation [288].

The main study limitation relates to patient selection bias during follow-up, which was apparent at two stages. The first bias relates to the patient subgroup undergoing blood sampling. In the derivation cohort and validation centre 1 not all asymptomatic patients underwent blood metal ion sampling, which may affect the generalisability of the optimal thresholds. However validation centre 2 did perform universal blood sampling, which provides some support that the devised thresholds can be used in centres undertaking either universal or targeted blood sampling. The second selection bias present at all centres was the use of targeted cross-sectional imaging in asymptomatic patients with raised blood metal ions. Although this reflects the approach used in the clinical setting where follow-up resources must be rationalised [22-24, 102, 145], some asymptomatic patients may have had silent ARMD though would have been incorrectly assigned to the non-ARMD group. The use of targeted imaging in asymptomatic patients could therefore influence the discriminatory characteristics reported in both the derivation and validation cohorts. However a subgroup analysis performed on all patients undergoing cross-sectional imaging in both the derivation and validation cohorts supported the results obtained in the whole cohort (Chapters 4.4.4.1. and 4.4.6.1.).

Another limitation is that only one blood metal ion result per patient was used in this study. Therefore recommendations cannot be made about the intervals for repeat blood testing. Patient and laboratory factors known to influence blood metal ion concentrations may have affected the reported findings (Chapter 2.3.1.1.). The bilateral BHR group in the validation cohort (n=93) was smaller compared with the derivation cohort (n=185), which may have affected the discriminatory characteristics reported in the bilateral validation cohort. Finally, the devised optimal thresholds only apply to

patients with unilateral or bilateral BHRs and cannot be assumed to be applicable to other MoMHA implant designs.

4.6. Discussion

4.6.1. Deriving blood metal ion thresholds

Regulatory authorities and numerous clinical studies have proposed blood metal ion thresholds for concern in HRA patients [23, 112, 145, 162, 167, 168]. However there is considerable variability between the thresholds reported (3 μg/l to 10 μg/l), with most ion thresholds not undergoing any form of validation prior to clinical implementation.

The present large derivation cohort demonstrated blood metal ion concentrations were effective for discriminating between unilateral and bilateral BHR patients with and without ARMD. The cobalt, chromium, and maximum cobalt or chromium all produced high and similar AUCs (90% or above), which were significantly greater compared with the AUC for the cobalt-chromium ratio. This suggests that the cobalt-chromium ratio has no additional role in the management of BHR patients compared with the other three parameters. Of all the parameters investigated, cobalt maximised the AUC in unilateral BHR patients and therefore was the best ion to test for in this group, with maximum cobalt or chromium representing the best parameter in bilateral BHR patients. This variation in the optimal ion parameter by implant laterality may reflect differences in bearing diameter and component positioning, which become increasingly complex in bilateral BHR patients but will affect the overall blood metal ion concentrations in different ways [155, 170]. Although cobalt, chromium, and maximum cobalt or chromium all produced similar AUCs, and therefore any could theoretically be used in

unilateral and bilateral BHR patients, it is important to acknowledge that the optimal blood metal ion thresholds for discriminating between BHR patients with and without ARMD were dependent on the specific parameter used. This is clinically relevant as if surgeons wish to use, for example, cobalt alone in preference to the maximum cobalt or chromium parameter in bilateral BHR patients, it is important the respective optimal threshold is applied as these do differ (cobalt 5.72 μg/l vs. maximum cobalt or chromium 5.52 μg/l).

The optimal blood metal ion thresholds were observed to be most effective for identifying unilateral and bilateral BHR patients who were at low risk of ARMD (almost perfect NPVs), rather than identifying those patients with ARMD (poor PPVs). This is clinically important as it is favourable to exclude the many patients without ARMD, thereby allowing resources to be focussed on the smaller subgroup of patients who may have ARMD. The newly devised BHR thresholds in unilateral and bilateral patients were both lower than the 7 μg/l that was considered optimal in patients with ASR HRAs [168]. The ASR represents the worst performing HRA implant worldwide whilst the BHR represents the best performing implant [2, 11, 66]. As blood metal ion concentrations reflect bearing wear [155] it is reasonable to assume that the higher ion thresholds observed in ASR devices relate to its poorer design and clinical performance compared with the BHR [148, 168]. This observation could explain some of the variability reported between optimal blood metal ion thresholds in previous studies, with most cohorts including a variety of different HRA designs rather than one single implant design [145, 162, 167].

The new unilateral and bilateral BHR thresholds provided more balanced sensitivity and specificity compared with thresholds identified in previous studies, which demonstrated good specificity but poor sensitivity [145, 167, 168]. In addition to the variety of implant designs included in previous studies, it is suspected that the definition used for failure was a major factor for the more favourable diagnostic test characteristics observed with the BHR thresholds. Previous studies have included all symptomatic patients, and revisions for non-ARMD indications as failures [145, 167, 168]. This represents a heterogeneous group of patients with a problematic HRA regardless of cause. The present study definition for identifying MoM bearing-related complications was more robust, and included patients who had not undergone a revision but had ARMD on imaging as failures. Therefore the devised BHR thresholds are specific for identifying patients with ARMD.

The optimal blood metal ion thresholds were higher in bilateral BHR patients compared with unilateral BHR patients. This supports the only other study assessing blood metal ion concentrations in both unilateral and bilateral HRA patients [145], and is unsurprising given the increased propensity for metal ion generation from two implants compared with one. However the difference between unilateral and bilateral BHR thresholds is of clinical relevance. Doubling the optimal thresholds for unilateral BHRs could be considered a crude estimate for bilateral thresholds. However doubling the optimal unilateral BHR thresholds (4.3 μg/l) gives a considerably different value compared with the optimal bilateral BHR threshold (5.52 μg/l). This highlights the importance of using the respective threshold when managing patients with unilateral and bilateral BHRs.

Application of fixed blood metal ion thresholds recommended by the UK MHRA and US surgeons missed more patients with ARMD compared with the new BHR thresholds. The 7 μg/l (UK MHRA) [23] and 10 μg/l (US high risk threshold) thresholds [112] missed significantly more BHR patients with ARMD compared with the newly devised thresholds. When using 3 μg/l (US low risk threshold) [112] although the difference was not statistically significant it was considered to be clinically significant given the potentially destructive nature of ARMD and poor outcomes initially reported following ARMD revision surgery [28]. It is important not to miss patients with this potentially destructive condition. Therefore minimising the false negative rate is clinically beneficial and supports the use of the new BHR thresholds over current fixed authority thresholds.

4.6.2. Validating blood metal ion thresholds

In reality very few validation studies are performed to confirm or refute the findings from original studies [289]. When validation studies are performed they can be limited by including much smaller cohorts compared with the derivation study, and by not involving a truly external cohort [288]. The present large validation study can therefore be considered methodologically strong (Chapter 4.5.). Very similar findings were observed when the optimal unilateral and bilateral BHR ion thresholds identified in the derivation cohort were applied to the validation cohort. These thresholds were again effective for identifying patients at low risk of ARMD, and missed fewer patients with ARMD compared to current fixed thresholds [23, 112].

In the validation cohort the unilateral and bilateral BHR thresholds were good at distinguishing between patients with and without ARMD (high AUC), with good and

balanced sensitivity and specificity. Furthermore these thresholds were effective for identifying patients without ARMD (high NPV). Although the BHR thresholds performed marginally better overall in the derivation cohort compared with the validation cohort, this is to be expected. Statistical models invariably perform better in the devised cohort compared with any external cohort because they will be intimately related to the patient, surgical, and unmeasured factors from the derivation study, which will almost never be reproducible externally [290, 291]. However the degree to which the findings from the validation cohort parallel those from the derivation cohort is reassuring.

In the validation cohort more patients with ARMD were missed when using fixed regulatory thresholds proposed by the UK MHRA and US surgeons compared with the BHR thresholds. The differences were statistically significant when applying the 7 μg/l and 10 μg/l thresholds, but did not reach significance when applying the 3 μg/l threshold. These observations confirm the findings in the derivation cohort and can also be considered clinically significant differences (Chapter 4.6.1.). Furthermore applying a low threshold such as 3 μg/l in all bilateral HRA patients may be considered inappropriate given such patients are considered at increased risk of ARMD [24, 141].

4.6.3. Clinical implications of findings

The management of HRA patients is complex (Chapter 2.) [292]. There are recognised risk factors for developing ARMD, which include being female, young at primary surgery, and having non-BHR implants (Chapter 3.). These factors are useful for risk stratifying HRA patients for surveillance [112]. The present study has demonstrated that blood metal ions have an important role when investigating BHR patients; namely

at identifying patients at low risk of ARMD. However blood metal ions cannot be relied upon as the only investigation when assessing HRA patients, as based on the data presented they cannot conclusively identify patients with ARMD or stratify according to the severity of ARMD [278]. Therefore blood metal ions play an important part in the complete clinical assessment of HRA patients.

Currently there is no international consensus on the acceptable blood metal ion threshold(s) for concern in HRA patients. Although fixed ion thresholds proposed by worldwide authorities are widely used clinically [23, 112], it is worth considering how they were formulated. In 2010, the MHRA first proposed 7 μg/l as the blood metal ion threshold for clinical concern, which was reasonable at this time given that ARMD was a new entity [67]. However the MHRA still recommend 7 μg/l [184] despite evidence to the contrary, with lower thresholds being preferable for managing patients according to previous reports [145, 162, 167, 185] and the present study. The more recent consensus statement from US surgeons attempts to risk stratify patients to assist management, which is advisable [112]. However the blood metal ion thresholds recommended (3 μg/l and 10 μg/l) appear to have been arbitrarily selected rather than being supported by data [112]. By contrast the newly devised ion thresholds have now been assessed in nearly 1,000 BHR patients from three different European centres. The results of applying these thresholds have been consistent in both the derivation and validation cohorts in that they were effective for identifying patients at low risk of ARMD, and they missed fewer patients with ARMD compared with currently recommended fixed thresholds [23, 112]. It is therefore hoped that clinicians will consider using the newly devised and validated evidence-based thresholds over fixed regulatory thresholds in BHR patients.

It is suggested that the new blood metal ion thresholds can be used clinically to rationalise follow-up resources in unilateral and bilateral BHR patients. Symptomatic patients also require cross-sectional imaging and are likely to remain under regular follow-up (typically annually) whilst symptomatic in line with current recommendations [22-24, 141, 142]. Therefore if symptomatic patients have blood metal ions below the respective ion thresholds and normal cross-sectional imaging, this can provide some reassurance to both the clinician and patient that the symptoms are not related to ARMD. The regularity of follow-up in such patients could therefore be reduced. Symptomatic patients with ions above the respective thresholds will either require more regular surveillance or revision surgery depending on their cross-sectional imaging findings.

The rationalisation of follow-up resources is most relevant in asymptomatic BHR patients, which represents the largest subgroup requiring surveillance. Asymptomatic BHR patients with blood metal ions above the respective ion thresholds require cross-sectional imaging. The use of cross-sectional imaging in asymptomatic BHR patients with blood metal ions below the respective ion thresholds will depend on the patient's risk of ARMD. On the basis of the work presented (Chapter 3.) females should be considered high risk of ARMD and therefore should undergo cross-sectional imaging regardless of blood metal ion concentrations. However males with BHRs are at low risk of ARMD, therefore it is reasonable to recommend that these patients do not need cross-sectional imaging if they are asymptomatic with blood metal ions below the respective thresholds, and normal clinical examinations and radiographs. Such patients would require less regular surveillance. However it is uncertain whether asymptomatic

HRA patients with sub-threshold blood metal ions can definitively be excluded from future surveillance due to a lack of studies with extended follow-up (Chapter 2.). Most BHR patients remain asymptomatic at long-term follow-up [13, 68-70, 72]. Furthermore most BHR patients in this study and previous reports are male (and therefore low risk) [13, 70, 72], and 78% of BHR patients in the present study had blood metal ion concentrations below the devised ion thresholds. Therefore reducing the follow-up frequency in this large BHR cohort will have substantial financial and resource savings (Chapter 2.5.3.). For completeness the regularity of surveillance in various HRA patient subgroups has been formally assessed in a longitudinal study (Chapter 5.).

Future considerations include only sampling cobalt in all BHR patients, rather than both cobalt and chromium, as cobalt produced similar results to maximum cobalt and chromium in bilateral BHRs. Although clinicians wishing to do this would need to adjust the bilateral BHR threshold they employ (cobalt 5.72 μg/l), sampling cobalt alone would save one-third of blood test costs (£10 per test saving) which is substantial. Finally as the optimal blood metal ion thresholds for identifying HRA patients with ARMD is suspected to relate to the clinical performance of the specific implant design (Chapter 4.6.1.), it is recommended that future studies devise similar thresholds for other commonly used HRA implants (such as the Conserve Plus, Recap, and Cormet designs) using the same methodology described here. These thresholds may also prove to be preferable to currently recommended fixed thresholds [23, 112].

4.7. Summary

- Optimal blood metal ion thresholds for discriminating between unilateral and bilateral BHR patients with and without ARMD were dependent on the specific metal ion parameter used (cobalt, chromium, maximum cobalt or chromium, or the cobalt-chromium ratio). The highest discriminatory ability in unilateral BHR patients was achieved with cobalt, and in bilateral BHR patients was achieved with maximum cobalt or chromium.

- The cobalt threshold for discriminating between unilateral BHR patients with and without ARMD providing the optimal diagnostic test characteristics was 2.15 μg/l. The maximum cobalt or chromium threshold for discriminating between bilateral BHR patients with and without ARMD providing the optimal diagnostic test characteristics was 5.52 μg/l.

- Patients with unilateral and bilateral BHRs and blood metal ion concentrations below the respective newly devised thresholds were at a low risk of having ARMD.

- Currently proposed fixed regulatory authority thresholds from the UK MHRA (7 μg/l) and US surgeons (3 μg/l and 10 μg/l) missed more patients with ARMD compared with the new BHR thresholds.

- The findings from the derivation cohort for both unilateral and bilateral BHR patients were successfully validated in a large cohort of BHR patients from two external European centres.

- It is therefore recommended that the newly devised BHR blood metal ion thresholds are used in preference to fixed regulatory authority thresholds when managing patients with BHR implants.

- Along with blood metal ion sampling, hip imaging (radiographs, ultrasound, and MARS-MRI) represents the other commonly used investigation worldwide in the surveillance of HRA patients. The specific roles of these various imaging modalities will be considered in the subsequent chapter.

5. INVESTIGATION – HIP IMAGING*

5.1. Introduction

Worldwide regulatory authorities recommend using radiographs and cross-sectional imaging (ultrasound and MARS-MRI) when investigating HRA patients [22-24, 141, 142]. Radiographs provide information on component position, bone quality, and implant fixation [192]. However given that pseudotumours can be solid or cystic lesions associated with soft-tissue damage and high wear [14, 16], clinicians often favour blood metal ions and cross-sectional imaging over radiographs [22-24, 141, 142]. Radiographic risk factors for pseudotumours have been identified (Chapter 2.2.6.3.), although the significance of femoral neck narrowing remains unclear (Chapter 2.3.2.). However previous studies have been limited by including small numbers of revisions, and by assessing relatively few radiographic parameters [20, 113, 150, 153]. The need for radiographs in HRA follow-up therefore remains unclear. However it is important to establish the role of radiographs during surveillance given their numerous advantages (Chapter 2.3.2.), and the fact that current follow-up recommendations are not evidence-based and very costly (Chapter 2.5.).

---

* Work presented in this chapter has been published in:

Matharu GS, Blanshard O, Dhaliwal K, Judge A, Murray DW, Pandit HG. Patient and radiographic factors help to predict metal-on-metal hip resurfacings with evidence of a pseudotumor. *J Bone Joint Surg Am* 2017; **99**(3): 214-222.

Matharu GS, Mansour R, Dada O, Ostlere S, Pandit HG, Murray DW. Which imaging modality is most effective for identifying pseudotumours in metal-on-metal hip resurfacings requiring revision: ultrasound or MARS-MRI or both? *Bone Joint J* 2016; **98-B**(1): 40-48.

Low AK, Matharu GS, Ostlere SJ, Murray DW, Pandit HG. How should we follow-up asymptomatic metal-on-metal hip resurfacing patients? A prospective longitudinal cohort study. *J Arthroplasty* 2016; **31**(1): 146-151.

Ultrasound or MARS-MRI is currently recommended for assessing peri-prosthetic soft-tissues around HRAs [22-24, 141, 142]. However no preference between modalities has been identified, with the choice left to clinicians. This appears reasonable given both modalities have been acceptable for identifying pseudotumours [25, 199, 202], with the relative merits of both described (Chapter 2.3.3.). Studies comparing ultrasound with MARS-MRI for detecting pseudotumours have largely suggested that ultrasound was effective (Chapter 2.3.3.3.). Although these studies had limitations, the major universal deficiency was either using MARS-MRI as the gold standard, or assuming agreement between modalities was acceptable [201, 229-232]. The benchmark for diagnostic accuracy should be correlation with revision and histopathological findings. Furthermore the role of combined imaging (ultrasound and MARS-MRI) for identifying patients with intra-operative pseudotumours has not been investigated, however it may be important in complex cases.

Currently the role of repeat cross-sectional imaging in the surveillance of HRA patients is unclear. This issue is most relevant in asymptomatic patients, with this subgroup being the most difficult to manage clinically [73, 202]. Most HRA patients remain asymptomatic and do not develop pseudotumours [68, 70, 72]. It is therefore important to identify the patients who do not need regular surveillance and could potentially be discharged. However a smaller but significant number of asymptomatic patients do develop pseudotumours [69, 73, 202]. It is therefore vital to identify such individuals early to prevent bone and soft-tissue damage. Presently there is no consensus on how to follow-up asymptomatic HRA patients between worldwide regulatory authorities. Recommendations range from as per local protocol [23] to annual radiographs and blood metal ions [22], with the latter having substantial cost implications (Chapter

2.5.3.). Well-designed prospective longitudinal studies would inform decisions regarding which asymptomatic patients require monitoring, at what intervals, and when follow-up should be repeated. However few longitudinal studies exist specifically involving asymptomatic HRA patients, with most reporting on small cohorts at short-term follow-up [187, 223, 224]. Identifying which asymptomatic patients subsequently develop pseudotumours at extended follow-up in a large HRA cohort would assist in devising evidence-based recommendations regarding the surveillance required. Conversely if asymptomatic HRA patients that do not develop pseudotumours could be identified based on their initial assessment, it would be possible for these individuals to either undergo less regular surveillance or potentially be discharged from follow-up.

5.2. Objectives

The objectives of this chapter were:

1. To determine whether a combination of patient and radiographic factors predicted HRA patients with evidence of a pseudotumour.
2. To compare the diagnostic test characteristics of ultrasound, MARS-MRI, and combined imaging (ultrasound and MARS-MRI) for identifying HRA patients with intra-operative pseudotumours.
3. To identify risk factors associated with developing new pseudotumours on ultrasound in asymptomatic HRA patients.

These three objectives were investigated in three separate clinical studies described in this chapter.

5.3. Patients and methods: radiographic study

5.3.1. Study design

A retrospective single centre multi-surgeon case-control study was performed at the NOC. All patients receiving primary HRAs at the NOC between 1st June 1999 and 31st December 2009 with a minimum follow-up of 1 year were eligible for inclusion (1,429 hips; Chapter 3.3.1.). Any HRAs performed at other institutions subsequently undergoing revision surgery at the NOC for pseudotumour were also eligible (19 hips). Hips meeting the inclusion criteria were divided into pseudotumour (case) and non-pseudotumour (control) groups.

Information regarding ethical approval (Chapter 3.3.1.), the data source (Chapter 3.3.2.), the surgery performed (Chapter 3.3.3.), and patient follow-up (Chapter 3.3.4.) at the NOC have been detailed. Any HRAs performed at other institutions subsequently undergoing revision surgery at the NOC were also added to the electronic research database at the time of revision (Chapter 3.3.2.).

5.3.2. Pseudotumour group

By 24th August 2015, 111 HRAs implanted at the NOC had been revised for pseudotumour (Table 3.1.), and 19 HRAs performed at other institutions were also revised for pseudotumour. The pseudotumour group therefore included 130 HRAs all undergoing revision surgery for pseudotumour.

All patients in the pseudotumour group were symptomatic and underwent comprehensive investigation before revision (Chapter 3.3.4.). The decision to perform

revision was made by the patient's surgeon based on their symptoms and the investigative findings. All pseudotumours were diagnosed on cross-sectional imaging prior to revision surgery, and were subsequently confirmed both intraoperatively and histopathologically (Chapter 3.3.5.).

5.3.3. Non-pseudotumour group

The non-pseudotumour group included all patients with non-revised HRAs implanted at the NOC, regardless of symptoms, with normal findings on cross-sectional imaging (no evidence of a pseudotumour on ultrasound and/or MARS-MRI). This group included 254 HRAs, of which 128 were symptomatic (median OHS=32; IQR 24-38) and 126 were asymptomatic (median OHS=47; IQR 45-48). Symptomatic patients were investigated during the 2010 and 2012 recalls (Chapter 3.3.4.). Asymptomatic patients were investigated between 2007 and 2008 in an ethically approved study, which originally consisted of 201 HRAs in 158 patients [202]. All asymptomatic patients were assessed with radiographs, ultrasound, and blood metal ions.

5.3.4. Radiographic assessments

Standardised anteroposterior pelvic radiographs were taken with patients positioned supine. Radiographs were stored and accessed using the NOC electronic radiology database (PACS). All radiographic parameters were assessed using a single radiograph for each HRA apart from femoral neck narrowing. The latter also required the radiograph taken immediately following HRA surgery. For the pseudotumour group, the radiograph assessed was the one taken closest to, but immediately before, revision surgery. At this time the hip was symptomatic and the pseudotumour had been

diagnosed radiologically. In the non-pseudotumour group, the radiograph assessed was taken at the time that cross-sectional imaging excluded a pseudotumour.

5.3.4.1. Radiographic abnormalities

Each radiograph was systematically analysed for the presence or absence of abnormalities described in HRAs [192]. These abnormalities included component loosening, osteolysis, femoral neck notching, fracture, dislocation, subluxation, impingement causing bone erosion, and HO [54, 192, 293, 294]. Loosening was defined as the presence of a radiolucent line of 2 mm or greater at the bone-implant interface in any of the three zones described by DeLee and Charnley (acetabular) [293] or Amstutz *et al*. (femoral) [54] (Figure 5.1.). Femoral neck notching was defined as a notch of greater than 2 mm present at the femoral component and femoral neck junction.

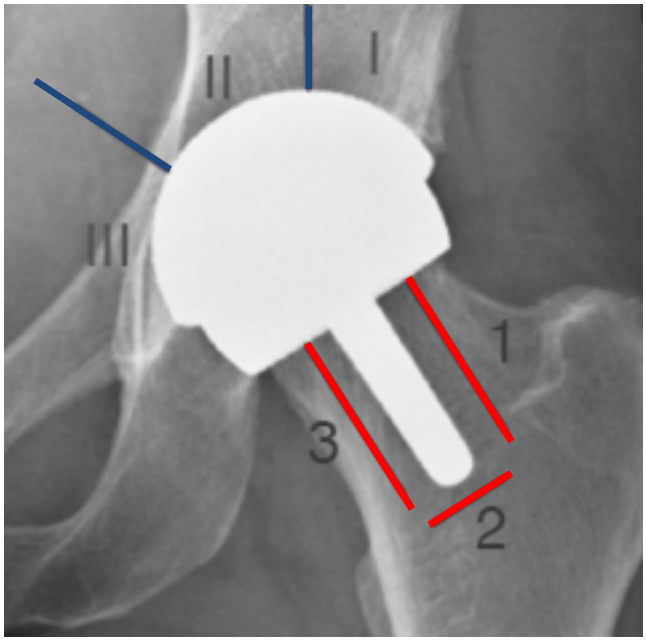


Figure 5.1 Anteroposterior hip radiograph illustrating the three zones on the acetabular side (blue) and the three zones on the femoral side (red), which were used when assessing radiolucent lines.

5.3.4.2. Measuring acetabular component position

Acetabular component inclination and version were measured using the open source image processing software, ImageJ (National Institutes of Health, Bethesda, Maryland, US) [295]. An ellipse was fitted to the open face of the acetabular component (Figure 5.2.). The inclination was the angle measured between the long axis of the ellipse and the pelvic inter-teardrop line (the tangent to the inferior limits of the teardrops). The version angle was calculated using the length and width ellipse measurements as described by Lewinnek (version = $\sin^{-1}$(width/length)) [296]. This method does not differentiate between anteversion and retroversion, which requires a lateral hip radiograph. Acetabular components were considered malpositioned if one or both parameters were outside the recommended optimal zone for HRAs (inclination of 35°-55° and anteversion 10°-30°) [150].

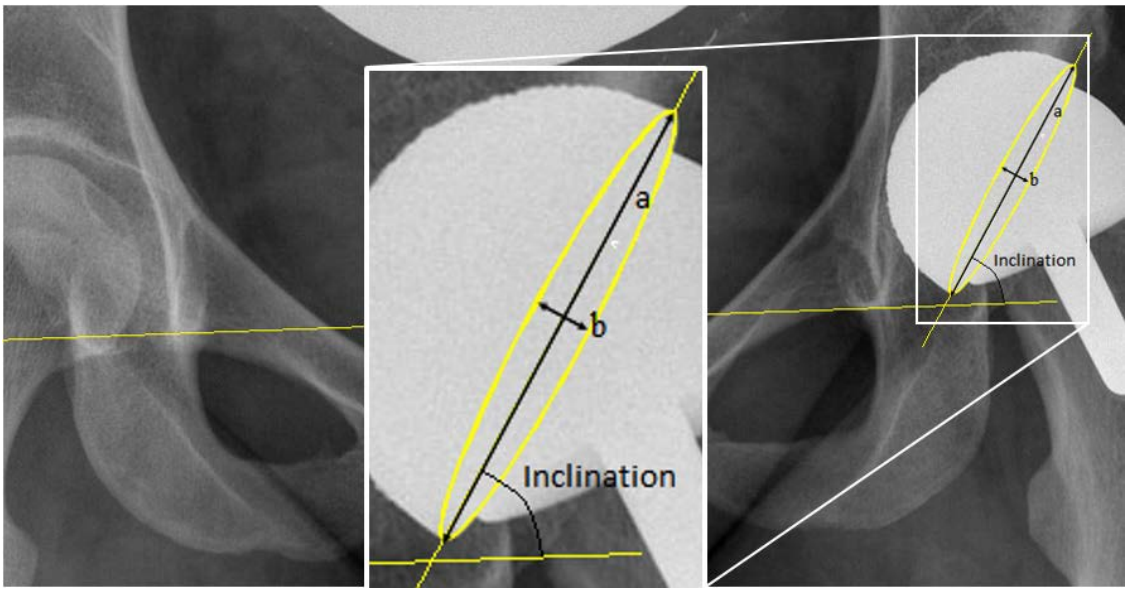


Figure 5.2 Method for measuring acetabular component inclination and version on an anteroposterior pelvic radiograph using ImageJ software.
An ellipse was fitted to the open face of the acetabular component. The length (a) and width (b) of the ellipse were required to calculate the version angle. The yellow horizontal line represents the pelvic inter-teardrop line, which was required to calculate the inclination angle (measured relative to the long axis of the ellipse).

5.3.4.3. Measuring femoral neck narrowing

Femoral neck narrowing was assessed as described by Hing *et al*. [197]. The femoral neck diameter was measured on each radiograph at the junction of the neck and the femoral component (Figure 5.3.). This measurement was divided by the measured femoral component diameter, which allowed correction for radiographic magnification. The difference between these ratios from the most recent radiograph and the radiograph taken immediately following HRA allowed calculation of the degree of femoral neck narrowing since the index procedure. The degree of narrowing was expressed as a percentage of the initial neck diameter.

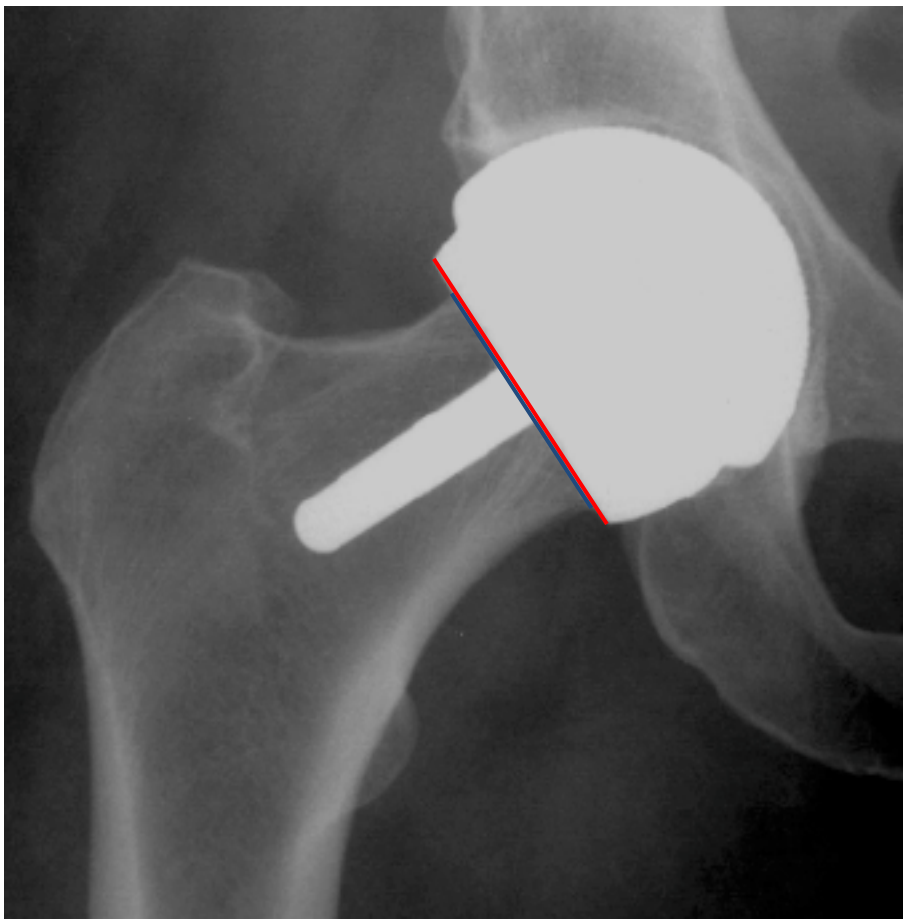


Figure 5.3 Method for measuring femoral neck narrowing on an anteroposterior hip radiograph. Blue line = femoral neck diameter (at the junction of the femoral neck and the femoral component). Red line = femoral component diameter.

5.3.4.4. Observer assessments

All radiographs were assessed by two independent observers. The second observer was an orthopaedic surgical trainee. Each observer assessed the radiographs in isolation, with these assessments performed at a different location and during a different time period to the other observer. The observers were both given a list of radiographs to assess, with this list containing only the necessary details to perform the assessment (patient identifiable information, hip laterality, and date of radiograph for assessment). The list was also randomly ordered to ensure that the radiographs from either the pseudotumour group or non-pseudotumour group were not all grouped together. Therefore radiographs were assessed in a random sequence, with both observers blinded to all clinical information that was practically possible, including the study group. Both observers assessed the same 100 radiographs, with the remainder divided between the two observers for assessment. For the presence or absence of abnormalities in the 100 commonly assessed radiographs, inter-observer agreement was excellent (Cohen kappa statistic=0.88-1.00) [297]. Any discrepancies between the two observers regarding the presence or absence of abnormalities were settled by a third observer who was an orthopaedic surgeon, with this final assessment used for analyses. For continuous radiographic data, intra-class correlation coefficients between the two observers were excellent: inclination=0.979 (95% CI 0.955-0.990), version=0.968 (95% CI 0.947-0.988), and femoral neck narrowing=0.941 (95% CI 0.861-0.987). When both observers had assessed the same radiograph the mean of the two observer measurements were used for the analyses of continuous radiographic variables.

5.3.5. Statistical analysis

The study outcome of interest was a binary variable (hip with or without a pseudotumour). Patient factors (gender, age, and implant design) and radiographic factors (Chapter 5.3.4.) were compared between groups. Numerical data were compared using either two-sided unpaired t-tests (parametric data) or the Wilcoxon rank-sum test (non-parametric data), with categorical data compared using either the Chi-squared test or Fisher's exact test.

Logistic regression modelling was used to identify predictors of outcome. Univariable models explored the association between each predictor and outcome. For continuous predictors, linearity was assessed using likelihood ratio tests, with data grouped if the relationship between a predictor and the outcome was non-linear. A multivariable logistic regression model was formulated using backward selection, with predictors retained in the final model if $p<0.10$. Regression diagnostics were assessed to ensure that all assumptions underlying the model were met [298, 299]. Internal validation of the final multivariable model was performed using calibration, discrimination, and bootstrapping (Appendix 12.) [299-301]. Predictors from the final multivariable model were used to develop a clinical risk scoring tool which quantified a patient's risk of having a pseudotumour (Appendix 12.) [300-302].

5.4. Results: radiographic study

5.4.1. Patient factors

The final cohort included 384 HRAs in 329 patients, with 130 hips in the pseudotumour group and 254 hips in the non-pseudotumour group (Table 5.1.). Compared with the

non-pseudotumour group, the pseudotumour group was significantly more likely to be younger and female. There were also significant differences in HRA implant design between the groups.

Covariate	All hips (n=384; 100%)	Non-pseudotumour group (n=254; 66.1%)	Pseudotumour group (n=130; 33.9%)	p-value
Gender				
Male	204 (53.1)	157 (61.8)	47 (36.2)	**<0.001**
Female	180 (46.9)	97 (38.2)	83 (63.9)	
Mean age at radiograph (range) in years	59.4 (30.3-88.2)	60.2 (30.3-88.2)	58.0 (37.5-77.6)	**0.0286**
Mean time from primary to radiograph (range) in years	5.1 (0.25-15.1)	4.8 (2.0-15.1)	5.8 (0.25-14.1)	**<0.001**
Mean time from primary to revision (range) in years	NA	NA	6.3 (0.31-15.5)	NA
Primary Implant				
BHR	180 (46.9)	117 (46.1)	63 (48.5)	**0.022**
Conserve Plus	174 (45.3)	122 (48.0)	52 (40.0)	
Recap	26 (6.8)	15 (5.9)	11 (8.5)	
Cormet	4 (1.0)	0 (0.0)	4 (3.1)	
Median blood metal ion concentrations (IQR) in µg/l				
Cobalt	2.2 (1.5-4.6)	2.1 (1.4-3.6)	6.3 (3.1-22.5)	**<0.0001**
Chromium	2.4 (1.3-4.9)	2.2 (1.3-4.1)	7.6 (2.6-29.2)	**<0.0001**

Table 5.1 Patient demographics for all hip resurfacing arthroplasties, and subdivided by study group.
BHR = Birmingham Hip Resurfacing; IQR = interquartile range; NA = not applicable. P-values <0.05 in bold text.

5.4.2. Radiographic factors

The pseudotumour group was significantly more likely to have abnormal radiographic findings compared with the non-pseudotumour group (80.0% vs. 63.4%; p=0.001) (Table 5.2.). The abnormalities more frequently observed in the pseudotumour group

compared with the non-pseudotumour group were acetabular osteolysis (21.5% vs. 4.3%; p<0.001), femoral osteolysis (20.0% vs. 1.6%; p<0.001), acetabular component loosening (13.1% vs. 4.7%; p=0.003), higher acetabular inclination (mean 49.5° vs. 46.0°; p=0.0013), femoral fracture (3.1% vs. 0%; p=0.013), dislocation (2.3% vs. 0%; p=0.038), and subluxation (2.3% vs. 0%; p=0.038).

Acetabular component anteversion (mean 15.1° vs. 14.7°; p=0.697) was not different between the pseudotumour and non-pseudotumour groups. However the effect of acetabular component anteversion on the outcome was non-linear therefore this covariate was subsequently grouped (Chapter 5.4.3.). Femoral neck narrowing (mean 5.0% vs. 4.5%; p=0.556) was not different between the pseudotumour and non-pseudotumour groups. The non-pseudotumour group was significantly more likely to have HO compared with the pseudotumour group (15.7% vs. 2.3%; p<0.001).

Covariate	All hips (n=384; 100%)	Non-pseudotumour group (n=254; 66.1%)	Pseudotumour group (n=130; 33.9%)	p-value
Abnormal radiograph (≥1 abnormality)	265 (69.0%)	161 (63.4%)	104 (80.0%)	**0.001**
Cup position				
Mean inclination in ° (range)	47.2 (21.3-83.5)	46.0 (21.3-67.6)	49.5 (24.3-83.5)	**0.0013**
Mean anteversion in ° (range)	14.9 (1.2-40.3)	14.7 (1.2-40.3)	15.1 (1.2-39.5)	0.697
Malpositioned	183 (47.7%)	113 (44.5%)	70 (53.8%)	0.062
Femoral neck narrowing				
Mean narrowing (range)	4.7% (0.3%-22.2%)	4.5% (0.3%-22.2%)	5.0% (0.3%-17.4%)	0.556
>10% narrowing	13 (3.4%)	7 (2.8%)	6 (4.6%)	0.377
Osteolysis				
Acetabular	39 (10.2%)	11 (4.3%)	28 (21.5%)	**<0.001**
Femoral	30 (7.8%)	4 (1.6%)	26 (20.0%)	**<0.001**
Loosening				
Acetabular	29 (7.6%)	12 (4.7%)	17 (13.1%)	**0.003**
Femoral	13 (3.4%)	7 (2.8%)	6 (4.6%)	0.377
Heterotopic ossification	43 (11.2%)	40 (15.7%)	3 (2.3%)	**<0.001**
Femoral neck notching	12 (3.1%)	9 (3.5%)	3 (2.3%)	0.758
Fracture				
Acetabular	0 (0%)	0 (0%)	0 (0%)	NA
Femoral	4 (1.0%)	0 (0%)	4 (3.1%)	**0.013**
Dislocation	3 (0.78%)	0 (0%)	3 (2.3%)	**0.038**
Subluxation	3 (0.78%)	0 (0%)	3 (2.3%)	**0.038**
Impingement	0 (0%)	0 (0%)	0 (0%)	NA

Table 5.2 Radiographic features for all hip resurfacing arthroplasties, and subdivided by study group.
NA = not applicable. P-values <0.05 in bold text.

5.4.3. Logistic regression modelling

Five factors (female gender, high acetabular component inclination, acetabular osteolysis, femoral osteolysis, and acetabular component loosening) significantly predicted being in the pseudotumour group in both the univariable and multivariable logistic regression analyses (Table 5.3.). There were two factors (acetabular component anteversion of 5° or greater and HO) which significantly predicted being in the non-pseudotumour group.

Covariate	Univariable model Odds Ratio (95% CI)	p-value	Multivariable model Odds Ratio (95% CI)	p-value
Gender				
Male	1.00	**<0.001**	1.00	**<0.001**
Female	3.00 (1.92-4.68)		3.14 (1.85-5.35)	
Age at radiograph (per year)	0.98 (0.95-1.00)	**0.044**	*	
Primary Implant				
BHR	1.00		*	
Conserve Plus	0.79 (0.50-1.25)	0.309		
Other (Recap and Cormet)	1.89 (0.86-4.11)	0.111		
Cup position				
Inclination (per °)	1.04 (1.02-1.07)	**0.001**	1.04 (1.01-1.07)	**0.006**
Anteversion				
<5°	1.00		1.00	
5° to <10°	0.37 (0.17-0.82)	**0.014**	0.31 (0.12-0.77)	**0.012**
≥10°	0.49 (0.25-0.94)	**0.033**	0.32 (0.15-0.70)	**0.004**
Femoral neck narrowing	1.03 (0.98-1.08)	0.259	*	
Osteolysis				
Acetabular	5.91 (2.82-12.4)	**<0.001**	5.06 (2.14-12.0)	**<0.001**
Femoral	14.4 (4.88-42.7)	**<0.001**	17.8 (5.09-62.2)	**<0.001**
Loosening				
Acetabular	3.15 (1.41-7.01)	**0.005**	3.35 (1.34-8.35)	**0.009**
Femoral	1.73 (0.57-5.26)	0.335	*	
Heterotopic ossification	0.13 (0.04-0.42)	**0.001**	0.19 (0.05-0.72)	**0.015**
Femoral neck notching	0.65 (0.17-2.45)	0.524	*	
Femoral fracture	1	NA	*	
Dislocation	1	NA	*	
Subluxation	1	NA	*	

Table 5.3 Logistic regression models identifying patient and radiographic factors predictive of pseudotumour revision.
BHR = Birmingham Hip Resurfacing; CI = confidence interval; NA = not applicable. Impingement and acetabular fracture were not analysed as no events occurred in the whole cohort.
The effect of acetabular component anteversion (p-value for non-linearity=0.0144) on outcome was non-linear, therefore this covariate was grouped. The effect of both patient age at radiograph (p-value for non-linearity=0.937) and acetabular component inclination (p-value for non-linearity=1.0) on outcome were linear, therefore these covariates were analysed as continuous variables. P-values <0.05 in bold text. *The final multivariable logistic regression model only included patient and radiographic factors where p<0.10.

5.4.4. Internal validation of the final multivariable model

The final multivariable model for distinguishing between HRAs with and without evidence of a pseudotumour was well calibrated (p=0.589) (Figure 5.4.) and demonstrated good discriminatory ability with an AUC of 0.801 (95% CI 0.752-0.849; 74.4% sensitivity and 71.7% specificity). Bootstrap validation of the final model provided a bias-corrected AUC of 0.784. The results of the clinical risk scoring tool have been summarised (Appendix 12.).

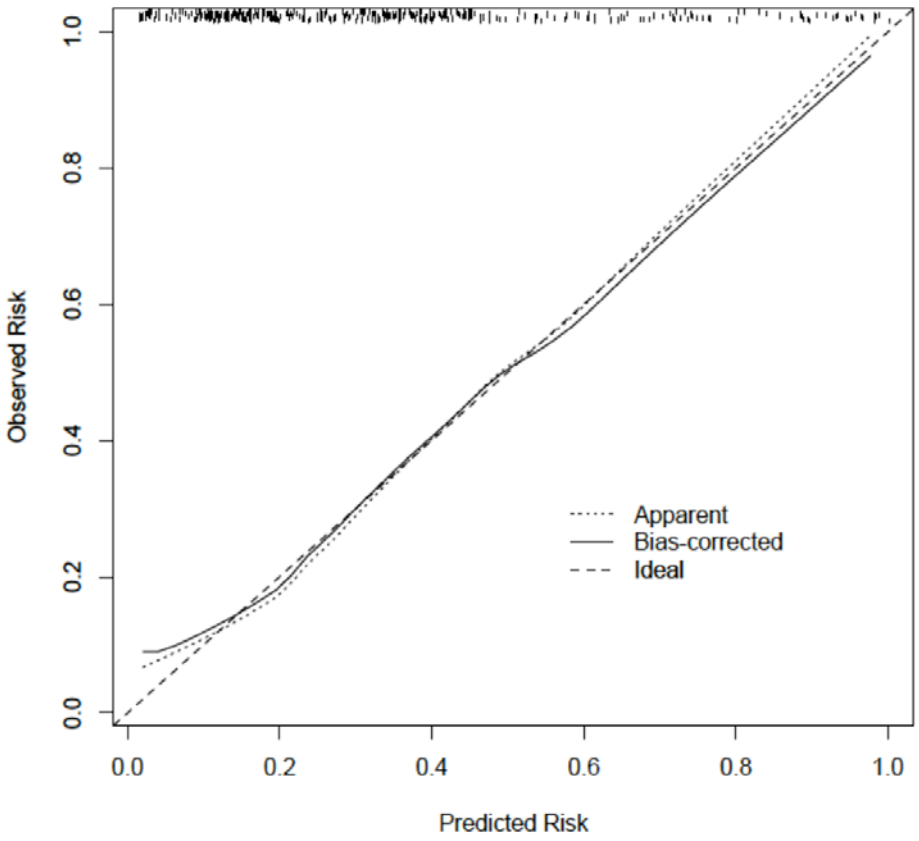


Figure 5.4 Plot to assess calibration of the final multivariable model.
Calibration measures how closely predicted risk (obtained from the final multivariable model) agrees with observed risk (0=non-pseudotumour group; 1=pseudotumour group). Calibration was assessed for each one-tenth of predicted risk using ten equally sized groups. For each group the predicted risk was plotted against the mean observed risk. The 'ideal' model has a line with a slope of 1 going through the graph origin. The 'apparent' line represents the final multivariable model from this study and demonstrated that the model was well calibrated throughout the complete range of pseudotumour risk. The 'bias-corrected' line was obtained from 200 bootstrap repetitions used to internally validate the study model. The 'bias-corrected' line closely follows the 'apparent' line for the final multivariable model.

5.5. Study strengths and limitations: radiographic study

The main study strengths are the large cohort size (n=384) and the assessment of all major hip radiographic parameters. Previous studies are limited by being small (between 42-163 hips) with few failures for comparison, and by only assessing selected radiographic factors such as acetabular component position or femoral neck narrowing [113, 150, 153, 197]. Both the pseudotumour and non-pseudotumour groups were large in this study, which is further strengthened by the latter including both symptomatic and asymptomatic patients with no cross-sectional imaging evidence of a pseudotumour. Such patients with non-revised HRAs are typical of the many currently under regular surveillance worldwide, which therefore increases the generalisability of the findings. The final multivariable model underwent robust internal validation and was not overfitted.

Limitations include this being a retrospective study and therefore the findings are subject to bias for this type of design. The radiographs in the pseudotumour group were taken significantly later after primary surgery compared with the non-pseudotumour group, which should be considered when interpreting the findings. Conclusive statements cannot be made about the significance of femoral neck narrowing as only 49% of hips had adequate quality immediate postoperative radiographs. Although inter-observer agreement was assessed intra-observer agreement was not. Due to the large number of radiographs requiring assessment, most radiographs were assessed by only one observer which may have influenced the findings. The software used to measure acetabular component position was unable to distinguish between component anteversion and retroversion. This problem exists with other types of software, such as Einsel-Bild-Roentgen analysis (EBRA), and requires a lateral hip radiograph (not

routinely performed) or CT scanning to establish component version. When such additional imaging was available these were used to confirm the version, however this limitation may have influenced the findings regarding acetabular component version. This study only predicts the presence or absence of a pseudotumour at the time of radiography, and not subsequent pseudotumour development, which would require a longitudinal study. It is possible that the study findings may have been influenced by not matching the pseudotumour and non-pseudotumour groups. However matching would have substantially reduced the study cohort given the known baseline differences between these patients (Chapter 3.). Finally, the study findings cannot be assumed to apply in MoM THA patients, and the final multivariable model requires external validation.

5.6. Patients and methods: ultrasound vs. MRI study

5.6.1. Study design

A retrospective diagnostic accuracy study was performed at the NOC. All patients undergoing revision surgery of HRAs for any indication between 1st December 2000 and 30th November 2014 were eligible for inclusion. Patients having their primary surgery at a different centre were included provided they satisfied the other criteria. During this time 216 consecutive HRA revisions (199 patients) were performed for all indications. Information regarding ethical approval (Chapter 3.3.1.), the data source (Chapters 3.3.2. and 5.3.1.), the primary surgery performed (Chapter 3.3.3.), and patient follow-up (Chapter 3.3.4.) at the NOC have been detailed.

5.6.2. Study inclusion and exclusion criteria

Patients undergoing both an ultrasound and MARS-MRI examination of the same hip before revision were eligible for inclusion (53 hips in 51 patients). Revisions did not necessarily have to be performed for pseudotumour as the ability of both imaging modalities to exclude a pseudotumour was being investigated. Imaging was performed at least one-year following HRA, because of difficulties in distinguishing post-surgical changes from soft-tissue abnormalities early on (no hips excluded) [199]. Furthermore both scans needed to be performed within one-year before revision, given recent observations that MARS-MRIs performed more than one-year before revision have poor sensitivity (29%) for detecting intra-operative pseudotumours (five hips excluded) [220]. Ultrasound and MARS-MRI needed to be performed in each patient within three months of each other for inclusion (eight hips excluded), as this interval has been considered acceptable previously [201, 220, 229-231]. The final cohort included 40 HRAs in 39 patients undergoing both ultrasound and MARS-MRI before revision surgery (Table 5.4.).

Covariate		40 hips in 39 patients
Gender	Male / Female	17 (42.5%) / 23 (57.5%)
Age at primary	Mean (range) in years	52.8 (28.7-71.8)
Body mass index	Mean (range) in kg/m^2	29.9 (22.2-50.1)
Metal-on-metal hip arthroplasty	Unilateral / Bilateral	23 (57.5%) / 17 (42.5%)
Hip resurfacing arthroplasty design	Birmingham Hip Resurfacing Conserve Plus Recap Articular Surface Replacement	21 (52.5%) 15 (37.5%) 3 (7.5%) 1 (2.5%)
Component sizes	Mean (range) in mm	Femoral 47 (40-54) Acetabular 53 (46-60)
Acetabular component position	Inclination - Mean (range) in ^{o} Anteversion - Mean (range) in ^{o}	49.6 (34.0-74.1) 10.8 (2.0-33.0)
Abnormal radiographic features	Acetabular loosening Femoral loosening Acetabular osteolysis Femoral osteolysis Femoral neck narrowing (>10%) Femoral neck fracture Heterotopic ossification	7 (17.5%) 2 (5.0%) 9 (22.5%) 4 (10.0%) 6 (15.0%) 2 (5.0%) 3 (7.5%)
Blood metal ion concentration	Median (interquartile range) in µg/l	Cobalt 12.9 (2.4-61.5) Chromium 11.4 (3.4-54.9)
Time between ultrasound and MRI	Mean (range) in days	14.6 (0-90)
Time between initial imaging and revision	Mean (range) in months	5.3 (0.06-12)
Time from primary to revision	Mean (range) in years	6.3 (1.7-11.5)
Age at revision	Mean (range) in years	59.1 (40.0-77.7)
Indication for revision surgery	Pseudotumour ARMD Acetabular loosening Femoral loosening	33 (82.5%) 5 (12.5%) 1 (2.5%) 1 (2.5%)

Table 5.4 Summary of patient demographics for revised hip resurfacing arthroplasties. ARMD = adverse reactions to metal debris; MRI = magnetic resonance imaging.

5.6.3. Cross-sectional imaging performed

Details regarding cross-sectional imaging performed (Chapter 3.3.4.1.) and the specific ultrasound and MARS-MRI protocols (Appendix 10.) have been described. Ultrasound was used first line with MARS-MRI reserved for equivocal or complex cases. All examinations were undertaken after obtaining informed patient consent. An

experienced musculoskeletal radiologist (who had performed over 500 ultrasounds for suspected HRA problems) performed all ultrasound examinations reviewed in this study, which has been recommended for achieving results comparable with MARS-MRI [229]. Sagittal ultrasound views along the axis of the femoral neck, anteriorly and posteriorly, and coronal views over the greater trochanter were routinely stored. Any abnormalities identified on ultrasound were examined in multiple planes, with images subsequently stored. All stored static and dynamic images were available on the hospitals electronic radiology database (PACS). Ultrasound examinations were performed within 15 minutes, with MARS-MRI examinations performed within 12 minutes.

5.6.4. Assessment of cross-sectional imaging

In April 2015 all ultrasound and MARS-MRI images (80 scans in total) were assessed by an experienced musculoskeletal radiologist who was blinded to all of the clinical information, including that patients were subsequently revised. This assessor was also different to the radiologist who performed the original ultrasound examinations. Images were reviewed in a random sequence to ensure that the radiologist was blinded to the results from the other modality when assessing the same HRA.

Each scan was assessed using a standard data collection form (Appendix 13.). The radiologist categorically stated whether scans were normal or abnormal, and whether or not a pseudotumour was present. For both imaging modalities a pseudotumour was defined as a cystic, solid, or mixed mass in continuity with the hip joint [16, 216]. For all pseudotumours, the volume (product of the maximum recorded dimension in each of three orthogonal planes in centimetres), consistency (solid, cystic, or mixed) and

location in relation to the prosthesis were recorded. Any other imaging abnormalities were also recorded, including effusions (distance of greater than 4 mm between the femoral neck and capsule) [199], isolated bursal distension and/or thickening (trochanteric and/or iliopsoas), muscle atrophy (involving gluteus medius, gluteus minimus, and iliopsoas), and tendon abnormalities (avulsion, atrophy, and ossification).

5.6.5. Assessment of revision procedures

All revisions were performed by one of six consultant arthroplasty surgeons. As described earlier, the threshold for revising a HRA has changed over the study period (Chapter 3.6.1. and 5.3.2.). All revision operation records were retrospectively assessed by one observer who was blinded to the imaging results. Findings were extracted from the records to determine the presence or absence of a pseudotumour, with intra-operative pseudotumours diagnosed as described (Chapter 3.3.5.). Pseudotumour location and consistency (solid, cystic, or mixed) were recorded, with the latter determined from the surgeon's observations of wall thickness and lesion content [220]. All abnormalities described other than pseudotumours were also recorded. Revisions were considered to be for ARMD rather than pseudotumour when no pseudotumour was present but features such as metallosis, synovitis, tissue damage and/or necrosis were observed with histopathology confirming a reaction to metal debris (Chapter 2.2.1.) [14, 98, 99, 220].

5.6.6. Statistical analysis

The diagnostic test characteristics (sensitivity, specificity, PPV, and NPV, all with 95% CIs) for identifying HRAs with intra-operative pseudotumours were calculated for ultrasound alone, MARS-MRI alone, and ultrasound and MARS-MRI combined. The

reference standard was the presence or absence of a histopathologically proven pseudotumour at revision surgery. Agreement of the imaging modalities (both alone and in combination) with the operative findings was assessed using the kappa statistic (κ), with these values representing slight (0-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80) and almost perfect (0.81-1.00) agreement [297].

5.7. Results: ultrasound vs. MRI study

5.7.1. Intra-operative findings

Of 40 HRAs undergoing revision surgery, 33 pseudotumours (82.5% prevalence; 95% CI 67.2%-92.7%) were diagnosed intra-operatively and confirmed histopathologically. The remaining hips without a pseudotumour were revised for ARMD (n=5) and component loosening (n=2). Intra-operative pseudotumours were most commonly cystic (n=23; 70%) or mixed (n=8; 24%), and were most frequently located anteriorly (n=19; 58%) followed by posteriorly (n=5; 15%).

5.7.2. Cross-sectional imaging findings

The abnormalities detected on ultrasound and MARS-MRI have been summarised (Table 5.5.). The prevalence of pseudotumour detected by both imaging modalities alone was identical at 85.0% (95% CI 70.2%-94.3%), and for ultrasound and MARS-MRI combined was 90.0% (95% CI 76.3%-97.2%).

	Ultrasound	**MRI**
Hips with ≥1 imaging abnormality	38 (95%)	38 (95%)
Hips with any pseudotumour on imaging	34 (85%)	34 (85%)
Median pseudotumour volume (IQR)	22.6 cm^3 (6.6 cm^3 – 111.2 cm^3)	42.8 cm^3 (6.9 cm^3 – 131.7 cm^3)
Pseudotumour consistency	9 (26%) Cystic 16 (48%) Mixed 9 (26%) Solid	16 (48%) Cystic 13 (38%) Mixed 5 (15%) Solid
Pseudotumour location	18 (53%) Anterior only 10 (29%) Anterior + posterior, lateral, or medial 5 (15%) Posterior only, or postero-lateral 1 (3%) Medial only	14 (41%) Anterior only 8 (24%) Anterior + posterior, lateral, or medial 8 (24%) Posterior only, or postero-lateral 3 (9%) Lateral only 1 (3%) Medial only
Details of non-pseudotumour abnormalities	15 (38%) joint effusions (1-8.1cm^3) 14 (35%) trochanteric and/or iliopsoas bursitis / distension 2 (5%) trochanteric bursal thickening 1 (3%) sub-gluteus medius bursitis 1 (3%) heterotopic ossification	25 (63%) trochanteric and/or iliopsoas bursitis / distension 19 (48%) muscle atrophy (gluteus medius, gluteus minimus, iliopsoas, iliacus, obturator) 8 (20%) tendon abnormalities (including tears, avulsions and/or ossification of gluteus medius, gluteus minimus, and hamstrings) 6 (15%) bony erosions (acetabulum, proximal femur, femoral neck, pubis) 4 (10%) joint effusions (1-13.2 cm^3)

Table 5.5 Findings identified on cross-sectional imaging performed prior to revision surgery. IQR = interquartile range; MRI = magnetic resonance imaging.

Imaging findings agreed with the intra-operative presence or absence of a pseudotumour as follows (values also represent the accuracy of imaging): 82.5% (n=33) for ultrasound alone, 87.5% (n=35) for MARS-MRI alone, and 92.5% (n=37) for ultrasound and MARS-MRI combined. There were nine hips (22.5%) where the findings from either ultrasound (n=4; Figure 5.5.), or MARS-MRI (n=2), or both modalities (n=3; Figure 5.6.) did not agree with the intra-operative findings (Table 5.6.).

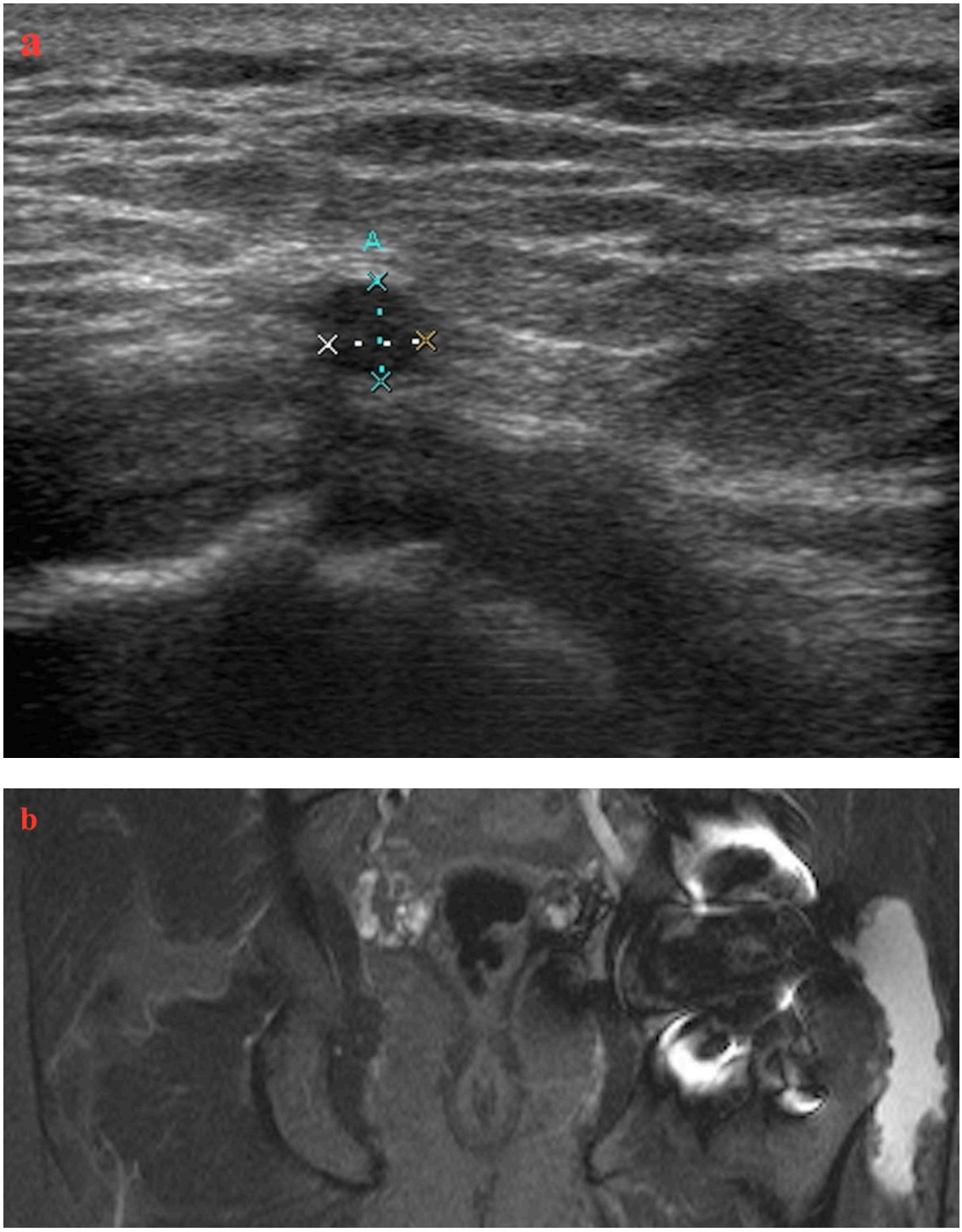
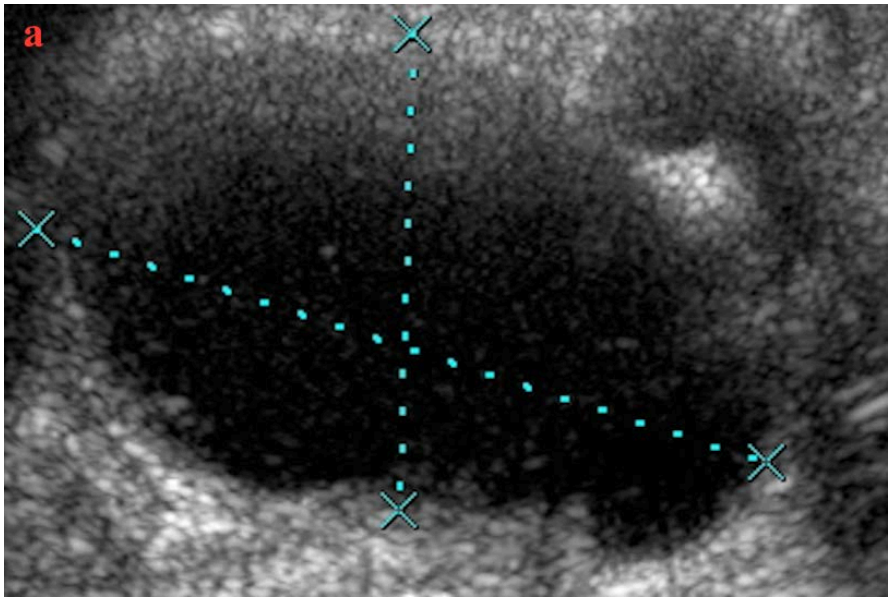


Figure 5.5 Disagreement of ultrasound with the presence of an intra-operative pseudotumour which was also confirmed on MARS-MRI.
A 60 year-old male with a Conserve Plus hip resurfacing arthroplasty underwent revision two years later. A large cystic pseudotumour was identified intra-operatively and confirmed histopathologically; a) ultrasound imaging demonstrated a small effusion and a separate 4 mm fluid collection laterally; however there was no pseudotumour. b) MARS-MRI was performed 61 days following ultrasound and confirmed a predominantly lateral cystic pseudotumour of 216 cm^3 (coronal image shown).



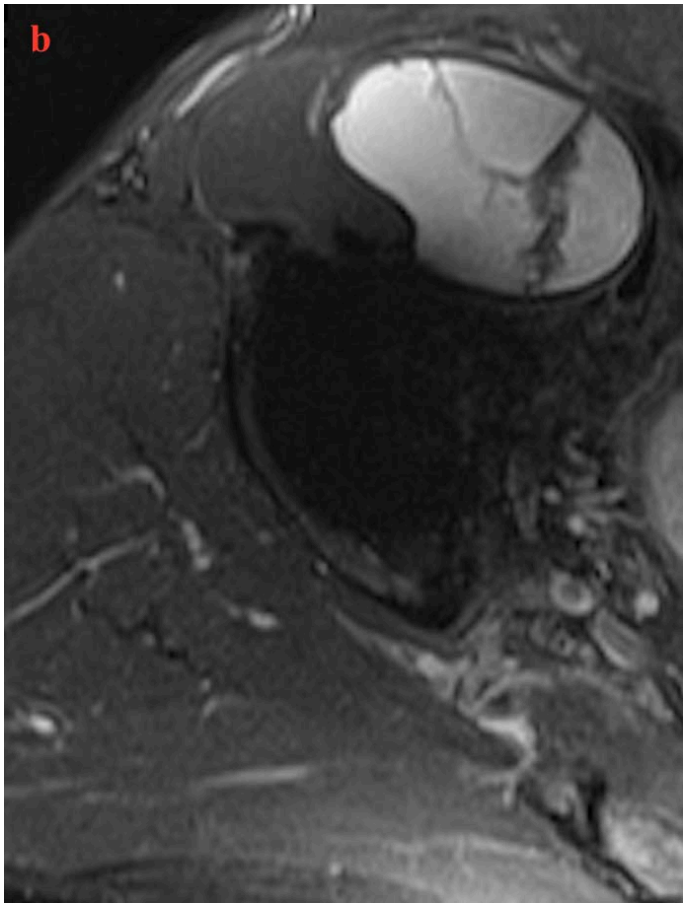


Figure 5.6 Disagreement of both ultrasound and MARS-MRI with the absence of a pseudotumour intra-operatively.
A 59 year-old male with a Recap hip resurfacing arthroplasty underwent revision 4.9 years later. At revision there was significant metallosis but no collection of fluid or pseudotumour. Histopathology confirmed a reaction to metal debris. Ultrasound demonstrated separate a) anterior, and posterolateral cystic pseudotumours with a total volume of 288 cm^3. MARS-MRI was performed 47 days following ultrasound and also demonstrated separate b) anterior, and posterior mixed pseudotumours with a total volume of 260 cm^3 (axial image shown).

Hip number/ Patient age at revision (years) / gender	Primary implant / time in-situ (years)	Time between ultrasound and MRI (days)	Time between imaging and revision (years)	Ultrasound findings	MRI findings	Intra-operative revision + histopathology findings
1 50.6 F	Conserve 1.7	29	0.5	Solid pseudotumour (anterior / 31 cm^3)	Fluid in trochanteric bursa. No pseudotumour	Large cystic pseudotumour (lateral)
2 33.0 F	BHR 4.3	37	0.4	Mixed pseudotumour (anterior / 9 cm^3)	Mixed pseudotumour (lateral / 33 cm^3)	ARMD with metallosis & granulation tissue. No pseudotumour
3 72.1 F	Conserve 3.6	17	1.0	No pseudotumour. Small effusion	Cystic pseudotumour (lateral / 103 cm^3)	Large mixed pseudotumour (postero-lateral)
4 74.3 M	BHR 2.5	16	0.3	No pseudotumour. Trochanteric bursal distension	Cystic pseudotumour (anterior / 114 cm^3)	Large cystic pseudotumour (anterior)
5 65.6 F	ASR 4.0	40	0.3	Mixed pseudotumour (anterior / 50 cm^3)	Normal scan	ARMD with joint effusion, synovitis & metallosis. No pseudotumour
6 54.7 M	Recap 7.0	0	0.3	Mixed pseudotumour (anterior / 23 cm^3)	Cystic pseudotumour (anterior / 5.4 cm^3)	ARMD with acetabular osteolysis, synovitis & metallosis. No pseudotumour
7 59.3 M	Recap 4.9	47	0.4	Cystic pseudotumour (anterior & postero-lateral / 288 cm^3)	Mixed pseudotumour (anterior & posterior / 260 cm^3)	ARMD with metallosis. No pseudotumour
8 59.7 M	Conserve 2.0	61	0.8	No pseudotumour. Small effusion and 4 mm fluid collection laterally	Cystic pseudotumour (anterior & lateral / 216 cm^3)	Large cystic pseudotumour (lateral)
9 55.4 F	BHR 10.2	0	0.6	Solid pseudotumour (anterior / 1 cm^3)	Normal scan	Small cystic pseudotumour (anterior)

Table 5.6 Cases where either or both cross-sectional imaging modalities disagreed with the presence or absence of an intra-operative pseudotumour.
ARMD = adverse reactions to metal debris; BHR = Birmingham Hip Resurfacing; F = female; M = male; MRI = magnetic resonance imaging.

5.7.3. Diagnostic test characteristics and kappa agreement: imaging vs. intra-operative findings

The diagnostic test characteristics for identifying the presence or absence of intra-operative pseudotumours for ultrasound alone and MARS-MRI alone reached similar sensitivities (90.9% vs. 93.9%) and PPVs (88.2% vs. 91.2%). Higher specificities (57.1% vs. 42.9%) and NPVs (66.7% vs. 50.0%) were achieved with MARS-MRI alone compared with ultrasound alone (Table 5.7.). When the findings from ultrasound and MARS-MRI were combined, the sensitivity and NPV both reached 100% with maintenance of the specificity (57.1%) and PPV (91.7%) (Appendix 14.). Similar findings were observed for agreement between the imaging modalities and the intra-operative presence or absence of a pseudotumour. Agreement was substantial (κ=0.69) for ultrasound and MARS-MRI combined, moderate for MARS-MRI alone (κ=0.54), and fair for ultrasound alone (κ=0.36).

	Ultrasound alone	**MRI alone**	**Ultrasound and MRI combined**
Area under the curve (%)	67 (46-87)	76 (55-96)	79 (59-98)
Sensitivity (%)	90.9 (75.7-98.1)	93.9 (79.8-99.3)	100.0 (89.4-100.0)
Specificity (%)	42.9 (9.9-81.6)	57.1 (18.4-90.1)	57.1 (18.4-90.1)
Positive predictive value (%)	88.2 (72.5-96.7)	91.2 (76.3-98.1)	91.7 (77.5-98.2)
Negative predictive value (%)	50.0 (11.8-88.2)	66.7 (22.3-95.7)	100.0 (39.8-100.0)
Kappa statistic (*κ*)	0.36 (-0.02-0.74)	0.54 (0.19-0.90)	0.69 (0.36-1.0)

Table 5.7 Diagnostic test characteristics for cross-sectional imaging identifying intra-operative pseudotumours in revised hip resurfacing arthroplasties.
MRI = magnetic resonance imaging. Values in brackets represent the 95% confidence intervals.

Combined imaging resulted in three hips with false positive findings (both modalities demonstrated a pseudotumour which was not confirmed intra-operatively) (Table 5.6.).

Although these cases may represent incorrect imaging diagnoses, given the retrospective study design these cases could alternatively reflect inaccuracies in surgeon reporting, or even inadvertent pseudotumour decompression during the surgical approach that was not subsequently recorded. Excluding these three cases improved the diagnostic test characteristics and agreement for all imaging combinations: ultrasound alone (sensitivity 90.9%, specificity 75.0%, PPV 96.8%, NPV 50.0%, AUC 83%, κ=0.54 representing moderate agreement); MARS-MRI alone (sensitivity 93.9%, specificity 100%, PPV 100%, NPV 66.7%, AUC 97%, κ=0.77 representing substantial agreement); combined imaging (sensitivity 100%, specificity 100%, PPV 100%, NPV 100%, AUC 100%, κ=1.0 representing perfect agreement).

5.8. Study strengths and limitations: ultrasound vs. MRI study

The main study strength was using the combination of intra-operative and histopathological findings as the reference standard for establishing the presence or absence of a pseudotumour. Previous studies have been limited by either using MARS-MRI as the reference standard, or by assuming agreement between imaging modalities was acceptable [201, 229-231]. Further strengths include only reporting on HRA patients, and applying stricter inclusion criteria for both the acceptable time intervals between performing imaging on the same hip, and between imaging and revision surgery [201, 220, 230]. This study is also the first assessing the diagnostic test characteristics for combined ultrasound and MARS-MRI imaging for detecting intra-operative pseudotumours.

Limitations include the retrospective design therefore the findings are subject to bias for this study type. The relatively small number of revisions included resulted in broad

confidence intervals around the estimated specificities and NPVs (Table 5.7.), which in turn reduces the precision of the reported findings. The reported diagnostic test characteristics for all imaging combinations may have been overestimated given that the true prevalence of pseudotumour was high. Ultrasound is a dynamic investigation, therefore the retrospective review of ultrasound reports and images could have affected interpretation. However as recommended [229] all original imaging was performed by an experienced musculoskeletal radiologist, with all retrospective reviews performed by a different but similarly experienced and blinded musculoskeletal radiologist. Furthermore in all cases the reports and images categorically confirmed the presence or absence of a pseudotumour, which was the primary outcome of interest. Delays between performing imaging within the same patient, and delays between imaging and revision may both explain some of the discrepancies between imaging and intra-operative findings. The NOC follow-up protocol may have introduced selection bias and underestimated the diagnostic characteristics for ultrasound given this modality was used first-line with MARS-MRI reserved for problematic cases. Ideally all patients would have undergone both ultrasound and MARS-MRI on the same day regardless of the results.

Imaging and surgery were not independent, with the imaging results used by surgeons to recommend revision. Therefore surgeons were not blinded to the imaging findings, which could influence the documented intra-operative findings; however this is an unavoidable limitation of this study type also encountered elsewhere [220]. Potential inaccuracies in surgeon reporting at revision could also influence the results and may explain some of the discrepancies between imaging and intra-operative findings. Including the false positive cases, which may have reflected inaccuracies in surgeon

reporting, resulted in an underestimation of the diagnostic characteristics and agreement for all imaging combinations (Chapter 5.7.3.). However all revision surgeons at the NOC were aware of pseudotumours since 2007 (Chapter 3.6.1.), with the presence or absence of a pseudotumour always clearly stated in the operation records assessed. A strict definition for pseudotumour was used, which included communication with the joint and incorporated histopathological findings. Earlier definitions of pseudotumour were broader [16, 25], which should be considered when interpreting the results. It is suspected that the findings presented only apply to symptomatic patients likely to require revision surgery, and therefore may not apply to the many asymptomatic patients with well-functioning arthroplasties under surveillance. Finally, the study findings cannot be assumed to apply in MoM THA patients.

5.9. Patients and methods: new pseudotumour study

5.9.1. Study design

A prospective longitudinal cohort study was performed involving patients with initially asymptomatic HRAs that were all implanted at the NOC. All primary procedures were performed between 1st December 2000 and 21st December 2005. Information regarding ethical approval (Chapter 3.3.1.), the data source (Chapter 3.3.2.), the primary surgery performed (Chapter 3.3.3.), and patient follow-up (Chapter 3.3.4.) at the NOC have been detailed.

5.9.2. Description of initial cohort and assessment

Between 2007 and 2008, an ethically approved prospective cohort study was performed, which included 201 asymptomatic HRAs in 158 patients (mean age 56.0 years; 61%

male) [202]. This study was designed to determine the prevalence of radiological pseudotumour in asymptomatic HRA patients, given little was known about this clinical entity at the time. Therefore patients participating in the initial study were assessed at variable time points from their index arthroplasty, though all patients were assessed at a minimum of three years from primary surgery. Patients were considered asymptomatic when initially assessed because they all: (1) denied experiencing any hip related symptoms; (2) were satisfied with the surgical results; and (3) had good or excellent initial OHSs ($\geq$34) [170, 262].

All patients underwent the same initial assessment between 2007 and 2008. This assessment included hip radiographs (Chapter 5.3.4.), blood metal ions (Chapter 3.3.4.2.), and hip ultrasound (Chapters 3.3.4.1 and 5.6.3. and Appendix 10.). Patients also completed an OHS for each HRA (Chapter 3.3.4. and Appendix 6.), and the University of California, Los Angeles (UCLA) activity score questionnaire was completed on a per patient basis [303, 304]. The UCLA activity score ranged from 1 (wholly inactive) to 10 (regular participation in impact sports) (Appendix 15.) [303].

5.9.3. Description of cohort undergoing repeat assessment

5.9.3.1. Cohort assessed

Between 2012 and 2013 a subgroup of patients from the initial study underwent repeat hip ultrasound examination and completed a further OHS and UCLA score questionnaire. As per locally agreed protocol between the NOC arthroplasty surgeons it was not possible to also repeat hip radiographs and blood metal ions in all patients. There were 152 HRAs from the initial study who underwent repeat assessment. Twelve asymptomatic HRAs undergoing repeat assessment had ultrasounds demonstrating a

pseudotumour at initial assessment and had not undergone revision surgery (Appendix 16.). As the study objective was to identify risk factors associated with developing new pseudotumours, these 12 cases were excluded from the present analysis. The final study cohort undergoing repeat assessment consisted of 140 HRAs (112 patients). Repeat assessments were performed at a mean time of 4.4 years (range 3.2-5.0 years) from the initial assessment, and at a mean time of 8.3 years (range 6.2-12.4 years) from primary HRA (Table 5.8.).

Covariate	Whole cohort (n=140)	Hips with a new pseudotumour (n=15)	Hips without a new pseudotumour (n=125)	P-value
Gender				
Female	45 (32.1%)	4 (26.7%)	41 (32.8%)	0.774
Male	95 (67.9%)	11 (73.3%)	84 (67.2%)	
Age at initial ultrasound (in years)	Mean 60.6 (range 33.3-74.7)	Mean 64.3 (range 53.3-74.7)	Mean 60.2 (range 33.3-73.1)	0.067
MoM hip laterality				
Unilateral	84 (60.0%)	10 (66.7%)	74 (59.2%)	0.781
Bilateral	56 (40.0%)	5 (33.3%)	51 (40.8%)	
Implant design				
BHR	79 (56.4%)	5 (33.3%)	74 (59.2%)	0.083
Conserve Plus	55 (39.3%)	10 (66.7%)	45 (36.0%)	
Recap	6 (4.3%)	0 (0%)	6 (4.8%)	
Time between hip resurfacing arthroplasty and initial ultrasound (in years)	Mean 3.9 (range 3.0–7.4)	Mean 3.7 (range 3.0–7.4)	Mean 3.9 (range 3.0–7.0)	0.596
Time between repeat ultrasound (in years)	Mean 4.4 (range 3.2-5.0)	Mean 4.3 (range 3.5–5.0)	Mean 4.4 (range 3.2–5.0)	0.361
Acetabular inclination (in ^{o})	Mean 46.1 (range 21.3-65.5)	Mean 44.7 (range 27.8-57.0)	Mean 46.2 (range 21.3-65.5)	0.508
Acetabular anteversion (in ^{o})	Mean 14.9 (range 2.0-33.0)	Mean 15.0 (range 2.7-30.4)	Mean 14.9 (range 2.0-33.0)	0.958
Blood cobalt (in µg/l)	Median 2.1 (IQR 1.5-4.0)	Median 3.8 (IQR 2.6–10.6)	Median 2.1 (IQR 1.4–3.6)	**0.004**
Blood chromium (in µg/l)	Median 2.4 (IQR 1.3-4.9)	Median 3.9 (IQR 2.3–8.3)	Median 2.3 (IQR 1.3–4.4)	**0.020**
OHS				
Initial	Median 47.0 (IQR 45.0-48.0)	Median 46.0 (IQR 44.5-47.5)	Median 47.0 (IQR 45.0-48.0)	0.212
Repeat	Median 46.5 (IQR 43.0-48.0)	Median 44.0 (IQR 43.0-48.0)	Median 47.0 (IQR 43.0-48.0)	0.632
Change	Mean -1.0 (range -24-10)	Mean -2.4 (range -19-5)	Mean -0.9 (range -24-10)	0.510
UCLA score				
Initial	Mean 7.3 (range 3-10)	Mean 6.8 (range 4-8)	Mean 7.3 (range 3-10)	0.340
Repeat	Mean 7.3 (range 2-10)	Mean 6.8 (range 4-9)	Mean 7.3 (range 2-10)	0.388
Change	Mean 0.1 (range -4-5)	Mean 0.1 (range 0-1)	Mean 0.1 (range -4-5)	0.880

Table 5.8 Patient demographics for the cohort undergoing repeat assessment.
BHR = Birmingham Hip Resurfacing; MoM = metal-on-metal; OHS = Oxford hip score; UCLA = University of California, Los Angeles. P-values <0.05 in bold text.

5.9.3.2. Ultrasound assessment and classification

The same experienced musculoskeletal radiologist performed all ultrasound examinations for both the initial and repeat assessments. This radiologist was blinded to all relevant clinical information, including patient symptoms and blood metal ion concentrations. The ultrasound protocol and examination techniques used have been detailed (Chapter 5.6.3. and Appendix 10.).

The radiologist assigned each ultrasound scan to one of four grades previously described [209]: (1) normal, (2) bursa (psoas bursa, trochanteric bursa/thickening), (3) pathological effusion, and (4) pseudotumour. A small amount of intra-articular fluid was considered normal, but when the fluid depth exceeded 15 mm at the anterior joint line, this was classified as a pathological effusion. Simple fluid collections in the anatomic psoas or trochanteric bursa were classified as such. The same definition was used for diagnosing a pseudotumour on ultrasound as described (Chapter 5.6.4.): a cystic, solid, or mixed mass in continuity with the hip joint. When lesions were present the volume (product of the maximum recorded dimension in each of three orthogonal planes in centimetres), consistency (solid, cystic, or mixed) and location were recorded.

5.9.3.3. Outcomes

The study outcome of interest was the proportion of HRAs developing a new pseudotumour between repeat ultrasounds. By definition these new pseudotumours could arise from hips with an initial scan grade of 1 to 3.

5.9.4. Cohort not undergoing repeat assessment

There were 49 HRAs from the initial study that did not undergo repeat assessment. Reasons for this included death (four hips), undergoing revision to conventional THA (16 hips), and declined to participate or did not attend scheduled ultrasound appointments (29 hips). The demographics for these cases and the results of the initial assessments have been summarised (Appendix 16.).

Mean time from initial assessment to revision for the 16 revised HRAs was 2.5 years (range 0.1-4.4 years), and the mean time from primary to revision surgery was 6.6 years (range 3.1-9.4 years). Revision indications included symptomatic pseudotumour (n=14), dislocation (n=1), and femoral component loosening (n=1). The 14 revisions for symptomatic pseudotumour all had blood metal ion concentrations over 2 µg/l and abnormal ultrasound imaging at initial assessment. All 14 pseudotumours were confirmed both intra-operatively and histopathologically (Chapter 3.3.5.).

5.9.5. Statistical analysis

For paired analyses, the change in volume between initial and repeat ultrasound scans was assessed using a Wilcoxon signed-rank test. To assess risk factors associated with the development of a new pseudotumour, statistical tests were chosen to reflect the exposure variable and data distribution. These included two-sided unpaired t-tests (age, acetabular inclination and anteversion, time from primary to first scan, time between repeat scans, change in OHS, UCLA score, initial lesion volume), the Wilcoxon rank-sum test (OHS, blood cobalt and chromium concentrations), and the Chi-squared test or Fisher’s exact test (gender, implant design, bilateral HRAs, grade of first ultrasound

scan). The Chi-squared test was also used to assess the proportion of HRAs developing new pseudotumours by initial ultrasound scan grade.

Diagnostic test characteristics (Chapter 5.6.6.) were calculated for HRAs that did not develop new pseudotumours on ultrasound by the repeat assessment. Three clinical scenarios were assessed based on the results of the initial assessment: (1) normal ultrasound alone, (2) low blood metal ions alone, and (3) normal initial ultrasound and low blood metal ions. Low blood metal ions were defined as both a cobalt and chromium concentration of below 2 μg/l given that this has been considered a safe upper-limit for well-functioning MoMHAs without clinical concern [22]. The newly devised BHR blood metal ion thresholds (Chapter 4.) were not used given that these only apply to patients with BHR implants, and the present cohort comprised of patients with a range of HRA designs.

5.10. Results: new pseudotumour study

5.10.1. Changes between repeat imaging

There were 140 HRAs undergoing repeat assessment. Changes in ultrasound grade that occurred between initial and repeat assessments have been summarised (Table 5.9.). An increase in ultrasound grade between scans was seen in 17.9% (n=25) of hips, 78.6% (n=110) had no grade change, and 3.6% (n=5) had a decrease in grade. The mean change in lesion volume between scans for all 140 HRAs was an increase of 4.5 cm^3 by the repeat assessment scan (range -3.0 cm^3 to 392 cm^3), which was statistically significant (p=0.0003).

	Repeat ultrasound grade				
Initial ultrasound grade	1	2	3	4	**Total**
1	97 (69.3%)	6 (4.3%)	3 (2.1%)	4 (2.9%)	110 (78.6%)
2	3 (2.1%)	11 (7.9%)	1 (0.7%)	8 (5.7%)	23 (16.4%)
3	2 (1.4%)	0 (0%)	2 (1.4%)	3 (2.1%)	7 (5.0%)
4	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Total	102 (72.9%)	17 (12.1%)	6 (4.3%)	15 (10.7%)	140 (100%)

Table 5.9 Change in ultrasound grade between initial and repeat assessments.

5.10.2. Development of new pseudotumours

New pseudotumours developed in 10.7% (n=15) of the cohort (Table 5.8. and 5.9.). For hips with an initial ultrasound scan grade of 1 to 3, the risk of developing new pseudotumours within five years of the initial assessment were as follows: grade 1 = 3.6% (4/110); grade 2 = 34.8% (8/23); grade 3 = 42.9% (3/7). New pseudotumours were significantly more likely to develop from HRAs with initial scan grades of 2 (p=0.0001) and 3 (p=0.0033) compared with normal initial scans (grade 1). There was no significant difference in new pseudotumour development between HRAs with initial scan grades of 2 and 3 (p=0.367).

The median volume of new pseudotumours was 18.0 cm^3 (IQR 6.2 cm^3 to 39.0 cm^3). High blood cobalt concentration (p=0.004) and high blood chromium concentration (p=0.020) at initial assessment were the only factors significantly associated with developing new pseudotumours (Table 5.8.). Presently four hips with new pseudotumours identified have been revised to conventional THAs, with the remaining 11 hips under regular clinical surveillance.

5.10.3. Diagnostic test characteristics for clinical scenarios

Diagnostic test characteristics were assessed for three clinical scenarios based on the results of the initial assessment (Chapter 5.9.5.). These scenarios aimed to identify HRAs that did not develop new pseudotumours following repeat assessment (Table 5.10.). The clinically optimal diagnostic test characteristics for identifying HRAs that did not develop new pseudotumours were obtained for hips with normal initial ultrasound scans (grade 1) and low blood cobalt and chromium concentrations (<2 µg/l) (n=50; 35.7% of current cohort) (Appendix 17.). The diagnostic test characteristics for not developing new pseudotumours for this clinical scenario were 40.0% sensitivity, 100% specificity, 100% PPV, and 16.7% NPV. These particular characteristics were considered clinically optimal compared with using the initial ultrasound or blood metal ion results alone because it was possible to definitively predict that HRAs with both normal initial ultrasounds and low blood metal ions did not develop new pseudotumours within five years (100% PPV).

	Results of initial assessment		
	Normal ultrasound and blood metal ions less than 2 µg/l	**Normal ultrasound alone**	**Blood metal ions less than 2 µg/l alone**
Number of hips meeting initial assessment criteria (% of cohort)	50 (35.7)	110 (78.6)	59 (42.1)
Area under the curve (%)	70 (66-74)	79 (67-91)	66 (56-76)
Sensitivity (%)	40.0 (31.3-49.1)	84.8 (77.3-90.6)	45.6 (36.7-54.7)
Specificity (%)	100.0 (78.2-100.0)	73.3 (44.9-92.2)	86.7 (59.5-98.3)
Positive predictive value (%)	100.0 (92.9-100.0)	96.4 (91.0-99.0)	96.6 (88.3-99.6)
Negative predictive value (%)	16.7 (9.6-26.0)	36.7 (19.9-56.1)	16.0 (8.8-25.9)

Table 5.10 Diagnostic test characteristics for asymptomatic hip resurfacing arthroplasties not developing new pseudotumours following repeat ultrasound for three clinical scenarios based on the initial assessment.
95% confidence intervals for all diagnostic test characteristic values are provided in brackets.

5.11. Study strengths and limitations: new pseudotumour study

The study strengths include the prospective design, and the large number of asymptomatic HRA patients undergoing repeat investigation after an extended time period following a comprehensive baseline assessment. Previous studies performing serial imaging in HRA patients have involved small cohorts (4-53 hips) with short-term follow-up between repeat imaging (mean of 0.7-2.2 years), and also have not uniformly included asymptomatic patients only [109, 223, 224]. Therefore this study represents the largest reported cohort with the longest follow-up period between repeat assessments in asymptomatic HRA patients. A further strength is an experienced musculoskeletal radiologist blinded to all the clinical information performed and

assessed all ultrasound scans, which is recommended for achieving results comparable to MARS-MRI [229].

Limitations include not performing hip radiographs and blood metal ions at the repeat assessment. Although this was not feasible due to locally agreed hospital protocols, it is recognised that the study findings and subsequent recommendations may have changed or have been enhanced if these investigations had been repeated. Given that ultrasound is operator dependent less experienced radiologists may not have obtained the same results as presented here. However it is suspected that MARS-MRI could be substituted for ultrasound (Chapter 5.7.) in institutions without adequate experience to perform and interpret ultrasound imaging. The findings presented and subsequent recommendations only apply to asymptomatic HRA patients undergoing initial assessments within three to seven years of surgery and repeat assessments within three to five years of initial assessment. Further studies are needed with long-term follow-up to inform clinical decisions about subsequent patient surveillance. A strict definition for pseudotumour was used in this study, which must be considered when making comparisons to earlier reports with broader definitions (Chapter 5.8.). Ideally a clinical risk scoring tool would have been produced to quantify a patient's risk of developing new pseudotumours based on the results of the initial assessment, such as the one described earlier (Appendix 12.). However this was not possible due to the small number of HRAs developing new pseudotumours (n=15). It was also not possible to investigate the newly devised BHR blood metal ion thresholds (Chapter 4.) specifically in the BHR subgroup due to the small number of BHRs developing new pseudotumours (n=5). Finally, the study findings cannot be assumed to apply in MoM THA patients.

5.12. Discussion

5.12.1. The role of radiographs

Patient and radiographic factors provided useful information for distinguishing between HRAs with and without evidence of a pseudotumour. Predictors of pseudotumour included female gender, acetabular component malposition, acetabular and femoral osteolysis, acetabular loosening, and the absence of HO.

Female gender was predictive of pseudotumour, which confirmed the earlier analysis and has been discussed (Chapter 3.6.2.). Young age at HRA and non-BHR implants predicted pseudotumour revision in the earlier analysis (Chapter 3.4.), but not in the radiographic study. This may reflect the smaller radiographic cohort and the different analytical methods (logistic vs. Cox regression). However these two factors are important and should be included along with female gender when risk stratifying HRA patients for surveillance (Chapter 3.6.3.).

Similar to previous studies [20, 113, 150], high inclination and inadequate anteversion (<5°) predicted pseudotumour revision. Hard-on-hard bearings have low tolerances to acetabular component positioning outside an “optimal zone,” with malposition associated with edge loading, high wear, and early failure [150]. However the relationship between acetabular component position and pseudotumours is complex given patients with pseudotumours can have adequately positioned components [20, 150, 151], and well-functioning HRAs can have malpositioned components (Table 5.2.) [13, 70, 72, 73]. Indeed acetabular component malposition was the primary reason for abnormal radiographs in non-pseudotumour hips. When malposition was excluded only

19% of the non-pseudotumour group had abnormal radiographs, which is consistent with previous studies (up to 25% of non-revised HRAs) [13, 70, 72, 73]. Pseudotumour development is therefore multifactorial and not only dependent on acetabular component malposition. Hence it is important to assess radiographs for other signs suggestive of failure.

Osteolysis and acetabular component loosening were highly predictive of pseudotumours. Intra-operative osteolysis and acetabular loosening have been reported in up to 33% of pseudotumour revisions [29, 195], with extensive osteolysis often requiring complex reconstruction that may result in poor outcomes following revision [12, 247]. Surgeons must carefully inspect radiographs for subtle osteolysis or acetabular loosening and investigate further were necessary, as early pseudotumour identification may improve patient outcomes following revision.

The significance of femoral neck narrowing remains unclear. Up to 28% of well-functioning HRAs reportedly have greater than 10% narrowing, with narrowing stabilising within five years [13, 54, 70, 197]. Therefore neck narrowing may simply reflect early bone remodelling [70, 153]. However neck narrowing has also been reported in up to 26% of pseudotumour revisions [192, 195]. The present work observed similar degrees of neck narrowing in HRAs with and without pseudotumours, suggesting narrowing is not necessarily pathological. However if radiographic narrowing is observed it should be interpreted in the context of other abnormalities.

Males more commonly have HO [305], therefore the higher HO prevalence in the non-pseudotumour group is likely related to significantly more males having well-

functioning HRAs [13, 70, 132]. Other factors that may influence HO development (surgical approach and peri-operative drugs) were not assessed due to lack of medication data and the risk of overfitting the model. Patients with HO may have a lower pseudotumour risk because the associated stiffness and reduced hip motion may decrease the risk of edge loading and pseudotumour formation, however this requires further investigation.

5.12.2. The role of cross-sectional imaging

Both ultrasound and MARS-MRI have a role when assessing symptomatic HRA patients requiring revision surgery. However when establishing the presence or absence of an intra-operative pseudotumour there was a trend supporting the combination of ultrasound and MARS-MRI as the most effective option, followed by MARS-MRI alone, then ultrasound alone.

Combined imaging with ultrasound and MARS-MRI provided substantial agreement (κ=0.69) for detecting intra-operative pseudotumours. Combined imaging demonstrated 100% sensitivity and NPV, whilst maintaining the specificity (57.1%) and PPV (91.7%) compared with ultrasound or MARS-MRI alone. Therefore combined imaging correctly identified all hips with a pseudotumour. Furthermore hips without a pseudotumour on combined imaging were correctly confirmed as having no such lesions intra-operatively. The use of both ultrasound and MARS-MRI to assess patients has the added advantage of identifying non-pseudotumour pathologies, which can be the source of symptoms [199, 201]. Although combined imaging did not miss any pseudotumours in the cohort investigated, both ultrasound and MARS-MRI results taken in isolation missed pseudotumours which included large lesions (Table 5.6.). This provides support for

using combined imaging in HRA patients considered to require revision surgery when there may be an initial diagnostic dilemma.

One previous study attempted to validate ultrasound and MARS-MRI findings with the intra-operative findings in MoMHA patients imaged with both modalities prior to revision. Although good correlation was reported between both modalities and the intra-operative findings, only five patients were revised with delays of up to one-year between performing ultrasound and MARS-MRI in the same patient [201]. In two separate studies involving different patient cohorts, Lainiala *et al*. examined the ability of ultrasound [206] and MARS-MRI [220] to detect intra-operative pseudotumours in MoMHA patients undergoing revision. Better agreement was reported for ultrasound (κ=0.52-0.64) [206] than MARS-MRI (κ=0.40) [220]. This is in contrast to the present study findings where MARS-MRI alone (κ=0.54) provided better agreement than ultrasound alone (κ=0.36). These differences may be explained by all previously reported patients having a withdrawn MoMHA system, which included MoM THAs, and that each patient was not imaged with both modalities [206, 220].

The best diagnostic test characteristics for the ability of cross-sectional imaging to identify pseudotumours have been reported by Garbuz *et al*. [229]. In 40 asymptomatic MoM THAs not undergoing revision surgery ultrasound and MRI findings were compared, with consensus between modalities used as the reference standard. The authors reported 100% sensitivity, 96% specificity, 92% PPV, and 100% NPV for detecting pseudotumours with ultrasound, and 92% sensitivity, 100% specificity, 100% PPV, and 96% NPV for detecting pseudotumours with MRI [229]. Apart from a lower specificity, the present findings for combined imaging produced similar results to those

reported by Garbuz *et al*. for ultrasound and MRI used in isolation [229]. However it is important to consider the methodological differences between the studies, which include the lack of an intra-operative and histopathologically confirmed reference standard in the earlier study. Furthermore the work by Garbuz *et al*. focussed on asymptomatic MoM THA patients whilst the present study was based on symptomatic HRA patients requiring revision surgery. Garbuz *et al.* concluded that ultrasound should be used to screen MoMHA patients for pseudotumours [229]. Whilst this may be acceptable for asymptomatic patients, the present data suggests that symptomatic HRA patients with a normal ultrasound require a normal MARS-MRI examination to exclude a pseudotumour. The converse is also true. For institutions preferring to use MARS-MRI initially, a normal examination requires a subsequently normal ultrasound to exclude a pseudotumour.

Although reasons have been considered which could account for the disagreement observed in this study between the imaging and intra-operative findings, such as inaccurate surgeon reporting and delays in imaging (Chapters 5.7.3. and 5.8.), it is worth considering one further explanation. Numerous studies have demonstrated pseudotumour regression (complete disappearance or reduction in pseudotumour grade and/or volume) on repeat ultrasound or MARS-MRI in both symptomatic and asymptomatic MoMHA patients [109, 209, 223, 224]. Whilst the explanation for why pseudotumours regress is unclear, such regression may explain some of the false positive imaging findings observed.

5.12.3. The role of repeat cross-sectional imaging

New pseudotumours were detected on ultrasound in 11% of initially asymptomatic HRA patients when undergoing repeat assessment within five years. High blood cobalt and chromium concentrations at initial assessment were the only risk factors associated with developing new pseudotumours. No asymptomatic HRAs with normal initial ultrasound scans and low blood metal ion concentrations (<2 µg/l) went on to develop new pseudotumours within five years of initial assessment.

Serial ultrasound imaging was useful for identifying the development and progression of abnormalities around HRAs, including pseudotumours. These observations support findings from smaller short-term longitudinal studies [109, 209, 229, 232]. Most asymptomatic HRA patients did not develop new pseudotumours within five years of initial assessment in the present cohort. Furthermore it was possible to reliably identify this subgroup based on the initial ultrasound and blood metal ion results.

To definitively exclude the development of new pseudotumours in HRA patients both a normal initial ultrasound and low blood metal ions (<2 µg/l) were required at initial assessment. Using the ultrasound or blood metal ion findings in isolation were less effective for excluding the development of new pseudotumours. These new pseudotumours were more likely to develop from HRAs with initial ultrasound scan grades of 2 and 3 compared with grade 1 scans. However a small proportion of hips (3.6%; 4/110) with normal initial ultrasound scans did develop new pseudotumours within five years of initial assessment if the blood metal ion results were not considered. By inference these patients with normal initial ultrasound scans all had blood metal ion concentrations above 2 µg/l at initial assessment. Similarly a small proportion of hips

(3.4%; 2/59) with blood cobalt and chromium concentrations below 2 µg/l at initial assessment did develop new pseudotumours within five years if the initial ultrasound results were not considered. Therefore on the basis of the data presented it is recommended that all asymptomatic HRA patients undergo a comprehensive baseline assessment, including both cross-sectional imaging and blood metal ions.

5.12.4. Clinical implications of findings

Current worldwide follow-up recommendations for HRA patients are not evidence-based and are very costly (Chapter 2.5.). This chapter has described three clinical studies, which have assessed the roles of hip imaging in the surveillance of HRA patients. Similar to earlier work (Chapter 4.6.3.) and observations by others [112] it is not possible to only rely on a single investigative modality when assessing HRA patients.

Pelvic radiographs form a useful part of the baseline assessment in HRA patients, though it is recognised that the MHRA currently do not specify a role for radiographs during surveillance [23]. However radiographs should not be considered a substitute for performing other investigations, given that 20% of HRAs revised for pseudotumour had normal radiographs. Therefore radiographs represent a useful adjunct to other investigative modalities. Similarly symptomatic patients who may eventually need revision surgery with a normal initial ultrasound or MARS-MRI should subsequently undergo combined cross-sectional imaging to definitively exclude the presence of a pseudotumour.

It is recommended that all HRA patients undergo a complete clinical assessment. In addition to blood metal ions (Chapter 4.), this assessment should include pelvic radiographs and cross-sectional imaging using either ultrasound or MARS-MRI. Subsequent management will depend on a combination of the results of this complete clinical assessment, hip symptoms, and the patient's pseudotumour risk (Chapter 3.6.3.). Although patient management following initial assessment will be considered further in the concluding chapter (Chapter 7.3.), it can broadly be grouped as follows:

- **Symptomatic patients with abnormal investigations**

Recommend revision surgery, or remain under close surveillance with further investigation (including combined cross-sectional imaging) within six months to one year of initial assessment. Follow-up regularity will depend on the nature of the abnormalities and the patient's pseudotumour risk.

- **Symptomatic patients with normal investigations**

Further investigations recommended including combined cross-sectional imaging, hip aspiration, and assessing for other causes of symptoms such as spinal pathology. If no cause is identified for the symptoms then these patients should undergo repeat clinical assessment within one to two years of initial assessment. Follow-up regularity will depend on the patient's pseudotumour risk.

- **Asymptomatic patients with abnormal investigations**

Remain under surveillance and undergo repeat clinical assessment (consider combined cross-sectional imaging) within three years of initial assessment. Follow-up regularity

will depend on the nature of the abnormalities and the patient's pseudotumour risk. In certain circumstances consider recommending revision surgery even if asymptomatic.

- **Asymptomatic patients with normal investigations**

If cross-sectional imaging is normal and blood metal ion concentrations are below 2 µg/l patients do not need repeat assessment within five years of initial assessment. Subsequent follow-up regularity will depend on the patient's pseudotumour risk, ranging from repeat assessment at five years in young females to potentially no further follow-up for male BHR patients whilst asymptomatic.

Given the significant costs associated with HRA surveillance (Chapter 2.5.) the financial implications of the proposed recommendations must be considered. Although all patients should undergo either ultrasound or MARS-MRI, the specific modality selected by each institution should consider the availability of ultrasound expertise and any local financial constraints. Due to the significant costs associated with combined cross-sectional imaging it is recommended that this option should be reserved for more complex cases with a diagnostic dilemma. Such instances include symptomatic patients with normal initial ultrasounds or MARS-MRIs, or even asymptomatic patients with normal initial cross-sectional imaging but with raised blood metal ion concentrations. On the basis of the data presented these recommendations only apply to the first year following initial imaging, after which time imaging may require repeating in symptomatic patients.

The lack of consensus between worldwide regulatory authorities regarding surveillance for the large group of asymptomatic HRA patients (Chapters 2.5. and 5.1.) makes the findings from the prospective longitudinal study both clinically and financially significant when considering the need for future follow-up. The present work suggests 36% of asymptomatic HRA patients do not require repeat assessment within five years of initial assessment. The European regulatory authority guidance for asymptomatic HRA patients is very resource intensive (Chapter 2.5.) and now appears unnecessary. The annual follow-up costs for asymptomatic HRAs recorded in the NJR are estimated to be nearly £8.3 million (Chapter 2.5.3.) when applying European guidance [22]. Excluding 36% of patients from surveillance for five years would result in huge savings of nearly £15 million over this period. As not all 201 asymptomatic HRAs initially assessed underwent repeat assessment in this longitudinal study (Chapter 5.9.4.) a more conservative saving estimate can be obtained be assuming all patients not undergoing repeat assessment would have had abnormal repeat investigations. In this instance 25% (50/201) of asymptomatic HRAs could still be released from surveillance for five years which would again produce significant savings of over £10.3 million. These savings could subsequently be concentrated towards patients requiring further investigations and/or revision surgery.

A final consideration relates to balancing the potential financial savings with the implications of missing pseudotumours in initially asymptomatic HRA patients. Although 36% could definitively be excluded from surveillance for five years based on the combined ultrasound and blood metal ion results with no risk of new pseudotumour development, a normal ultrasound alone could exclude many more patients (79%). Using imaging alone would more than double the potential savings when excluding

patients from surveillance for five years, with further savings achieved by not needing to perform an initial blood test (saving £30 per patient). However these financial savings need to be carefully balanced against the small risk of missing new pseudotumours within five years (3.6%).

The clinical implications of missing new pseudotumours are somewhat controversial. If patients with missed pseudotumours subsequently developed symptoms due to an aggressive and locally destructive lesion, which could result in a poor outcome following revision surgery [12, 28], then this would be considered clinically significant. In such cases a comprehensive initial assessment with imaging and blood metal ions would be preferable. However the position is less clear in patients who develop asymptomatic pseudotumours given that the natural history of these lesions remains unclear, with some studies reporting regression following repeat imaging [109, 110, 209, 223, 224]. Missing such pseudotumours may not be clinically significant, however further large longitudinal studies with extended follow-up are needed to make more definitive recommendations. Therefore on balance asymptomatic HRA patients require cross-sectional imaging and blood metal ions at initial assessment, rather than imaging alone, until more is known about the natural history of radiologically diagnosed pseudotumours.

5.13. Summary

- A combination of patient and radiographic factors provided useful information for distinguishing between HRAs with and without evidence of a pseudotumour. Factors predictive of HRAs with pseudotumours included female gender,

acetabular component malposition, acetabular osteolysis, femoral osteolysis, acetabular loosening, and the absence of HO.

- Baseline radiographs should be recommended for all HRA patients in the follow-up guidance issued by worldwide regulatory authorities. Currently the MHRA do not specify a role for radiographs.

- However radiographs should not be considered a substitute for performing blood metal ions and cross-sectional imaging, given that 20% of HRAs revised for histopathologically confirmed pseudotumours had normal radiographs.

- Both ultrasound and MARS-MRI have a role when assessing HRA patients requiring revision surgery. However there was a trend supporting the combination of ultrasound and MARS-MRI as the most effective option, followed by MARS-MRI alone, then ultrasound alone.

- The choice of cross-sectional imaging modality for investigating HRA patients will depend on local financial constraints and the availability of ultrasound expertise. However combined imaging should be considered in more complex cases, such as symptomatic patients with a normal initial ultrasound or MARS-MRI.

- Combined ultrasound and MARS-MRI imaging was effective at identifying intra-operative pseudotumours if present. Furthermore if neither imaging modality detected a pseudotumour, one was not found intra-operatively.

- New pseudotumours developed in 11% of asymptomatic HRAs undergoing repeat assessment within five years of initial assessment. High blood cobalt and chromium concentrations at initial assessment were the only risk factors associated with developing new pseudotumours.

- No asymptomatic HRAs with normal initial ultrasound scans and low blood metal ion concentrations (<2 µg/l) went on to develop new pseudotumours (36% of the current cohort) within five years of initial assessment. This patient subgroup does not require repeat follow-up during this time period, which will result in considerable financial savings.

- Asymptomatic HRA patients with either normal initial ultrasound scans alone, or low blood metal ion concentrations alone still have a small risk of developing new pseudotumours within five years of initial assessment. As this risk could be clinically significant it is advised that such patients do undergo some form of surveillance during this time period.

- Now that the specific roles for blood metal ions and the various hip imaging modalities have been defined for HRA patient surveillance, it is important to consider the expected outcomes following revision surgery. The subsequent chapter uses NJR data to establish the outcomes following MoMHA revision surgery, and to identify prognostic factors of outcome.

6. MANAGEMENT – REVISION SURGERY*

6.1. Introduction

Despite the high failure rates of primary MoMHAs [9, 10], it is estimated that at least 80% of these implants remain in-situ worldwide [2, 11]. Given the prevalence of ARMD revision surgery is increasing [2, 11, 203] (Chapter 3.) it is expected that many more MoMHA patients will undergo future revision surgery. However there is little good quality evidence available regarding the outcomes following MoMHA revision surgery performed for ARMD [28] (Chapter 2.4.), which represents the commonest cause of revision (Chapter 3.). This makes it difficult for surgeons to informatively counsel patients about the risks of undergoing further procedures.

Early observations suggested short-term outcomes following ARMD revision surgery were poor, and inferior compared with both MoMHA revisions for non-ARMD indications (fracture, loosening, infection), and matched patients undergoing primary conventional THA [12]. The poor outcomes were thought to relate to the invasive and destructive nature of ARMD [12, 14, 16, 48]. These initial observations [12, 243] led regulatory authorities and orthopaedic surgeons to widely recommend performing early revision in MoMHAs with ARMD [12, 23, 24, 29, 48], as it was thought that this strategy would improve subsequent outcomes. Over a period of time surgeons subsequently adopted a lower threshold for performing revision for ARMD. However

* Work presented in this chapter has been published in:

Matharu GS, Judge A, Pandit HG, Murray DW. Surgical factors influence failure rates following metal-on-metal hip arthroplasty revision surgery performed for adverse reactions to metal debris: analysis of 2,535 revisions from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. *Bone Joint J* (In Press: accepted 21st February 2017).

the impact of performing early surgery for ARMD is not known. Although it may prevent further bone and soft-tissue destruction, the surgery can be extensive [12, 203] and may be associated with increased morbidity and mortality. There is also a concern that widely recommending early revision surgery may have resulted in many unnecessary operations, which are not without risk.

Furthermore surgeons currently have no robust information regarding the type of revision surgery to perform in MoMHA patients with ARMD, apart from exchanging the bearing to a non-MoM articulation [195, 203]. Knowledge of any prognostic factors of outcome following ARMD revision would assist surgeons when making decisions about the reconstruction to perform. Some studies have assessed factors predictive of poor outcomes following ARMD revision, but these were small and underpowered (Chapter 2.4.3.4.). Therefore large cohort studies are needed to address these important clinical issues.

6.2. Objectives

The objectives of this chapter were:

1. To determine the outcomes following MoMHA revision surgery performed for ARMD, and to identify pre-operative and intra-operative factors predictive of re-revision surgery.
2. To compare the outcomes following MoMHA revision surgery performed for ARMD with the outcomes following MoMHA revision surgery performed for non-ARMD indications.

These two objectives were investigated in two separate clinical studies using data from the NJR, with both studies subsequently described in this chapter.

6.3. Patients and methods: ARMD outcome study

6.3.1. Study design

A retrospective observational cohort study was performed using data from the NJR. The subset of data used for the present study included all primary MoMHAs subsequently undergoing revision surgery for ARMD. A formal application was submitted to the NJR Research Sub-committee (RSC) to access the relevant data and perform the work described within this chapter. This application was approved on 2nd January 2015 (Appendix 18.).

6.3.2. Data source

6.3.2.1. Background

The NJR for England and Wales was set up in 2002 by the Department of Health and the Welsh Government. The NJR commenced collecting data on primary and revision hip and knee arthroplasties performed at both National Health Service (NHS) and private hospitals in England and Wales from 1st April 2003. Data on ankle (since April 2010), and elbow and shoulder (since April 2012) arthroplasties have been collected more recently. Procedures performed in Northern Ireland (since February 2013) and the Isle of Man (since July 2015) have also been collected.

The purpose of the NJR is described as being "To collect high quality and relevant data about joint replacement surgery in order to provide an early warning of issues relating to patient safety. In a continuous drive to improve the quality of outcomes and ensure the quality and cost effectiveness of joint replacement surgery, the NJR will monitor

and report on outcomes, and support and enable related research" [306]. The NJR is the world's largest arthroplasty registry, containing details of over two million joint replacement procedures. Of these, 800,683 primary THAs and 89,023 revision THAs have been recorded [2].

Data from the NJR is communicated in four main ways: (1) publication of an annual report available to all individuals, (2) individual surgeon reports containing data on arthroplasty activity and outcomes for the cases performed, (3) manufacturer reports detailing the outcomes of their respective implants, and (4) research requests which are approved by the NJR RSC.

6.3.2.2. Data collection

Voluntary and informed signed patient consent is required for arthroplasty procedures to be included in the NJR with patient identifiable information (name, date of birth, postcode, and NHS number) (Appendix 19.). This consent is usually requested pre-operatively. If consent is not provided only the anonymous surgical data can be included in the NJR, and in these cases linkage to any further procedures is not possible (detailed below). The latest annual report states 93% of all arthroplasty patients provided consent for their personal details to be recorded within the NJR [2].

Standardised data capture forms are completed by the operating surgeon at the time of any primary (Appendix 8.) and revision (Appendix 9.) hip arthroplasty procedure. These forms are completed by hand and broadly collect data on patient demographics at the time of surgery and the procedure performed. Surgeons may choose one or more indications when performing revision procedures. Component labels supplied by the

surgical team allow detailed implant data to be recorded in the NJR. Data from the completed forms are subsequently entered onto the electronic NJR data entry system via a portal by trained staff at each hospital. The unique patient identifiable information from individuals providing consent allows linkage of primary arthroplasty procedures to any subsequent revision procedures in which components were removed or exchanged. The NJR currently reports a linkability of 94.5% [2].

The NJR data capture forms contain a number of data fields, referred to as the minimum dataset (MDS). Over time the MDS has been modified. Therefore some data available on the most recent forms may not have been collected at all previously, or the data may have been collected previously but grouped in a different manner. Both of these issues have implications for data analysis when assessing a cohort where procedures have been captured using more than one MDS form. Most major changes to the MDS occurred before 2008 during the early phase of the NJR. Therefore these changes would not influence the revision procedures assessed in the present chapter given the dates of the data extract provided (Chapter 6.3.2.3.).

The most important and relevant modification to the MDS for the present work is the formal addition of adverse soft tissue reaction to particulate debris as a revision indication in June 2008 (Appendix 9.). This term selected by the NJR is considered synonymous with ARMD (Chapter 2.2.1.), therefore revisions recorded in the NJR as performed for adverse soft tissue reaction to particulate debris have subsequently been referred to in this thesis as revisions performed for ARMD. Prior to June 2008, ARMD revisions were either: (1) diagnosed and coded using the “Other” option with or without

further details provided as free text, or (2) undiagnosed or missed, with the revision coded using non-ARMD indications.

6.3.2.3. Data extract

An extract of anonymised patient data was received from the NJR on 14th August 2015. This dataset contained details of all primary MoMHAs recorded in the NJR that subsequently underwent revision surgery for any indication between 1st June 2008 and 14th August 2014 (n=5,867). The former date represents when ARMD was introduced on the NJR data capture forms as a revision indication for surgeons to select. The latter date was chosen to provide a minimum 1-year follow-up period for determining outcomes following revision surgery. Prior to obtaining the dataset the entire NJR database was linked with the Office for National Statistics (ONS) database using unique patient identifiers. The ONS provides data on all-cause patient mortality. Therefore the extract included both linked re-revision and mortality data.

The dataset received from the NJR included procedures performed in England and Wales only, as there are currently no patient tracing services in Northern Ireland and the Isle of Man. Furthermore the data extract did not include MoMHA revision procedures where the primary arthroplasty had not been recorded in the NJR.

6.3.2.4. Data cleaning

The data extract, which included 5,867 MoMHA revision procedures, was loaded into Stata Version 13.1 and subsequently coded and cleaned. Each numerical and categorical variable was systematically examined to assess for missing data and anomalies, such as

implausible values. All but three variables contained complete data with plausible values. The three variables with missing data were revision body mass index (BMI), revision femoral head size, and revision bearing surface. Revision BMI contained a large proportion of missing data in addition to implausible values, and therefore required significant cleaning prior to analysis. The other two variables had much smaller proportions of missing data. The proportion of missing data for each variable has been detailed subsequently as this varied between the specific cohorts assessed (Table 6.1. and 6.4.). As part of sensitivity analyses missing data was also addressed using multiple imputation techniques as described (Chapters 6.3.6.1. and 6.6.5).

All 5,867 procedures had been assigned a revision BMI value. These values ranged from 0 kg/m^2 to 395 kg/m^2 with their distribution illustrated (Figure 6.1.). In 2,040 hips (34.8%) the assigned value was 0 kg/m^2, with these values considered to represent missing data. Revision BMI values of below 16 kg/m^2 or above 60 kg/m^2 were considered improbable (17 hips; 0.3%) and therefore were subsequently excluded. Following data cleaning there were 3,810 hips (64.9% of whole cohort) with complete revision BMI data available and plausible values (range 17 kg/m^2 to 60 kg/m^2). The distribution of these revision BMI values following cleaning confirmed that this data was normally distributed (Figure 6.2.).

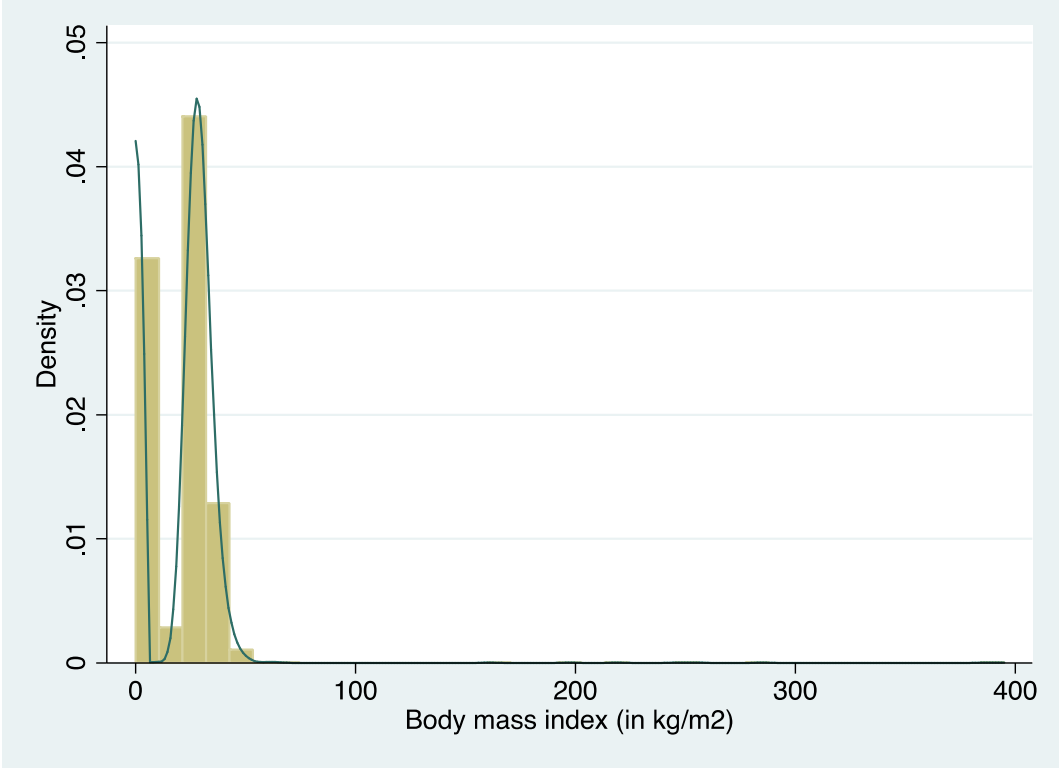


Figure 6.1 Histogram of revision body mass index values before data cleaning. Smoothed kernel density distribution estimate illustrated.

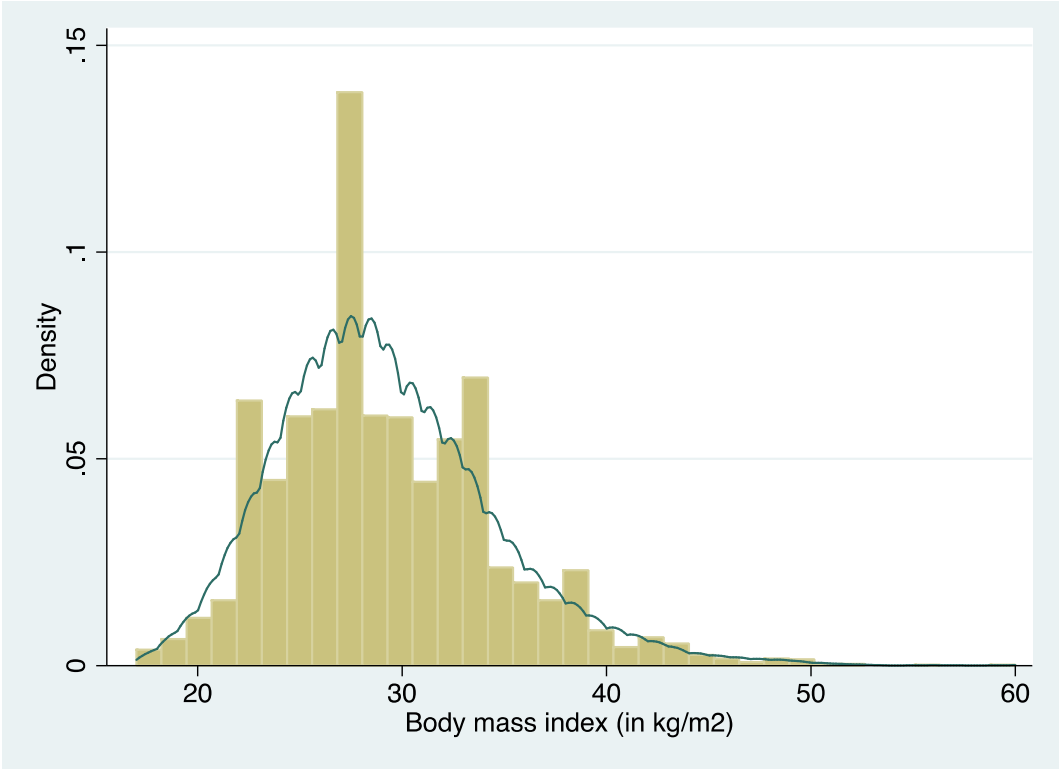


Figure 6.2 Histogram of revision body mass index values after data cleaning. Smoothed kernel density distribution estimate illustrated.

Following data cleaning the distributions for all numerical variables were assessed graphically for normality, with the data for each variable subsequently summarised based on the respective distribution. Binary categorical variables, such as gender, were maintained as such. Multi-level categorical variables were made binary where possible to assist data processing and analysis. For example there are five different options to select for surgical approach used at revision on the latest NJR data capture form (Appendix 9.). As the majority of hip revisions are performed using the posterior approach, this multi-level categorical variable was converted to a binary variable (i.e. posterior versus other). The same is true for revision surgeon grade, which was converted to consultant versus other. The remaining multi-level categorical variables were grouped to retain three or four levels because the additional levels were considered clinically important to differentiate from each other. Examples of such variables included American Society of Anesthesiologists (ASA) grade, chemical venous thromboembolism prophylaxis, revision procedure performed, and revision bearing surface (Table 6.1.).

6.3.3. Study inclusion and exclusion criteria

In the NJR data extract of 5,867 MoMHA revisions (Chapter 6.3.2.3.) there were 2,566 MoMHA revision procedures (43.7% of whole cohort) with ARMD recorded as either the only revision indication or one of multiple revision indications. These cases were initially eligible for study inclusion.

Hips were subsequently excluded for the following reasons: (1) it was not possible to confirm that the primary arthroplasty had a MoM bearing surface (three hips excluded); (2) the ARMD revision was recorded as a two-stage procedure with the excision

arthroplasty (first stage) captured in the NJR, but the re-implantation procedure (second stage) was absent (18 hips excluded); and (3) a two-stage revision was performed with both stages captured within the NJR, where the first stage was performed for ARMD but the second stage was performed for non-ARMD indications such as infection (ten hips excluded). The latter criterion was considered appropriate given that surgeons performing the second stage would have the results available from the histopathological and microbiological analyses performed on samples excised at the first stage procedure. Such results would allow a more accurate revision indication to be recorded at the second stage procedure. The final study cohort included 2,535 primary MoMHAs undergoing revision surgery for ARMD.

6.3.4. Exposure variables

For all procedures the NJR collects data on patient demographics (age, gender, BMI, ASA grade) and the arthroplasty procedure performed (date, indication, venous thromboembolism prophylaxis, surgeon grade, approach, components implanted including bearing surface, size, fixation, and the use of bone graft). Access to manufacturer data for the components removed and implanted was not authorised for this study, therefore this information was not available for subsequent analyses. In addition to ARMD, the other indications that could be recorded by the operating surgeon for revision procedures included aseptic loosening, infection, peri-prosthetic fracture, lysis, dislocation/subluxation, implant malalignment, implant mismatch (head/socket), acetabular component wear, implant fracture, liner dissociation, unexplained pain, and other.

The patient and surgical factors related to the ARMD revision procedure were used as covariates when assessing prognostic factors of outcome. The revision procedure performed was grouped into four to assist analysis: (1) all component revisions; (2) cup only revisions (with or without femoral head, liner, taper adapter revision), (3) stem only revisions (with or without femoral head, liner, taper adapter revision), and (4) modular component only revisions (both the cup and stem were retained with revision of the femoral head and liner, with or without the use of a taper adapter). Conventional methods for describing bearing surfaces were used, for example metal-on-polyethylene represents a metal femoral head articulating with a polyethylene liner or socket. The three major revision articulations (metal-on-polyethylene; ceramic-on-polyethylene; ceramic-on-ceramic) were grouped separately, with less frequently used revision bearing surfaces, such as MoM and ceramic-on-metal, combined in a fourth category.

6.3.5. Outcomes

The three study outcomes of interest were: (1) intra-operative complications during the MoMHA revision procedure; (2) all-cause mortality following revision surgery, and (3) all-cause re-revision surgery following revision.

Intra-operative complications were recorded at the time of the revision procedure by the operating surgeon (Appendix 9.). These complications included calcar crack, pelvic and/or femoral shaft penetration, trochanteric and/or femoral shaft fracture, and other complications. Mortality and re-revision surgery were linked to the MoMHA revision procedural data using unique patient identifiable information prior to receiving the data extract from the NJR (Chapter 6.3.2.3.). All-cause re-revision surgery was defined as any subsequent procedure in which hip arthroplasty components were removed or

exchanged for any indication following the MoMHA revision surgical procedure.

6.3.6. Statistical analysis

Differences in the patient and surgical factors between re-revised and non re-revised hips were assessed using either two-sided unpaired t-tests or the Wilcoxon rank-sum test (numerical data), and either the Chi-squared test or Fisher's exact test (categorical data). Cumulative patient and implant survival rates following ARMD revision surgery were determined using the Kaplan-Meier method. The endpoint for implant survival was re-revision surgery. Patients not undergoing re-revision surgery who remained alive were censored on the study end date (14th August 2015).

Fine and Gray competing-risk regression modelling was used to identify predictors of re-revision surgery. Mortality can be considered a competing-risk for re-revision surgery, therefore this method was considered appropriate when examining predictors of re-revision. Univariable models explored the association between each patient and surgical predictor, and re-revision surgery. Linearity of continuous predictors with outcome (re-revision) was assessed using fractional polynomial regression modelling, with data grouped if affects were non-linear. The proportional sub-hazards assumption was examined and satisfied for all predictors. The final multivariable competing-risk regression model was developed using stepwise selection methods. The p-values for removing and including predictors in the final multivariable model were $p>0.20$ and $p<0.10$ respectively. The discriminatory ability of the final multivariable model for distinguishing between hips undergoing and not undergoing re-revision surgery was analysed using the concordance (*c*) statistic. The *c* statistic (range 0-1; useful prognostic

models=0.60-0.85) provides a global assessment of fit for the survival model, and is equivalent to the area under a ROC curve [300, 307].

6.3.6.1. Sensitivity analysis: missing data

There were four covariates with missing data (revision BMI=682; revision procedure=5; revision femoral head size=29; revision bearing surface=119). Excluding all cases with missing data would substantially reduce the cohort size, which would subsequently result in a significant loss of statistical power and precision. In addition excluding cases with missing data could potentially introduce selection bias, given that the characteristics of patients with complete data may differ from those patients without complete data [300, 308]. To overcome this bias multiple imputation methods were used, which represents an accepted statistical process for managing missing data [309].

Multiple imputation allows for the uncertainty regarding missing data by creating several plausible imputed datasets and appropriately combining the results. As a sensitivity analysis, the competing-risk regression analyses were repeated using a complete dataset where all missing data values were imputed. An overall estimate of the regression model was calculated from the 50 imputed datasets, which was produced by averaging the results obtained from each dataset (Appendix 20.) [310, 311].

6.4. Results: ARMD outcome study

6.4.1. Intra-operative complications

There were 2,535 MoMHAs revised for ARMD that were included in this study (Table 6.1.). Intra-operative complications occurred in 1.6% (n=40) of these cases. Calcar cracks (27.5%; n=11) and greater trochanteric fractures (20.0%; n=8) were the commonest intra-operative complications.

6.4.2. Mortality

Overall mortality following revision surgery was 1.6% (n=41). Mean time from ARMD revision surgery to death was 1.9 years (range 0.1-5.9 years), with no hips undergoing re-revision before death. The cumulative patient survival rates following ARMD revision surgery were 99.4% (95% CI=99.1-99.7%) at 1 year, and 95.9% (95% CI=92.3%-97.8%) at five years.

Covariate	All ARMD hip revisions (n=2,535) (100%)	ARMD hip revisions not undergoing re-revision surgery (n=2,343) (92.4%)	ARMD hip revisions undergoing re-revision surgery (n=192) (7.6%)	p-value
Gender				
Female *vs. male*	1,509 (59.5)	1,388 (59.2)	121 (63.0)	0.305
Age at revision (years)				
Mean (SD)	63.6 (10.1)	63.7 (10.0)	62.3 (11.0)	0.1045
Revision BMI (kg/m^2)*	29.0 (5.1)	28.9 (5.1)	30.1 (5.7)	**0.0179**
Mean (SD)				
Bilateral revisions for ARMD	224 (8.8)	206 (8.8)	18 (9.4)	0.784
Primary arthroplasty				
MoM THA	1,716 (67.7)	1,572 (67.1)	144 (75.0)	**0.024**
Hip resurfacing	819 (32.3)	771 (32.9)	48 (25.0)	
Primary femoral head size (mm)				
≤36	470 (18.5)	429 (18.3)	41 (21.4)	0.737
38-44	751 (29.6)	696 (29.7)	55 (28.7)	
46-50	1,102 (43.5)	1,020 (43.5)	82 (42.7)	
≥52	212 (8.4)	198 (8.5)	14 (7.3)	
Time from primary to revision (years)				
Mean (SD)	5.2 (1.8)	5.3 (1.8)	4.7 (1.9)	**<0.0001**
Revision ASA grade				
1	476 (18.8)	438 (18.7)	38 (19.8)	0.769
2	1,827 (72.1)	1,688 (72.0)	139 (72.4)	
3 or above	232 (9.2)	217 (9.3)	15 (7.8)	
VTE – chemical				
LMWH (+/-other)	1,309 (51.6)	1,219 (52.0)	90 (46.9)	0.167
Aspirin only	68 (2.7)	62 (2.7)	6 (3.1)	
Other	922 (36.4)	852 (36.4)	70 (36.5)	
None	236 (9.3)	210 (9.0)	26 (13.5)	
VTE – mechanical				
Any *vs. none*	2,447 (96.5)	2,263 (96.6)	184 (95.8)	0.584
Number of ARMD hip revision procedures performed by each of 237 centres				
Median (range)	28 (1-169)	28 (1-169)	34 (1-169)	0.0532
Revision surgeon grade				
Consultant *vs. other*	2,430 (95.9)	2,247 (95.9)	181 (95.3)	0.693
Surgical approach				
Posterior *vs. other*	2,084 (82.2)	1,929 (82.3)	155 (80.7)	0.577

CONTINUED /...

Covariate	All ARMD hip revisions (n=2,535) (100%)	ARMD hip revisions not undergoing re-revision surgery (n=2,343) (92.4%)	ARMD hip revisions undergoing re-revision surgery (n=192) (7.6%)	p-value
Revision indications / intra-operative findings				
ARMD only	1,604 (63.3)	1,474 (62.9)	130 (67.7)	0.185
2 to 6 indications	931 (36.7)	869 (37.1)	62 (32.3)	
ARMD	2,535 (100)	2,343 (100)	192 (100)	NA
Pain	554 (21.9)	522 (22.3)	32 (16.7)	0.070
Aseptic loosening (any)	219 (8.6)	206 (8.8)	13 (6.8)	0.338
Acetabular	125 (4.9)	118 (5.0)	7 (3.7)	0.392
Femoral	112 (4.4)	105 (4.5)	7 (3.7)	0.588
Osteolysis (any)	177 (6.9)	166 (7.1)	11 (5.7)	0.479
Acetabular	110 (4.3)	104 (4.4)	6 (3.1)	0.390
Femoral	101 (4.0)	95 (4.1)	6 (3.1)	0.527
Other abnormality	85 (3.4)	81 (3.5)	**	0.405
Implant malalignment	79 (3.1)	76 (3.2)	**	0.278
Acetabular component wear	43 (1.7)	40 (1.7)	**	1.00
Peri-prosthetic fracture	43 (1.7)	39 (1.7)	**	0.563
Dislocation / subluxation	37 (1.5)	29 (1.2)	8 (4.2)	**0.005**
Infection	22 (0.9)	19 (0.8)	**	0.230
Head / socket mismatch	9 (0.4)	8 (0.3)	**	0.508
Liner dissociation	9 (0.4)	9 (0.4)	0 (0.0)	1.00
Implant fracture	6 (0.2)	6 (0.3)	0 (0.0)	1.00
Revision procedure*				
All components	1,032 (40.8)	968 (41.4)	64 (33.5)	**0.031**
Cup only	1,163 (46.0)	1,071 (45.8)	92 (48.2)	
Stem only	78 (3.1)	73 (3.1)	**	
Modular component only	257 (10.2)	227 (9.7)	30 (15.7)	
Revision femoral head size (mm)*				
<36	882 (35.3)	820 (35.5)	62 (33.0)	0.751
36	1,486 (59.5)	1,371 (59.4)	115 (61.2)	
>36	130 (5.2)	119 (5.2)	11 (5.9)	
Bearing surface*				
CoP	1,104 (45.7)	1,040 (46.5)	64 (35.4)	**0.017**
CoC	765 (31.7)	693 (31.0)	72 (39.8)	
MoP	533 (22.1)	488 (21.8)	45 (24.9)	
CoM, MoM or MoC	14 (0.6)	14 (0.6)	0 (0.0)	
Revision cup fixation				
Uncemented	1,919 (87.5)	1,780 (87.3)	139 (89.1)	0.522
Cemented	275 (12.5)	258 (12.7)	17 (10.9)	
Revision stem fixation				
Uncemented	723 (65.3)	674 (64.9)	49 (71.0)	0.299
Cemented	385 (34.8)	365 (35.1)	20 (29.0)	
Bone graft (femoral)	76 (3.0)	71 (3.0)	**	0.739
Bone graft (acetabular)	509 (20.1)	457 (19.5)	52 (27.1)	**0.012**

Table 6.1 Patient and surgical factors for all metal-on-metal hip arthroplasties revised for adverse reactions to metal debris.
ARMD = adverse reactions to metal debris; ASA = American Society of Anesthesiologists; BMI = body mass index; CoC = ceramic-on-ceramic; CoM = ceramic-on-metal; CoP = ceramic-on-polyethylene; LMWH = low molecular weight heparin; MoC = metal-on-ceramic; MoM = metal-on-metal; MoP = metal-on-polyethylene; NA = not applicable; SD = standard deviation; THA = total hip arthroplasty; VTE = venous thrombo-embolism prophylaxis.
Values in brackets are percentages unless otherwise indicated. P-values <0.05 in bold text (for comparisons between re-revised and non re-revised hips).
*** Missing data for stated number of hips: revision BMI (n=682); revision procedure (n=5); revision femoral head size (n=29); revision bearing surface (n=119).**
**** Data suppressed due to small count within the cell. The actual number was between 1 and 5.**

6.4.3. Re-revision surgery

Re-revision surgery was performed in 7.6% (n=192) of hips at a mean time of 1.2 years (range 1 day to 5.7 years) following ARMD revision. In re-revised hips, 37.0% (n=71) had more than one indication for failure. The commonest re-revision indications were pain (27.1%; n=52), ARMD (25.0%; n=48), dislocation/subluxation (22.4%; n=43), infection (16.1%; n=31), and acetabular component loosening (15.1%; n=29) (Table 6.2.). The mean follow-up time for hips not undergoing re-revision surgery was 2.9 years (range 1.0-6.6 years). The cumulative implant survival rate following ARMD revision surgery was 89.5% (95% CI=87.3%-91.3%) at five years (Figure 6.3.).

Re-revision indications and intra-operative findings	Number of hips (%)
All indications - Hips with 1 re-revision indication - Hips with 2 to 4 re-revision indications	291 in 192 re-revised hips 121 (63.0%) 71 (37.0%)
Pain	52 (27.1%)
Adverse reactions to metal debris	48 (25.0%)
Dislocation / subluxation	43 (22.4%)
Infection	31 (16.1%)
Aseptic loosening – acetabular	29 (15.1%)
Aseptic loosening – femoral	15 (7.8%)
Implant malalignment – acetabular	15 (7.8%)
Osteolysis – femoral	10 (5.2%)
Other abnormality	10 (5.2%)
Peri-prosthetic fracture – femoral	8 (4.2%)
Peri-prosthetic fracture – acetabular	8 (4.2%)
Acetabular component wear	8 (4.2%)
Other (including femoral component malalignment, implant fracture, and liner dissociation)	8 (4.2%)
Osteolysis – acetabular	6 (3.1%)

Table 6.2 Indications for hips undergoing re-revision surgery following revision surgery performed for adverse reactions to metal debris.

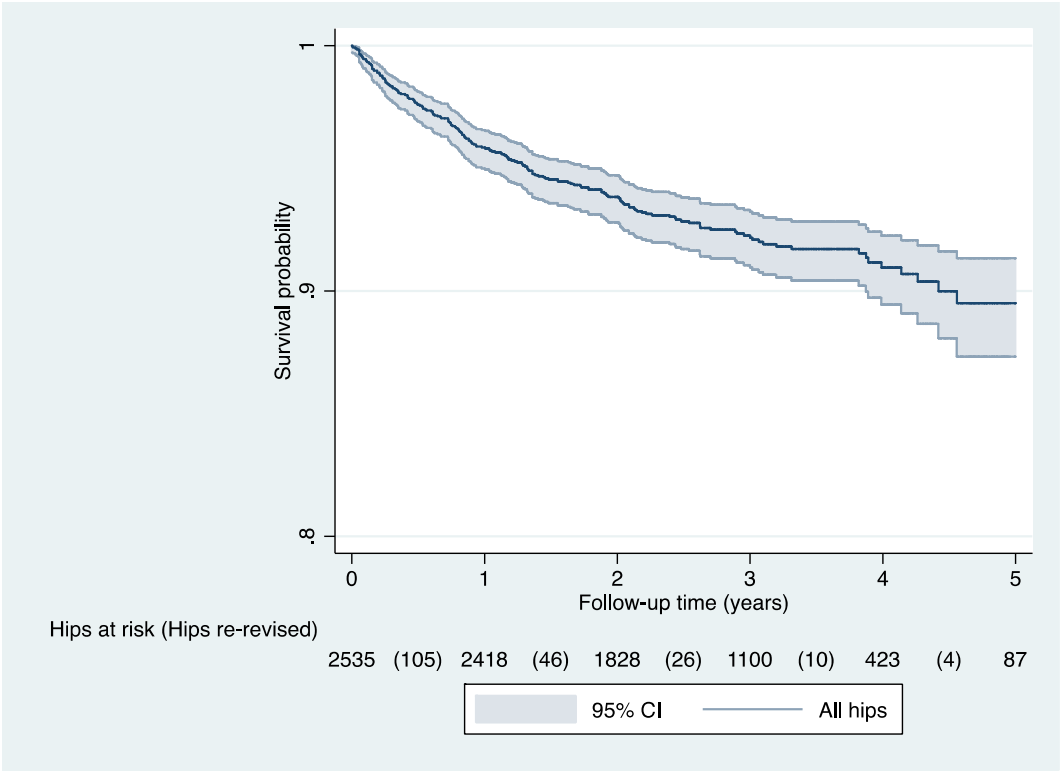


Figure 6.3 Kaplan-Meier cumulative implant survival rate following revision surgery performed for adverse reactions to metal debris at up to five years.
CI = confidence interval. Risk table indicates the number of hips at risk at one year intervals, with the corresponding number in brackets detailing the number of hips undergoing re-revision surgery during each one year interval. One re-revision procedure was performed more than five years following revision surgery, therefore has not been included in the risk table.

6.4.4. Predictors of re-revision surgery

6.4.4.1. Univariable models

Univariable analysis identified six significant predictors of re-revision surgery (Table 6.3.). These predictors were: high BMI at revision; primary THA (versus HRA); shorter time interval between primary and revision surgery; modular component only revisions; ceramic-on-ceramic revision bearings; and acetabular bone grafting.

Covariate	Univariable sub-hazard ratio (95% CI)	p-value	Multivariable sub-hazard ratio (95% CI)	p-value
Gender				
Female *vs. male*	1.14 (0.85-1.55)	0.382	^	
Age at revision (per year)	0.99 (0.98-1.01)	0.455	^	
Revision BMI (per kg/m^2)	1.05 (1.02-1.08)	**0.003**	1.06 (1.02-1.09)	**0.001**
Bilateral revisions for ARMD	1.03 (0.63-1.70)	0.901	^	
Primary arthroplasty				
Hip resurfacing arthroplasty	1.00	Ref	^	
MoM THA	1.47 (1.05-2.06)	**0.023**		
Time from primary to revision (per year)	0.87 (0.79-0.95)	**0.002**	0.91 (0.81-1.01)	0.067
Revision ASA grade				
1	1.00	Ref	1.00	Ref
2	1.00 (0.69-1.45)	0.989	*	0.946
3 or above	0.92 (0.50-1.70)	0.798	0.58 (0.28-1.21)	0.147
VTE – chemical				
None	1.00	Ref	^	
LMWH (+/-other)	0.66 (0.42-1.04)	0.074		
Aspirin only	0.87 (0.35-2.12)	0.753		
Other	0.72 (0.45-1.15)	0.173		
VTE – mechanical				
Any *vs. none*	0.79 (0.39-1.60)	0.509	^	
Number of ARMD revisions performed per centre (per ten cases)	1.02 (0.99-1.05)	0.118	1.03 (0.99-1.06)	0.056
Revision surgeon grade				
Consultant *vs. other*	0.83 (0.42-1.62)	0.583	^	
Surgical approach				
Posterior *vs. other*	0.91 (0.63-1.32)	0.624	^	
Revision indications				
ARMD only	1.00	Ref	^	
2 to 6 indications	0.80 (0.58-1.09)	0.156		
Revision procedure performed				
Cup only	1.00	Ref	1.00	Ref
All components	0.78 (0.56-1.07)	0.126	*	0.878
Stem only	0.22 (0.30-1.59)	0.133	**	**
Modular component only	1.59 (1.04-2.43)	**0.032**	2.01 (1.19-3.38)	**0.009**
Revision femoral head size (mm)				
<36	1.00	Ref	^	
36	1.12 (0.82-1.53)	0.491		
>36	1.04 (0.53-2.02)	0.915		

CONTINUED /...

Covariate	Univariable subhazard ratio (95% CI)	p-value	Multivariable subhazard ratio (95% CI)	p-value
Revision bearing surface				
CoP	1.00	Ref	1.00	Ref
CoC	1.54 (1.10-2.16)	**0.013**	1.86 (1.23-2.80)	**0.003**
MoP	1.44 (0.98-2.10)	0.062	1.45 (0.88-2.39)	0.141
CoM, MoM or MoC	**	**	**	**
Revision cup fixation				
Uncemented	1.00	Ref	^	
Cemented	0.89 (0.54-1.47)	0.656		
Revision stem fixation				
Uncemented	1.00	Ref	^	
Cemented	0.76 (0.44-1.30)	0.316		
Bone graft (femoral)	0.74 (0.28-1.96)	0.541	^	
Bone graft (acetabular)	1.52 (1.10-2.11)	**0.011**	2.10 (1.43-3.07)	**<0.001**

Table 6.3 Competing-risk regression analysis to identify predictors of re-revision surgery following revision surgery performed for adverse reactions to metal debris.
ARMD = adverse reactions to metal debris; ASA = American Society of Anesthesiologists; BMI = body mass index; CoC = ceramic-on-ceramic; CoM = ceramic-on-metal; CoP = ceramic-on-polyethylene; CI = confidence interval; LMWH = low molecular weight heparin; MoC = metal-on-ceramic; MoM = metal-on-metal; MoP = metal-on-polyethylene; Ref = reference group; THA = total hip arthroplasty; VTE = venous thrombo-embolism prophylaxis. P-values <0.05 in bold text.
^ Covariate was not eligible for inclusion in the final multivariable model.
*** Specific subgroup did not meet inclusion criteria for final multivariable model (p>0.20).**
**** Unable to calculate value as no hips in this subgroup underwent re-revision surgery.**
All univariable analyses (apart from revision BMI) were based on a cohort of 2,416 hips with complete data available for all variables, excluding revision BMI (181 hips undergoing re-revision surgery and 39 deaths). Multivariable analysis, and the univariable analysis for revision BMI were based on a cohort of 1,766 hips with data available for all variables, including revision BMI (132 hips undergoing re-revision surgery and 28 deaths).

6.4.4.2. Final multivariable model

The final multivariable model included 1,766 hips (70% of whole cohort) with data available for all variables, including revision BMI. Four predictors from the univariable analysis remained significant predictors of re-revision surgery in the multivariable model (Table 6.3.): high BMI at revision (sub-hazard ratio (SHR)=1.06 per kg/m^2 increase; 95% CI=1.02-1.09; p=0.001); modular component only revisions (SHR=2.01; 95% CI=1.19-3.38; p=0.009); ceramic-on-ceramic revision bearings (SHR=1.86; 95% CI=1.23-2.80; p=0.003); and acetabular bone grafting (SHR=2.10; 95% CI=1.43-3.07; p<0.001). Type of primary MoMHA was not eligible for final model inclusion

(p=0.670). Time between primary and revision surgery was included in the final model but did not reach statistical significance (p=0.067). The final multivariable model had reasonable discriminatory ability for distinguishing between hips undergoing and not undergoing re-revision surgery (*c* statistic=0.66).

6.4.4.3. Analysis of multilevel predictors in the final multivariable model

The type of revision procedure performed and the revision bearing surfaces implanted both significantly influenced the risk of re-revision surgery in the final multivariable model. Both of these covariates contained four different levels. The reference groups used in the final multivariable model for each of these covariates were cup only revisions for the type of revision procedure, and ceramic-on-polyethylene for the revision bearing surface. To account for the multiple levels within each covariate the reference groups were changed to allow other clinically relevant comparisons to be made. Although modular component only revisions predicted re-revision surgery in the final multivariable model, only MoM THAs, and not HRAs, could undergo this type of revision procedure. Therefore the multivariable analysis was repeated in the MoM THA subgroup alone.

Modular component only revisions had an increased risk of re-revision compared with hips undergoing all component revisions (SHR=2.03; 95% CI=1.17-3.54; p=0.012), and compared with hips undergoing cup only revisions (SHR=2.01; 95% CI=1.19-3.38; p=0.009). Re-revisions following modular component only revisions (n=30) were most frequently for dislocation/subluxation (33.3%; n=10) and infection (26.7%; n=8). Multivariable analysis based on MoM THAs alone confirmed that modular component

only revisions remained a risk factor for re-revision (SHR=1.58; 95% CI=1.02-2.45; p=0.038).

Ceramic-on-ceramic revision bearings had an increased risk of re-revision compared with hips revised using ceramic-on-polyethylene bearings (SHR=1.86; 95% CI=1.23-2.80; p=0.003), but not compared with hips revised using metal-on-polyethylene bearings (SHR=1.27; 95% CI=0.80-2.02; p=0.308). Re-revisions following the use of ceramic-on-ceramic revision bearings (n=72) were most frequently for pain (34.7%; n=25), ARMD (27.8%; n=20), and acetabular component loosening (22.2%; n=16).

6.4.4.4. Sensitivity analysis

Regression analyses were repeated using the complete dataset of 2,535 MoMHAs revised for ARMD after the missing data values were estimated using multiple imputation. The same four factors were identified as significant predictors of re-revision surgery in the multivariable model based on the complete imputed dataset (Appendix 21.): high BMI at revision, modular component only revisions, ceramic-on-ceramic revision bearings, and acetabular bone grafting. The effect sizes for these four predictors in this sensitivity analysis were similar to those observed when the analysis was based on the smaller subgroup with complete data available (Chapter 6.4.4.2.).

Shorter time interval between primary and revision surgery, and centres performing higher volumes of ARMD revision surgery were both statistically significant predictors of re-revision in the multivariable model based on the complete imputed dataset (Appendix 21.). These two factors were not statistically significant predictors in the smaller subgroup with complete data available (Table 6.3.). Although the effect sizes

for these two covariates were almost identical between the two models, the difference in statistical significance was related to the imputed dataset having greater statistical power and precision (imputed dataset contained 2,535 observations compared to 1,766 in the complete case analysis).

6.5. Study strengths and limitations: ARMD outcome study

Study strengths include using linked data from the world's largest arthroplasty registry, which provides details of the primary MoMHA and all subsequent procedures even when performed at different institutions. The large cohort size ensures adequate statistical power, and given the large number of re-revisions the final regression model was not over-fitted. Furthermore assessing an unselected population reduces the likelihood of sampling bias. Therefore it is suspected that the findings have good external validity and generalisability, though this requires validation. Finally, recent studies validating NJR data using MoMHA retrievals reported that when the revision procedure was captured within the NJR the data completion and accuracy were excellent [96, 137].

Registries have recognised limitations, which could affect the study findings. Using observational data means causality cannot be inferred. Non-registry validation is therefore recommended; however validating the predictive factors with adequate statistical power will be difficult given specialist centres perform relatively few ARMD revisions [12, 28, 29, 191, 195, 203, 247]. Recent studies validating NJR data using MoMHA retrievals concluded that NJR revision rates may be considerably underestimated [96, 137]. Therefore the present cohort only involved a sample of all ARMD revisions actually performed nationwide. Furthermore the subsequent re-

revision rates may also have been underestimated; hence the re-revision rates reported in this chapter represent the best-case scenario. As ARMD only became a revision indication in the NJR from June 2008, any ARMD revisions performed prior to this time would not be included within the present cohort (Chapter 6.3.2.2.). These excluded early revisions may have represented the most severe ARMD revisions with the worst outcomes [12], which must be considered when interpreting the outcomes presented. The NJR does not collect histopathological data, therefore some MoMHAs truly revised for ARMD but recorded in the NJR using different indications have been missed. Conversely some ARMD revisions included in this cohort may have been misdiagnosed intra-operatively and truly been revised for non-ARMD indications, such as infection. Registries do not collect data on non-revision procedures or patient-reported outcomes following revision [2, 11]. Further surgery represents an important outcome measure but some procedures performed for dislocations (closed reductions), infections (debridement and washout), and peri-prosthetic fractures (internal fixation) occurring after ARMD revision will not have been captured within the NJR.

This study lacks a formal control group, such as patients undergoing conventional THA revision surgery. It was therefore not possible to explore whether the predictors of re-revision surgery identified would also be important in a control group undergoing revision surgery. However the study outcomes were compared with aggregated data from the same registry on patients undergoing conventional THA revision surgery (Chapter 6.9.1.4.) [2], with a formal comparison with MoMHA patients undergoing non-ARMD revision surgery also performed (Chapter 6.7.). A further limitation is that institutions and surgeons may have used different criteria for both diagnosing ARMD and performing revision. The ARMD revisions analysed may therefore comprise a

heterogeneous group, ranging from small peri-prosthetic collections and/or metallosis to large invasive and destructive lesions [12, 28, 29, 191, 195, 203, 247]. Although missing data for some variables (namely revision BMI) could potentially affect the final model, the study findings were supported when missing data had been imputed (Appendix 21.). Finally, the analysis could not be stratified by MoMHA implant manufacturer given access to this sensitive data was not authorised by the NJR.

6.6. Patients and methods: ARMD vs. non-ARMD outcome study

6.6.1. Study design and data source

A retrospective observational cohort study was performed using data from the NJR. The data used for the present study included all primary MoMHAs subsequently undergoing revision surgery for any indication. Information about the data source, including the data extract and the cleaning process, has been described in detail (Chapter 6.3.2.).

6.6.2. Study inclusion and exclusion criteria

There were 5,867 MoMHA revision procedures initially eligible for study inclusion (2,566 ARMD and 3,301 non-ARMD). Hips were subsequently excluded for the following reasons: (1) it was not possible to confirm that the primary arthroplasty had a MoM bearing surface (three hips excluded); (2) the revision was recorded as a two-stage procedure with the excision arthroplasty (first stage) captured in the NJR, but the re-implantation procedure (second stage) was absent (73 hips excluded); and (3) the intra-operative findings included “other” but ARMD was not recorded as one of the revision indications (895 hips excluded). The latter criterion was considered appropriate because the recognition of ARMD as a clinical problem associated with MoMHAs

occurred over time [12, 14, 16]. Therefore it was possible that some or all of these "other" cases may have represented undiagnosed ARMD or incorrectly recorded ARMD. The final study cohort included 4,896 primary MoMHAs undergoing revision surgery for any indication (2,535 ARMD and 2,361 non-ARMD). These cases have subsequently been referred to as the unmatched cohort.

6.6.3. Exposure variables

The study exposure of interest was whether a MoMHA revision procedure was performed for ARMD or non-ARMD indications. When revision procedures were performed in two-stages the revision indications listed at the second stage procedure were used to assign cases to either the ARMD or non-ARMD group, given that these were considered more accurate compared with revision indications recorded at the first stage procedure (Chapter 6.3.3.).

Differences in baseline characteristics between the ARMD and non-ARMD groups could potentially cause confounding by revision indication, which could subsequently affect any comparison of the study outcomes between the two groups. Revision procedures for ARMD and non-ARMD indications were matched for multiple potential confounding factors using propensity score techniques [312, 313]. The matching process aims to reduce confounding by revision indication and has been detailed subsequently (Chapter 6.6.5.).

Further analysis was performed with the non-ARMD group subdivided into specific revision indications. When multiple revision indications were recorded, the following hierarchy was used which was developed using previous studies [140, 314]: (1) ARMD;

(2) infection; (3) peri-prosthetic fracture; (4) aseptic loosening and/or lysis; (5) dislocation/subluxation; (6) other (includes implant malalignment, head/socket mismatch, acetabular component wear, implant fracture, and liner dissociation); (7) unexplained pain. Any given revision indication may include anything listed after it but nothing listed before it. Revisions for peri-prosthetic fracture, aseptic loosening, lysis, implant malalignment, and implant fracture all included failures of the acetabular and/or femoral side [140, 314].

6.6.4. Outcomes

The three study outcomes of interest were: (1) intra-operative complications during the MoMHA revision procedure; (2) all-cause mortality following revision surgery, and (3) all-cause re-revision surgery following revision. Details about intra-operative complications and the definition used for all-cause re-revision surgery have been described (Chapter 6.3.5.).

6.6.5. Statistical analysis

Multiple imputation techniques were used to estimate missing data values for revision BMI, revision femoral head size, and revision bearing surface (Chapter 6.3.6.1.). There was no missing data for revision procedure in this study as a simpler grouping was used (complete versus incomplete).

Propensity score matching was performed using R. Revision procedures for ARMD and non-ARMD indications were matched for multiple potential confounding factors using propensity score techniques [312, 313]. The following covariates were used in the matching process: gender, age, revision BMI, bilateral MoMHA revision procedures,

primary hip diagnosis, primary arthroplasty (THA versus HRA), year of revision, time from primary to revision surgery, revision ASA grade, revision venous thromboembolism prophylaxis, revision surgeon grade, revision surgical approach, number of revision indications (single versus multiple), revision procedure (complete versus incomplete), revision bearing surface, revision femoral head size, and the use of bone graft. Logistic regression was used to generate a propensity score. The propensity score represented the probability that a MoMHA was revised for ARMD. The ARMD and non-ARMD groups were subsequently matched based on the propensity score generated. The matching ratio applied was one-to-one, with a standard deviation caliper width of 0.02 used. Any revision procedure that could not be matched to a procedure in the other group based on the propensity score was automatically dropped.

Cumulative patient and implant survival rates following MoMHA revision were determined using the Kaplan-Meier method. The endpoint for implant survival was re-revision surgery. Patients not undergoing re-revision surgery who remained alive were censored on the study end date (14th August 2015). Outcomes following revision surgery were compared between ARMD and non-ARMD groups using logistic (intra-operative complications) and Cox (mortality and re-revision surgery) regression modelling. For Cox regression analysis, proportional hazards assumptions were assessed using Schoenfeld's residuals. As the ARMD and non-ARMD groups were matched, univariable regression models were used for the analysis. Adjusted regression models were also assessed to account for any residual covariate imbalance following matching. Cox regression analysis was also used to compare the different indications for re-revision surgery between the matched ARMD and non-ARMD groups. As a

sensitivity analysis, outcomes were also assessed using a complete case dataset, which excluded revision BMI only given the large proportion of missing BMI data (35%).

6.7. Results: ARMD vs. non-ARMD outcome study

6.7.1. Propensity score matching

The unmatched cohort included 4,896 primary MoMHAs undergoing revision surgery for any indication (2,535 ARMD and 2,361 non-ARMD). Baseline characteristics differed between the ARMD and non-ARMD groups in the unmatched cohort (Table 6.4.). This was confirmed when examining the distribution of the propensity scores before matching (Figure 6.4.). The matching process improved the balance of the baseline characteristics between the ARMD and non-ARMD groups (Table 6.4.). This was confirmed when examining the distribution of the propensity scores following matching (Figure 6.5.). After matching there remained some residual imbalance between the ARMD and non-ARMD groups for four covariates: bilateral MoMHA revisions; primary MoMHA; time from primary to revision surgery; and year of revision surgery. To account for this residual imbalance adjusted regression models have also been assessed for each outcome of interest.

As a complete dataset was generated using multiple imputation before performing the matching process, the distribution of the propensity scores for each of the three outcomes of interest was slightly different. This in turn resulted in the overall matched cohort sizes for each outcome of interest being slightly different: 2,588 hips for intra-operative complications (1,294 ARMD vs. 1,294 non-ARMD); 2,558 hips for mortality

(1,279 ARMD vs. 1,279 non-ARMD); 2,576 hips for re-revision surgery (1,288 ARMD vs. 1,288 non-ARMD).

	Unmatched cohort			**Matched cohort § (For re-revision surgery analysis)**		
Covariate	**All hip revisions (n=4,896) (100%)**	**ARMD revisions (n=2,535) (52%)**	**Non-ARMD revisions (n=2,361) (48%)**	**All hip revisions (n=2,576) (100%)**	**ARMD revisions (n=1,288) (50%)**	**Non-ARMD revisions (n=1,288) (50%)**
Gender	^					
Female *vs. male*	2,764 (56.5)	1,509 (59.5)	1,255 (53.2)	1,451 (56.3)	713 (55.4)	738 (57.3)
Age at revision (yr)	^					
Mean (SD)	62.2 (10.7)	63.6 (10.1)	60.8 (11.2)	62.0 (10.6)	62.2 (10.1)	61.8 (11.1)
Revision BMI (kg/m^2)	^ *****					
Mean (SD)	29.2 (5.2)	29.0 (5.1)	29.4 (5.4)	29.3 (5.3)	29.3 (5.4)	29.3 (5.2)
Bilateral MoM hip revision surgery	^			^^		
	436 (8.9)	286 (11.3)	150 (6.4)	199 (7.7)	84 (6.5)	115 (8.9)
Primary diagnosis						
Primary OA *vs. other*	4,579 (93.5)	2,378 (93.8)	2,201 (93.2)	2,420 (93.9)	1,213 (94.2)	1,207 (93.7)
Primary arthroplasty	^			^^		
MoM THA	2,903 (59.3)	1,716 (67.7)	1,187 (50.3)	1,373 (53.3)	660 (51.2)	713 (55.4)
Hip resurfacing	1,993 (40.7)	819 (32.3)	1,174 (49.7)	1,203 (46.7)	628 (48.8)	575 (44.6)
Primary femoral head size (mm)	^					
Mean (SD)	44.3 (5.9)	44.0 (5.5)	44.6 (6.2)	44.4 (5.8)	44.3 (5.6)	44.5 (6.0)
Time from primary to revision (yr)	^			^^		
Mean (SD)	4.6 (2.2)	5.2 (1.8)	3.9 (2.3)	5.0 (2.0)	5.3 (1.9)	4.8 (2.1)
Year of revision	^			^^		
2008	132 (2.7)	1 (0.04)	131 (5.6)	0 (0.0)	0 (0.0)	0 (0.0)
2009	393 (8.0)	31 (1.2)	362 (15.3)	58 (2.3)	31 (2.4)	27 (2.1)
2010	598 (12.2)	171 (6.8)	427 (18.1)	264 (10.3)	114 (8.9)	150 (11.7)
2011	1,025 (20.9)	568 (22.4)	457 (19.4)	495 (19.2)	184 (14.3)	311 (24.2)
2012	1,248 (25.5)	730 (28.8)	518 (21.9)	712 (27.6)	318 (24.7)	394 (30.6)
2013	994 (20.3)	698 (27.5)	296 (12.5)	589 (22.9)	334 (25.9)	255 (19.8)
2014	506 (10.3)	336 (13.3)	170 (7.2)	458 (17.8)	307 (23.8)	151 (11.7)
Revision ASA grade	^					
1	966 (19.7)	476 (18.8)	490 (20.8)	512 (19.9)	254 (19.7)	258 (20.0)
2	3,376 (69.0)	1,827 (72.1)	1,549 (65.6)	1,767 (68.6)	872 (67.7)	895 (69.5)
3 or above	554 (11.3)	232 (9.2)	322 (13.6)	297 (11.5)	162 (12.6)	135 (10.5)
VTE – chemical	^					
LMWH (+/-other)	2,709 (55.3)	1,309 (51.6)	1,400 (59.3)	1,481 (57.5)	760 (59.0)	721 (56.0)
Aspirin only	217 (4.4)	68 (2.7)	149 (6.3)	96 (3.7)	48 (3.7)	48 (3.7)
Other	1,435 (29.3)	922 (36.4)	513 (21.7)	783 (30.4)	387 (30.1)	396 (30.8)
None	535 (10.9)	236 (9.3)	299 (12.7)	216 (8.4)	93 (7.2)	123 (9.6)
VTE – mechanical	^					
Any *vs. none*	4,529 (92.5)	2,447 (96.5)	2,082 (88.2)	2,422 (94.0)	1,207 (93.7)	1,215 (94.3)
Revision surgeon						
Consultant *vs. other*	4,693 (95.9)	2,430 (95.9)	2,263 (95.9)	2,473 (96.0)	1,236 (96.0)	1,237 (96.0)
Surgical approach	^					
Posterior *vs. other*	3,802 (77.7)	2,084 (82.2)	1,718 (72.8)	1,972 (76.6)	969 (75.2)	1,003 (77.9)

CONTINUED /...

Covariate	Unmatched cohort			Matched cohort § (For re-revision surgery analysis)		
	All hip revisions (n=4,896) (100%)	ARMD revisions (n=2,535) (52%)	Non-ARMD revisions (n=2,361) (48%)	All hip revisions (n=2,576) (100%)	ARMD revisions (n=1,288) (50%)	Non-ARMD revisions (n=1,288) (50%)
Number of revision indications	^					
1 indication +/- pain	4,066 (83.0)	1,966 (77.6)	2,100 (89.0)	2,242 (87.0)	1,130 (87.7)	1,112 (86.3)
Multiple indications	830 (17.0)	569 (22.4)	261 (11.1)	334 (13.0)	158 (12.3)	176 (13.7)
Revision procedure	^					
All components	2,314 (47.3)	1,032 (40.7)	1,282 (54.3)	1,384 (53.7)	715 (55.5)	669 (51.9)
Incomplete revision**	2,582 (52.7)	1,503 (59.3)	1,079 (45.7)	1,192 (46.3)	573 (44.5)	619 (48.1)
Revision femoral head size (mm)	^ *					
Mean (SD)	34.7 (4.4)	34.3 (3.2)	35.1 (5.4)	34.1 (3.5)	34.0 (3.3)	34.1 (3.7)
Bearing surface	^ *					
CoP	1,581 (34.5)	1,104 (45.7)	477 (22.1)	848 (32.9)	434 (33.7)	414 (32.1)
CoC	1,400 (30.6)	765 (31.7)	635 (29.4)	891 (34.6)	434 (33.7)	457 (35.5)
MoP	1,282 (28.0)	533 (22.1)	749 (34.7)	809 (31.4)	406 (31.5)	403 (31.3)
CoM, MoM or MoC	313 (6.8)	14 (0.6)	299 (13.8)	28 (1.1)	14 (1.1)	14 (1.1)
Bone graft (femoral)	146 (3.0)	76 (3.0)	70 (3.0)	68 (2.6)	40 (3.1)	28 (2.2)
Bone graft (acetabular)	^ 860 (17.6)	509 (20.1)	351 (14.9)	450 (17.5)	219 (17.0)	231 (17.9)

Table 6.4 Patient and surgical factors in the unmatched cohort and in the matched cohort.
ARMD = adverse reactions to metal debris; ASA = American Society of Anesthesiologists; BMI = body mass index; CoC = ceramic-on-ceramic; CoM = ceramic-on-metal; CoP = ceramic-on-polyethylene; LMWH = low molecular weight heparin; MoC = metal-on-ceramic; MoM = metal-on-metal; MoP = metal-on-polyethylene; NA = not applicable; OA = osteoarthritis; SD = standard deviation; THA = total hip arthroplasty; VTE = venous thrombo-embolism prophylaxis.

Values in brackets are percentages unless otherwise indicated.

§ The matched cohort presented was based on the analysis performed for re-revision surgery. Similar distributions for all covariates were obtained when assessing intra-operative complications and mortality in the respective matched cohorts.

*** Missing data for stated number of hips: revision BMI (n=1,731); revision femoral head size (n=111); revision bearing surface (n=320). Multiple imputation was performed before matching, therefore there was no missing data in the matched cohort.**

**** Incomplete revisions included: cup only revisions; stem only revisions; and modular component only revisions.**

^ Statistically significant difference ($p<0.05$) in covariate between ARMD and non-ARMD revision groups in the unmatched cohort. Unpaired t-tests were used to compare numerical data, and Chi-squared tests were used to compare categorical data.

^^ Statistically significant difference ($p<0.05$) in covariate between ARMD and non-ARMD revision groups in the matched cohort. Unpaired t-tests were used to compare numerical data, and Chi-squared tests were used to compare categorical data.

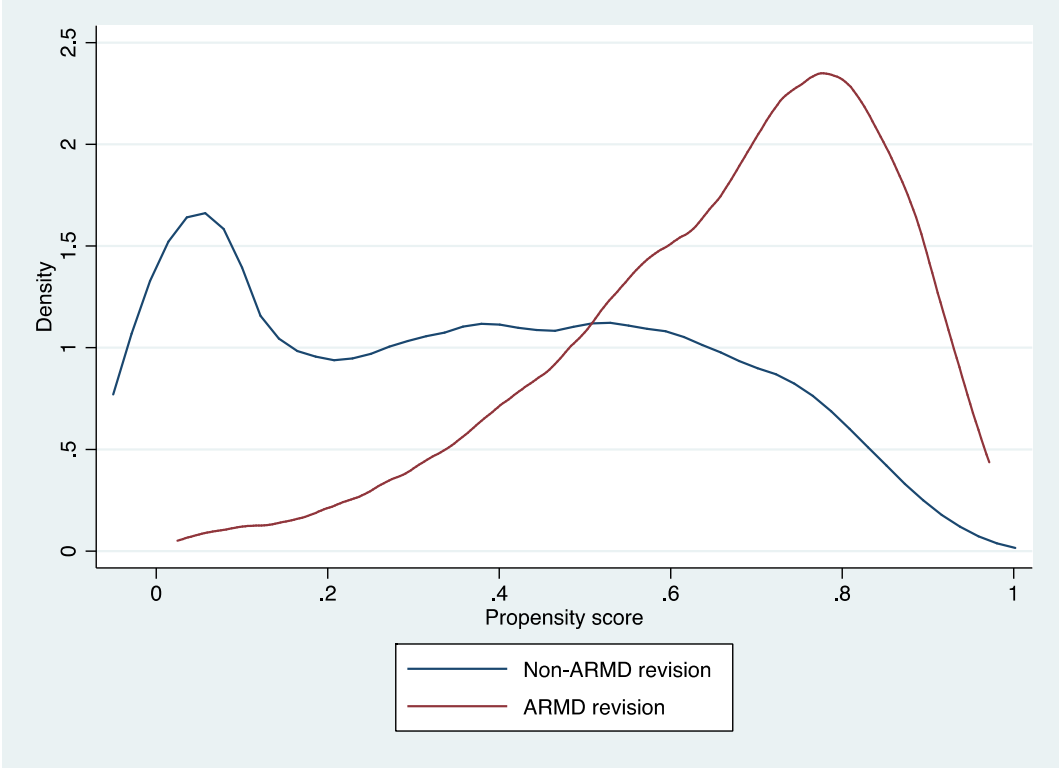


Figure 6.4 Kernel density plot illustrating the distribution of the propensity scores before matching.
ARMD = adverse reactions to metal debris.
This plot relates to the matching process when assessing re-revision surgery. Similar plots were obtained when examining the other two outcomes of interest (intra-operative complications and mortality). The plot illustrates that the distribution of the propensity scores before matching were markedly different between the ARMD and non-ARMD revision groups.

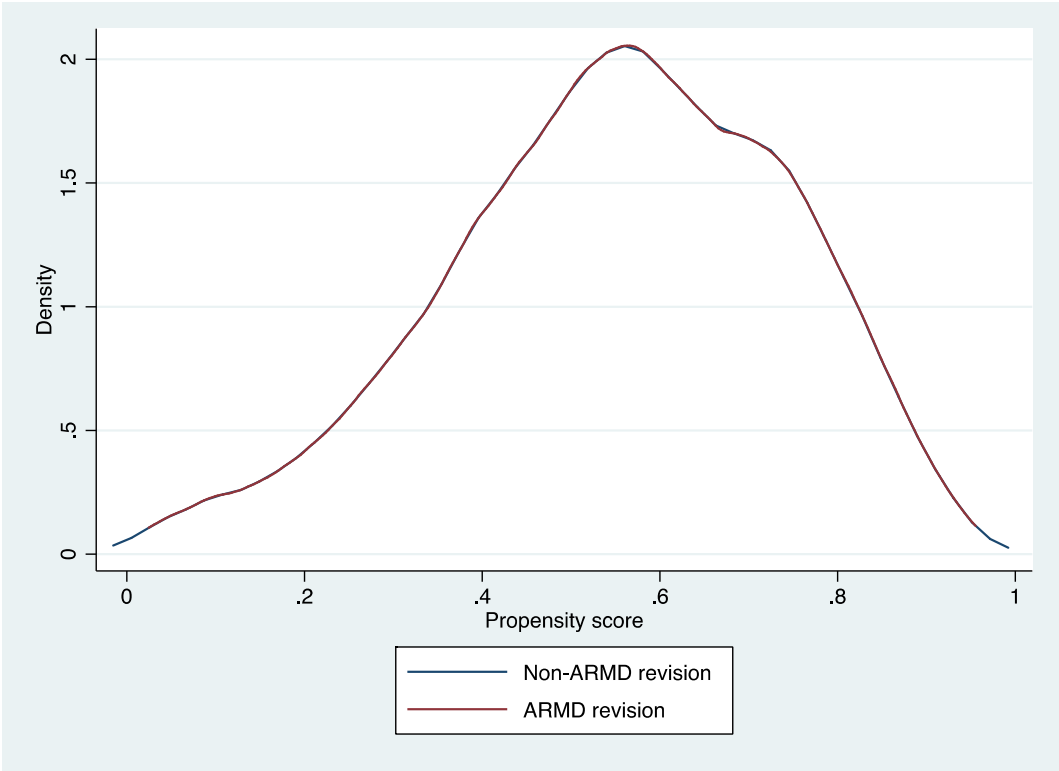


Figure 6.5 Kernel density plot illustrating the distribution of the propensity scores after matching. ARMD = adverse reactions to metal debris.
This plot relates to the matching process when assessing re-revision surgery. Similar plots were obtained when examining the other two outcomes of interest (intra-operative complications and mortality). The plot illustrates that the distribution of the propensity scores after matching were identical between the ARMD and non-ARMD revision groups, with the two distribution lines completely overlapping.

6.7.2. Intra-operative complications

The overall risk of intra-operative complications in the matched cohort was 2.4% (n=63), with calcar cracks and greater trochanteric fractures representing the commonest complications. The risk of intra-operative complications between ARMD (2.4%) and non-ARMD (2.5%) revision groups was similar (odds ratio=0.97; 95% CI=0.59-1.60; p=0.899) (Table 6.5.). The adjusted regression model produced similar results (adjusted odds ratio=0.95; 95% CI=0.57-1.59; p=0.851).

	Risk of intra-operative complications	5 year implant survival (95% CI)	5 year patient survival (95% CI)
Overall cohort	2.4% (n=63)	92.3% (90.3%-93.8%)	96.1% (93.9%-97.6%)
ARMD group	2.4% (n=31)	94.3% (91.7%-96.1%)	96.9% (91.5%-98.9%)
Non-ARMD group	2.5% (n=32)	90.5% (87.6%-92.7%)	95.5% (92.8%-97.2%)
Regression analysis (95% CI)	OR = 0.97 (0.59-1.60) p = 0.899 Adjusted OR = 0.95 (0.57-1.59) p = 0.851	HR = 0.52 (0.36-0.75) **p < 0.001** Adjusted HR = 0.58 (0.40-0.84) **p = 0.004**	HR = 0.43 (0.22-0.86) **p = 0.018** Adjusted HR = 0.45 (0.22-0.92) **p = 0.032**

Table 6.5 Outcomes following metal-on-metal hip arthroplasty revision surgery in the matched cohort.
ARMD = adverse reactions to metal debris; CI = confidence interval; HR = hazard ratio; OR = odds ratio. P-values <0.05 in bold text.

Odds ratios and hazard ratios greater than 1 represent an increased risk of the specified outcome in the ARMD revision group.

6.7.3. Mortality

The overall mortality risk in the matched cohort was 1.6% (n=41). Mean time from revision surgery to death was 2.2 years (range 0.1-5.9 years), with no hips undergoing re-revision before death. Mortality rates were significantly lower following ARMD revision surgery compared with non-ARMD revision surgery (Table 6.5.). The five year cumulative patient survival rate following ARMD revision surgery was 96.9% (95% CI=91.5%-98.9%) compared with 95.5% (95% CI=92.8%-97.2%) following non-ARMD revision surgery (HR=0.43; 95% CI=0.22-0.86; p=0.018). The adjusted regression model produced similar results (adjusted HR=0.45; 95% CI=0.22-0.92; p=0.032).

6.7.4. Re-revision surgery

The overall re-revision risk in the matched cohort was 5.2% (n=134). Mean time from revision surgery to re-revision was 1.4 years (range 1 day to 5.7 years). Re-revision indications included dislocation/subluxation (20.9%; n=28), ARMD (19.4%; n=26), infection (17.2%; n=23), aseptic loosening and/or lysis (17.2%; n=23), other causes (11.2%; n=15), unexplained pain (7.5%; n=10), and peri-prosthetic fracture (6.7%; n=9). The mean follow-up time for hips not undergoing re-revision surgery was 3.0 years (range 1.0-6.6 years).

Re-revision rates were significantly lower following ARMD revision surgery compared with non-ARMD revision surgery (Table 6.5.). The five year cumulative implant survival rate following ARMD revision surgery was 94.3% (95% CI=91.7%-96.1%) compared with 90.5% (95% CI=87.6%-92.7%) following non-ARMD revision surgery (HR=0.52; 95% CI=0.36-0.75; p<0.001) (Figure 6.6.). The adjusted regression model produced similar results (adjusted HR=0.58; 95% CI=0.40-0.84; p=0.004).

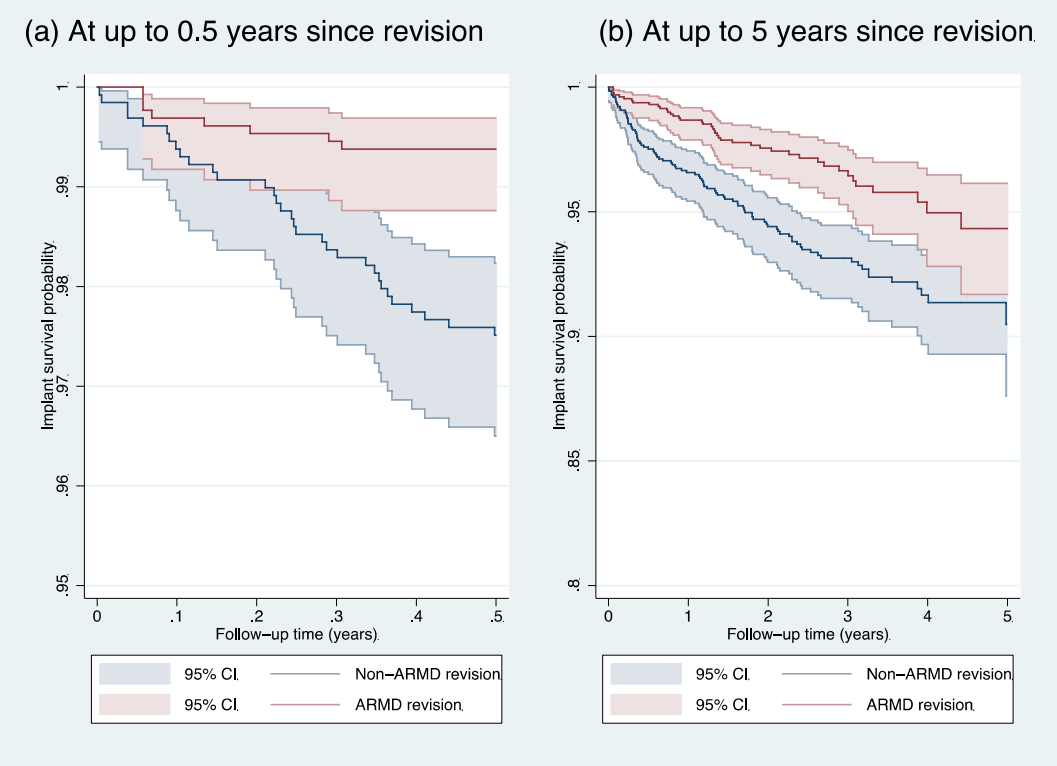


Figure 6.6 Kaplan-Meier cumulative implant survival rate following revision surgery at up to five years by revision indication.
ARMD = adverse reactions to metal debris; CI = confidence interval.
(a) During the first 0.5 years from revision surgery the proportional hazards assumption was not satisfied, with the non-ARMD revision group having a disproportionally higher rate of implant failure compared with the ARMD revision group. (b) From 0.5 to five years follow-up from revision surgery the proportional hazards assumption was satisfied.

Implant survival rates following MoMHA revision surgery have been summarised for specific revision indications (Table 6.6.). Compared with ARMD revision surgery, implant survival rates were significantly lower when MoMHA revisions were performed for infection, peri-prosthetic fracture, aseptic loosening and/or lysis, dislocation/subluxation, and unexplained pain. Revisions performed for infection (81.2%; 95% CI=55.3%-93.0%) and dislocation/subluxation (81.9%; 95% CI=68.8%-89.9%) had the lowest five year implant survival rates.

Indication for revision surgery	Number of hips (% of matched cohort)	5 year implant survival rate following revision surgery (95% CI)	Hazard ratio for Cox regression (95% CI) vs. ARMD revision	p-value
All causes	2,576 (100)	92.3% (90.3%-93.8%)	NA	NA
ARMD	1,288 (50.0)	94.3% (91.7%-96.1%)	Reference group	NA
Infection	104 (4.0)	81.2% (55.3%-93.0%)	2.81 (1.41-5.59)	**0.003**
Peri-prosthetic fracture *	86 (3.3)	89.9% (79.4%-95.2%)	2.35 (1.06-5.23)	**0.036**
Aseptic loosening and/or osteolysis *	499 (19.4)	92.6% (89.3%-94.9%)	1.60 (1.00-2.55)	**0.048**
Dislocation / subluxation	60 (2.3)	81.9% (68.8%-89.9%)	4.58 (2.30-9.13)	**<0.001**
Other indications **	157 (6.1)	91.5% (81.3%-96.2%)	1.57 (0.77-3.22)	0.218
Unexplained pain	382 (14.8)	91.4% (86.9%-94.4%)	1.79 (1.09-2.94)	**0.020**

Table 6.6 Implant survival rates following metal-on-metal hip arthroplasty revision surgery by specific revision indication.
ARMD = adverse reactions to metal debris; CI = confidence interval; NA = not applicable. P-values <0.05 in bold text.

*** Revision indication included failures on the acetabular and/or femoral side.**

**** Other indications included implant malalignment, head/socket mismatch, acetabular component wear, implant fracture, and liner dissociation.**

Hazard ratios greater than 1 represent an increased risk of implant re-revision surgery for the listed revision indication compared with hips revised for ARMD.

Apart from unexplained pain, any listed revision indication may include anything below it but nothing above it.

In the matched cohort the re-revision rates for most indications were not significantly different between ARMD revisions and non-ARMD revisions (Appendix 22.). Compared with ARMD revisions, non-ARMD revisions were at increased risk of undergoing re-revision surgery for dislocation/subluxation, and aseptic loosening and/or lysis.

6.7.5. Sensitivity analysis

Regression analyses were repeated using the complete dataset but with revision BMI excluded. The findings from these analyses when performed in a matched cohort of patients (Appendix 23.) were similar to those obtained in the complete imputed dataset of matched patients (Table 6.5.). As a further sensitivity analysis the regression models were repeated in the unmatched cohort (Appendix 24.), which demonstrated different findings compared with those obtained in the complete imputed dataset of matched patients (Table 6.5.).

6.8. Study strengths and limitations: ARMD vs. non-ARMD outcome study

Strengths of studies based on large registry datasets have been previously described (Chapter 6.5.). In addition the present study used robust statistical methods, which included propensity score matching. This method was employed to minimise the potential for confounding by revision indication, with this a priori decision supported by the findings in the unmatched cohort (Appendix 24.) being different to those observed in the matched cohort (Table 6.5.).

Limitations of studies based on registry datasets have been described (Chapter 6.5.). Although registries potentially underreport failures [96, 137], there is no reason to suspect that underreporting would differ between the ARMD and non-ARMD revision groups. The histopathological and microbiological results from intra-operative revision samples that are not included in the NJR represents a pertinent limitation in the present study. This is because the true diagnosis following sample analysis may have caused cases analysed in one revision group to crossover to the alternate revision group, and

vice-versa. Despite using multiple imputation, missing data for some variables (namely revision BMI) may have affected the study findings. However sensitivity analysis based on the complete dataset (excluding revision BMI) produced similar findings and effect sizes (Appendix 23.) compared with the analysis based on the complete imputed dataset of matched patients (Table 6.5.).

A further limitation relates to the potential selection bias introduced by the study inclusion criteria. Excluding incompletely recorded staged revisions may have affected the non-ARMD revision outcomes, given that most staged revisions are performed for infection [2, 11] and these cases had the lowest implant survival (Table 6.6.). The non-ARMD revisions with "other" indications that were also excluded may have represented undiagnosed or incorrectly recorded ARMD, especially during the early years [12, 14, 16]. Although this provided a more robust cohort for analysis, the earliest and most severe ARMD revisions with the worst outcomes [12] may have been excluded from this analysis. Despite matching there is potential for residual confounding. However the observed effect sizes were large, therefore it is unlikely any unmeasured factors would have been large enough to change the direction of the effect sizes. Matching may have also reduced the generalisability of the study findings. Many revisions performed before 2012 were excluded following matching. These early cases may have again represented the most difficult and aggressive ARMD cases with the poorest outcomes [12, 247, 249]. However the direction of the effect sizes for all outcomes were the same in the unmatched cohort (Appendix 24.) and the matched cohort (Table 6.5.).

6.9. Discussion

6.9.1. Outcomes following MoMHA revision surgery

6.9.1.1. Intra-operative complications

No studies have formally assessed the risk of intra-operative complications in patients undergoing MoMHA revision surgery. This is suspected to be because such events are considered rare and require large cohort studies to assess properly, which are currently lacking. However it is important to define the risk of intra-operative complications given the increasing burden of MoMHA revision surgery [2, 29, 203].

Few intra-operative complications occurred in the unmatched ARMD revision cohort (1.6%). Although the risk of intra-operative complications was slightly higher in the matched cohort (2.4%), this risk did not differ between ARMD and non-ARMD revisions. In both cohorts peri-prosthetic femoral fractures represented the commonest complication, which is consistent with previous revision THA cohorts [315]. The risk of intra-operative peri-prosthetic femoral fractures has been much higher in previous studies (12.3%) [315] compared with the present cohort. However this is suspected to relate to the present study including many revision procedures where the femoral component was retained, and also including revisions of HRAs (designed to conserve femoral bone stock).

6.9.1.2. Mortality

Similarly no studies have formally assessed patient mortality following MoMHA revision surgery. Although mortality may also be relatively rare it is important to

establish the risk, and compare this with patient mortality following other types of hip arthroplasty. Higher patient survival rates following MoMHA revision surgery (95.5%-96.9% at five years) were observed compared with those reported following primary conventional THA (90.5% at five years (95% CI=90.4%-90.6%)) [2]. However this is likely related to MoMHA revision patients (mean 62.0-63.6 years) being younger compared with primary conventional THA patients (mean 69.0 years) [2]. In a previous large propensity score matched study which assessed mortality following THA, the five year patient survival rates by type of arthroplasty were as follows: HRA (98%; mean age 57.2 years); uncemented conventional THA (96%; mean age 57.7 years); and cemented conventional THA (95.5%; mean age 59.4 years) [128]. Therefore the risk of mortality was similar after MoMHA revision surgery compared with patients of a similar age following primary uncemented or cemented conventional THA; however the mortality risk was higher after MoMHA revision surgery when compared to patients following primary HRA.

The matched analysis demonstrated ARMD revisions had approximately half the risk of death compared with non-ARMD revisions. The reason for revision may have contributed to these observed mortality differences. In conventional THA revisions high mortality rates have been reported when performing revisions for infection and peri-prosthetic fracture [316-318]. However this could not be explored further using the current dataset due to the relatively small number of overall patient deaths in the matched cohort (n=41) coupled with the small subgroups of MoMHAs revised for infection (n=104) and peri-prosthetic fracture (n=86).

As patients were matched, factors such as age, ASA grade, and type of revision cannot be responsible for the observed mortality differences. It has been suggested that the most active patients develop ARMD, with increased hip movement potentially contributing to impingement and/or edge loading (Chapter 3.6.2.). Higher general activity levels may therefore help reduce the mortality risk in patients revised for ARMD. Patients revised for ARMD may also have lower mortality compared with non-ARMD revisions because of differences in surgical complexity not captured within the NJR. Revising MoMHAs for ARMD early was adopted to reduce the potential tissue damage thought to be associated with poor outcomes [12, 29, 48, 247]. It is suspected that the surgical complexity of ARMD revisions would be reduced by performing procedures early. This in turn could have potential advantages such as less tissue trauma, reduced blood loss, and shorter operative time and hospital stays. These factors may therefore help explain the reduced mortality risk observed following ARMD revision surgery, although further investigation is recommended.

6.9.1.3. Re-revision surgery: indications for failure

In the unmatched ARMD revision cohort (Chapter 6.4.3.), re-revisions following ARMD surgery occurred early with re-revision indications similar to previous reports [28]. Dislocation/subluxation can occur due to destructive ARMD requiring extensive soft-tissue debridement, and reduction of the femoral head diameter [29, 247]. Repeat surgery and incomplete excision of metal debris and necrotic tissue can contribute to infection. Furthermore ARMD-induced osteolysis may lead to acetabular component loosening following revision [28, 247]. Although ARMD recurrence has been reported following ARMD revision in numerous studies [28], ARMD was detected in 25% of all re-revision cases. This is concerning given the short-term follow-up and exchange to

non-MoM bearings in hips subsequently re-revised. Recurrence may be due to incomplete excision of metal debris and/or ARMD generated from newly implanted modular THA junctions [16, 93, 256]. It is recommended that future studies explore the potential reasons for the observed high frequency of ARMD recurrence.

Similar findings were observed in the matched cohort of MoMHA revisions; namely re-revisions occurred early with the indications for failure consistent with previous studies [28]. Indications for performing re-revision surgery were broadly the same following ARMD revision procedures compared with non-ARMD revisions. Although the numbers were small, compared with ARMD revisions it was observed that non-ARMD revisions were at increased risk of undergoing re-revision surgery for dislocation/subluxation, and aseptic loosening and/or lysis.

6.9.1.4. Re-revision rates following ARMD vs. non-ARMD revision

Numerous studies have reported poor short-term outcomes following the initial experience with ARMD revision surgery, with some patients experiencing catastrophic outcomes [12, 247, 249, 256]. Furthermore implant survival and patient-reported outcomes following ARMD revision surgery have been inferior compared with MoMHA revisions for non-ARMD indications, and compared with matched patients undergoing primary conventional THA [12]. However these early studies had numerous limitations, which include being small underpowered single centre cohorts substantially influenced by selection bias (first MoMHA revisions performed), with the potential for confounding by revision indication to also influence the findings. The only large cohort assessing outcomes following MoMHA revisions (n=884) was reported by the Australian Joint Replacement Registry [243, 263]. This registry study observed high re-

revision rates (up to 23% at five years), with re-revision not influenced by type of revision procedure performed or bearing surface implanted [243, 263]. However important limitations included reporting on fewer than 200 ARMD revisions, excluding all MoM THAs and any HRA revisions performed for infection, including the initial MoMHA revisions performed which may have influenced outcomes, not stratifying outcomes by revision indication, and only reporting on re-revision rates [243, 263].

The present matched analysis demonstrated that ARMD revisions had approximately half the risk of re-revision surgery compared with MoMHAs revised for non-ARMD indications. This finding was completely contradictory to previous observations [12, 247, 249, 256]. Comparison of the present findings with the Australian Joint Replacement Registry cohort is hindered by the previous studies limitations. It is suspected that the reduced re-revision risk observed after ARMD revision surgery in the matched MoMHA cohort is related to the increased awareness of this problem by clinicians, regular MoMHA patient surveillance, and the widespread adoption of a lower threshold for performing ARMD revisions following recommendations from worldwide regulatory authorities and surgeons [12, 23, 24, 29, 48]. Even when acknowledging the study limitations (Chapter 6.8.), it can be concluded that MoMHA revision surgery performed for ARMD has not caused any additional morbidity compared to revisions performed for non-ARMD indications.

Despite the present study lacking a formal control group of patients undergoing conventional THA revision surgery, historical data can be used to make comparisons between these different revision cohorts. As registry datasets have recognised limitations, including the potential to underestimate subsequent failures (Chapter 6.5.),

it is suggested the fairest comparison of the present study findings would be with other registry results. The five year implant survival rates following all-cause conventional THA revision surgery recorded in the NJR range from 87.8% to 89.1% for primary metal-on-polyethylene THAs (depending on implant fixation), and are 88.2% for primary uncemented ceramic-on-ceramic THAs [2]. The five year implant survival rates in the matched MoMHA cohort (ARMD=94.3% and non-ARMD=90.5%) suggest that non-ARMD revisions have similar survival to conventional THA revisions, however ARMD revisions have better implant survival compared with conventional THA revisions [2]. The five year implant survival rate in the unmatched ARMD revision cohort (89.5%) appears comparable with the survival reported following conventional THA revisions for all indications [2]. The variance between implant survival in the unmatched and matched ARMD groups reported in the present chapter is potentially related to limitations in the matching process (Chapter 6.8.). Matching may have eliminated the earliest ARMD revision procedures, which could represent the most difficult and aggressive cases with the poorest outcomes [12, 247, 249]. However regardless of whether the matched or unmatched ARMD cohort from the present study is used, it can be concluded that the implant survival following MoMHA revision surgery performed for ARMD was no worse than that following conventional THA revision surgery. This therefore does suggest that improvements have been made in the outcomes following MoMHA revision surgery performed for ARMD given the extremely poor outcomes that were initially reported [12, 247, 249].

The inferior outcomes observed in matched MoMHA revisions performed for non-ARMD indications are concerning given that many such revisions continue to be performed [2, 11]. Although many non-ARMD indications (aseptic loosening and/or

lysis, peri-prosthetic fracture, and unexplained pain) had inferior implant survival compared with ARMD revisions, MoMHAs revised for infection and dislocation/subluxation had the poorest implant survival. Managing peri-prosthetic joint infection can be challenging with poor outcomes reported following conventional THA revisions for infection [317, 319]. The present study confirmed the same was true for infected MoMHA revisions. Contrary to conventional THA bearing surfaces where infected revisions occur early after primary arthroplasty, the revision risk for infection continues to increase with time since primary MoMHA implantation [2]. Although the reasons for this remain unknown, this is concerning given that many MoMHAs may require revision for infection in the future and the poor outcomes reported here for this subgroup. Dislocation is uncommon following large-diameter MoMHA (Chapter 3.4.1.) [247], but if it does occur the present findings suggest outcomes following revision surgery for instability are poor. It is suspected that controlling hip instability when reducing the femoral head size is challenging, which may become increasingly difficult if there is associated peri-prosthetic soft-tissue damage.

6.9.2. Predicting re-revision surgery following ARMD revision

In the unmatched ARMD revision cohort, the final multivariable model was only reasonable for distinguishing between hips undergoing and not undergoing re-revision. Registry based models are unlikely to have much better discriminatory ability given they do not collect data on important variables, such as blood metal ions and imaging. However the final model did include four predictors of re-revision surgery with large effect sizes.

High BMI at revision and acetabular bone grafting both predicted future re-revision. These are largely considered non-modifiable risk factors based on: (1) the potential for ARMD to be destructive [12, 16, 203], therefore suggesting it is inadvisable to delay revision surgery to lose weight, and (2) any associated acetabular defects requiring appropriate intra-operative reconstruction unless ARMD has been identified early. High BMI predicts re-revision following revision of conventional THAs performed for non-ARMD indications [320, 321]. The present work confirms that this is also true for ARMD revisions. Patients requiring acetabular bone grafting had twice the re-revision risk. Bone grafting suggests more complex reconstructions, which may be required because of ARMD-induced osteolysis and/or iatrogenic bone loss when removing well-fixed components [29, 195, 203]. Alternatively acetabular bone grafting may represent a modifiable risk factor if the reasons for the higher re-revision rates are related to this strategy being ineffective for managing ARMD-induced osteolysis compared with other reconstructive methods. Complications associated with bone grafts include infection, component migration and loosening, which can all require further surgery [322]. Future research is therefore needed to examine the higher failure rates observed following acetabular bone grafting.

Modular component only revisions and ceramic-on-ceramic revision bearings both predicted future re-revision, and represent modifiable surgical risk factors. Some surgeons advise retaining well-fixed and well-positioned THA components at ARMD revision, with exchange of the modular components to non-MoM bearings [191, 247]. Taper adapters have also been recommended if tapers are not severely damaged [93, 247]. Modular exchange has been attractive given it prevents the morbidity associated with removal of well-fixed components [253, 315]. However there is little evidence

supporting this approach [93, 247]. In the present cohort modular component only revisions had twice the re-revision risk compared to all component revisions and cup-only revisions. Failures following modular revisions were commonly due to dislocation. This supports previous studies employing this strategy [93, 247], suggesting hip stability is difficult to control after ARMD revision with modular component exchange alone. Component malposition represents an important risk factor for subsequent arthroplasty failure [150, 323]. By not correcting any component malposition at ARMD revision, hips undergoing modular only revisions would be expected to be at increased risk of subsequent failure, especially for dislocation/subluxation. However the NJR does not collect information on primary MoMHA component positioning, therefore this cannot be considered further.

There is no consensus on which non-MoM bearing surface to use at ARMD revision. Ceramic bearings are popular given MoMHA revision patients are generally young and active [29, 203, 247, 256]. However ceramic-on-ceramic bearings had an 86% increased risk of re-revision compared to ceramic-on-polyethylene bearings in this study. It is unclear why hard-on-hard ceramic bearings have inferior outcomes, with re-revision indications in this subgroup being similar to the remaining cohort. However the present analysis was limited by not knowing the specific type of ceramic head implanted. The observations that ceramic-on-polyethylene bearings have the lowest re-revision risk following ARMD revision support registry observations regarding the lower failure rate for this bearing surface in primary THA [2, 11].

Primary MoM THAs have inferior survivorship compared to HRAs, even when using identical bearing surfaces [2, 11, 179]. In MoMHAs undergoing ARMD revision, the

primary implant (MoM THA vs. HRA) did not influence future re-revision risk in this study. Although limited evidence is available, previous small studies have also not identified differences in outcome between MoM THAs and HRAs following ARMD revision [28, 195].

6.9.3. Clinical implications of findings

Early studies observed poor outcomes following MoMHA revision surgery performed for ARMD, which were highly concerning [12, 247, 249, 256]. Subsequently surgeons and worldwide authorities widely recommended early revision for ARMD, with a lower surgical threshold universally adopted [12, 23, 24, 29, 48]. However, the impact of early surgery for ARMD is unknown. This is important to establish given most MoMHAs currently remain in-situ in a young and active patient population [2, 11], and that many implants are likely to require future revision surgery given the high failure rates of most MoMHA designs [2, 11] and the regular patient surveillance recommended [22-24]. The observations in the present large nationwide study that MoMHA patients undergoing revision for ARMD have approximately half the risk of re-revision and death compared to matched patients undergoing non-ARMD revisions are clinically significant given that they are completely opposite to what was widely believed [12, 247, 249, 256]. It is suspected that worldwide regulatory authorities have positively influenced outcomes following ARMD revision by promoting regular patient surveillance and widely recommending that surgeons exercise a lower threshold for performing revision surgery [22-24].

The present work has provided surgeons with robust evidence, which can be used to informatively counsel patients about the risk associated with MoMHA revision surgery.

Patients and surgeons can both be somewhat reassured that revision of MoMHAs for ARMD is now relatively safe and successful compared with non-ARMD revision procedures. By contrast there were significant risks associated with MoMHA revision procedures performed for non-ARMD indications, especially when performed for infection and dislocation/subluxation. These observations are concerning and require further investigation. In the interim MoMHA patients undergoing revision procedures for infection and dislocation/subluxation must be appropriately counseled pre-operatively about the expected outcomes. It is also suggested these patients require regular surveillance following revision surgery to detect any complications early.

Surveillance bias of MoMHA patients and the universal adoption of a lower surgical threshold for performing ARMD revision surgery [12, 22-24, 29] are both suspected to have contributed to the increasing rate of ARMD revision surgery observed with extended follow-up (Chapter 3.6.1.). It is therefore feasible that some MoMHA revisions for ARMD may actually be performed too early, which has the potential for surgical risk to outweigh any benefits. Although further research is required to refine the thresholds for performing ARMD revision surgery, on the basis of the present study it is considered that the threshold for performing ARMD revision surgery need not be lowered further.

There is currently a paucity of information on prognostic factors of outcome following ARMD revision surgery [29, 195, 203]. The present large study identified important predictors of poor outcomes following ARMD revision, which included factors that can be modified intra-operatively. Although these predictors ideally require external validation, they do suggest that surgeons could reduce failure rates further following

ARMD revision. The higher re-revision rates observed with acetabular bone grafting may imply that acetabular osteolysis detected early in MoMHA patients represents a threshold for considering revision. This observation may alternatively imply that acetabular bone grafting is not a successful strategy for managing osteolysis following ARMD compared with other reconstructive methods, possibly because metal debris influences graft incorporation, and therefore could guide surgical decision making at revision. The type of revision reconstruction to perform in MoMHA patients with ARMD also influenced subsequent failure rates. When revising MoM THAs for ARMD, modular component only revisions should be avoided. In addition, revision of MoMHAs for ARMD should include exchange of the MoM bearing to a hard-on-soft articulation, preferably ceramic-on-polyethylene. Therefore the present study has provided surgeons with assistance regarding the type of revision procedure to perform and which bearing surface to implant when performing ARMD revision surgery for failed MoMHAs.

6.10. Summary

- Contrary to previous observations, analysis of a large NJR dataset demonstrated that MoMHA patients revised for ARMD have approximately half the risk of re-revision and death compared with matched patients undergoing non-ARMD revision surgery.

- The risk of re-revision surgery in unmatched MoMHA patients revised for ARMD was comparable to that reported in the NJR following all-cause conventional THA revision surgery.

- Four predictors of re-revision surgery following ARMD revision were identified: high BMI at revision, modular component only revisions, ceramic-on-ceramic revision bearings, and acetabular bone grafting. As these predictors included modifiable factors (type of revision procedure, revision bearing surface, and possibly the use of acetabular bone graft) surgeons may be able to reduce failure rates further following ARMD revision.

- It is suspected that surgeons and worldwide regulatory authorities have positively influenced outcomes following ARMD revision by promoting regular patient surveillance and widely recommending a lower threshold for performing revision surgery.

- The findings provide some reassurance that MoMHA revision surgery for ARMD is relatively safe and successful compared with non-ARMD revision surgery. However on the basis of the present work it is suggested that the threshold for performing ARMD revision surgery need not be lowered further.

- These findings provide surgeons with robust evidence, which can be used to informatively counsel patients about the risks associated with MoMHA revision surgery. Furthermore the findings may assist surgeons when making decisions about the type of reconstruction to perform when revising MoMHAs for ARMD.

- The high risk of re-revision surgery following MoMHA revisions performed for infection and dislocation/subluxation is concerning and requires further investigation.

7. CONCLUSIONS AND FUTURE WORK

7.1. Summary of the problem and thesis aims

Approximately 1.5 million MoMHAs were implanted worldwide largely in young and active patients. These devices have experienced unanticipated high short-term failure rates, and are now rarely used. Many MoMHAs have undergone revision surgery for ARMD and pseudotumour, however the clinical impact of this problem in the second decade after arthroplasty remains unknown. To identify problems early worldwide regulatory authorities currently recommend regular follow-up for most MoMHA patients. However these recommendations are not evidence-based and are very costly, with the specific roles of the main investigative modalities not yet established (Chapter 2.). The regular surveillance of many MoMHA patients may therefore be unnecessary. Although the prevalence of ARMD revision surgery is increasing, little remains known about outcomes following revision (Chapter 2.). Regulatory authorities and surgeons have widely recommended early revision surgery for ARMD given the poor short-term outcomes initially reported. However the impact of performing surgery early for ARMD is not known. Furthermore surgeons have no guidance available regarding the type of revision procedure to perform for ARMD due to a lack of evidence. Therefore currently there are many unanswered questions that are of high clinical importance when looking after MoMHA patients.

The broad aims of this thesis were to determine the clinical impact of pseudotumour in the second decade following arthroplasty, and to develop evidence-based guidance for the investigation and management of MoMHA patients (Appendix 25.).

7.2. Conclusions

The clinical impact of pseudotumour was assessed in a large prospective single centre study (Chapter 3.). This represents the only independent study assessing pseudotumour revision rates at 15 years in HRA patients. Pseudotumour was the commonest indication for HRA revision surgery and accounted for 61.7% of all revisions. The overall 15 year pseudotumour revision rate following HRA (14.0%) was high compared with all-cause revision rates reported in young patients following primary conventional THA (9.9%) [11]. Risk factors for pseudotumour revision surgery were gender specific. Young females (under 50 years at HRA) had the highest risk of pseudotumour revision (24.5% at ten years). In males the only predictor of pseudotumour revision was implantation with Conserve Plus (7.9% at nine years) and Recap designs (12.1% at nine years), with BHR implants having the lowest pseudotumour revision risk (4.6% at 15 years).

Given that pseudotumour revision rates are increasing into the second decade following HRA it is expected that many patients worldwide will require regular surveillance over future years. Therefore it is important to establish the specific roles of the various investigative modalities, namely blood metal ions and hip imaging.

Two large prospective cohort studies were undertaken to investigate the effectiveness of blood metal ion concentrations for identifying unilateral and bilateral HRA patients with ARMD (Chapter 4.). All patients had BHR implants, which represent the most commonly implanted HRA design worldwide. In the first single centre study (n=494), cobalt was the metal ion parameter with the highest discriminatory ability for

distinguishing between unilateral BHR patients with and without ARMD, with maximum cobalt or chromium having the highest discriminatory ability in bilateral BHRs. Patients with blood metal ions below the newly devised thresholds were at low risk of having ARMD (cobalt=2.15 μg/l for unilateral BHRs; maximum cobalt or chromium=5.52 μg/l for bilateral BHRs). However these new thresholds were not effective for identifying patients who had ARMD. Currently proposed fixed thresholds from the UK MHRA (7 μg/l) and US surgeons (3 μg/l and 10 μg/l) missed more patients with ARMD compared with the newly devised thresholds. The findings from this derivation study were successfully validated in a second study involving 416 BHR patients from two external European centres. Specifically patients in the validation study with blood metal ions below the newly devised thresholds were also at low risk of having ARMD, and fixed regulatory authority thresholds continued to miss more patients with ARMD compared with the newly devised thresholds in the validation cohort.

Three observational studies were undertaken to investigate the specific roles of various imaging modalities during HRA patient follow-up (Chapter 5.). A large retrospective case-control study identified a number of radiographic factors that were predictive of HRAs with evidence of a pseudotumour. These factors included acetabular component malposition, osteolysis, acetabular component loosening, and the absence of HO. Although radiographs were useful for distinguishing between HRAs with and without evidence of a pseudotumour, 20% of HRAs revised for pseudotumour had normal radiographs. Therefore radiographs should not be used as the only investigative modality during patient surveillance. A retrospective diagnostic accuracy study of 40 HRAs subsequently requiring revision surgery was performed to compare the cross-

sectional imaging findings prior to revision with the intra-operative findings. Ultrasound and MARS-MRI both had a role when assessing HRA patients prior to revision surgery. However combined imaging using ultrasound and MARS-MRI was most effective for both identifying and excluding intra-operative pseudotumours. These observations suggest that combined imaging may be useful in more complex clinical cases. A prospective longitudinal cohort study of 140 asymptomatic HRAs demonstrated that 11% of hips developed new pseudotumours on ultrasound examination within five years of initial assessment. No asymptomatic patients with a normal initial ultrasound and low blood metal ion concentrations (<2 μg/l) developed new pseudotumours within five years of initial assessment (36% of hips studied). These patients therefore do not need regular follow-up. Asymptomatic HRA patients with either a normal initial ultrasound scan alone, or low blood metal ion concentrations alone still had a small, but potentially clinically significant, risk of developing new pseudotumours during this time period. Therefore this patient subgroup requires some form of surveillance within five years of initial assessment.

Given the increasing prevalence of MoMHA revision surgery it is anticipated that a considerable proportion of patients will eventually undergo revision. Therefore it is important to consider the expected outcomes following revision surgery as well as identify prognostic factors of outcome. Two large retrospective observational studies were performed involving MoMHA revision procedures recorded in the NJR (Chapter 6.). This work represents the largest report of outcomes following MoMHA revision surgery. Contrary to previous observations [12], patients revised for ARMD had approximately half the risk of re-revision and death compared with matched patients undergoing non-ARMD revision surgery. In all revised MoMHAs the highest implant

survival rate was observed following ARMD revision (94.3% at five years), with revisions for infection (81.2% at five years) and dislocation (81.9% at five years) having the worst outcomes. The risk of re-revision surgery in unmatched MoMHA patients revised for ARMD (89.5% at five years) was comparable to that reported in the NJR following all-cause conventional THA revision surgery (87.8% to 89.1% at five years) [2]. Predictors of re-revision surgery following ARMD revision were high BMI at revision, modular component only revisions, ceramic-on-ceramic revision bearings, and acetabular bone grafting. Most predictors were factors that the surgeon can modify at the time of ARMD revision (type of revision procedure, revision bearing surface, and possibly acetabular bone graft).

Based on the work presented in this thesis it can be concluded that the clinical impact of pseudotumour following HRA is significant. The increasing revision rates observed in the second decade following arthroplasty are likely multifactorial with explanations including increased awareness of pseudotumours amongst surgeons, regular patient surveillance, and lowering of the revision threshold. This thesis subsequently developed an evidence-based approach for the investigation and management of MoMHA patients, which provides clinicians with important information when looking after these patients. No single investigative modality can universally be relied upon during patient surveillance. However important information has now been provided for risk stratifying patients for surveillance, interpreting blood metal ion concentrations, clarifying the specific roles of various imaging modalities, and providing evidence-based parameters that would allow many asymptomatic HRA patients to be safely excluded from regular follow-up. Furthermore surgeons now have robust evidence available regarding the risks associated with MoMHA revision surgery, which can be used to informatively

counsel patients prior to undergoing these procedures. The outcomes following ARMD revision surgery have improved since the threshold for performing revision was universally lowered. However as outcomes were influenced by modifiable factors it is suspected that surgeons could reduce failure rates further following ARMD revision.

7.3. Clinical recommendations

The high pseudotumour revision rates that continue to increase into the second decade following HRA highlight the importance of these patients undergoing some form of surveillance. Numerous protocols have been proposed for patient surveillance over the past seven years [22-24, 67, 112, 141, 142, 184]. Although current follow-up recommendations by worldwide regulatory authorities are not evidence-based (Chapter 2.5.), it is clear that any proposed patient surveillance algorithms will not be simple.

An algorithm has been developed for the investigation and management of HRA patients (Figure 7.1.). This algorithm is predominantly based on the work described within this thesis and studies published by other authors (Chapter 2.). However some assumptions have been made given that a number of clinical questions remain unanswered (Chapter 7.5.).

It is recommended that all HRA patients undergo a complete clinical assessment as a baseline. This assessment should include blood metal ion concentrations, cross-sectional imaging, and pelvic radiographs. The choice of cross-sectional imaging modality (ultrasound or MARS-MRI) used will depend on the availability of ultrasound expertise at the respective institution as well as local financial constraints.

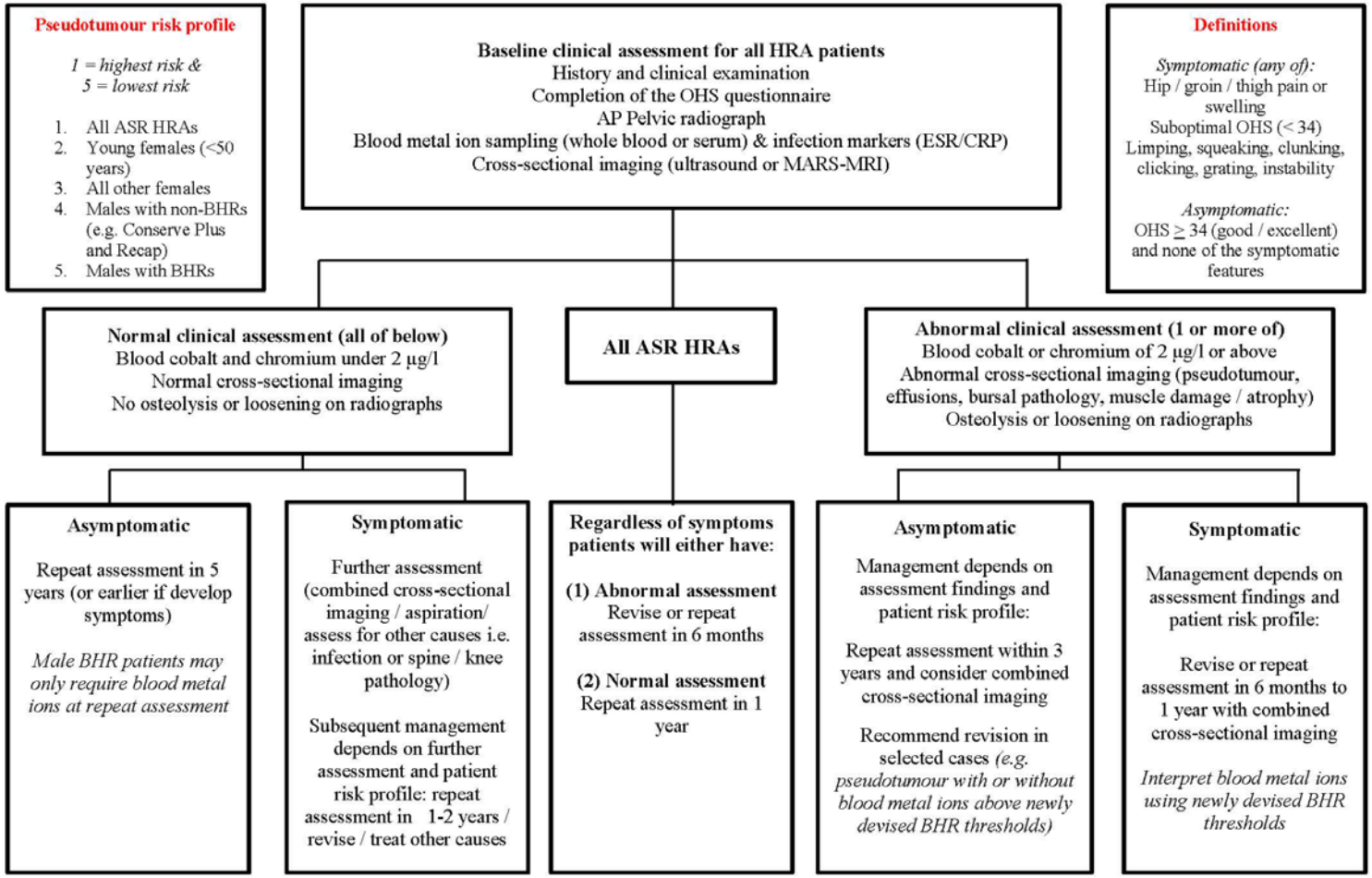


Figure 7.1 Algorithm for the investigation and management of hip resurfacing arthroplasty patients.

Subsequent patient management will depend on a combination of the results of the complete clinical assessment, hip symptoms, and the patient's pseudotumour risk (Chapter 3.). Patients with ASR HRA require regular surveillance (6 to 12 monthly) regardless of symptoms, which is in line with current worldwide recommendations [22-24, 141, 142]. The remaining patients with non-ASR HRAs can be divided into four subgroups based on the assessment results (normal or abnormal) and patient symptoms (symptomatic or asymptomatic). The follow-up guidance for these four subgroups has previously been detailed (Chapter 5.12.4.) and summarised (Figure 7.1.). Importantly there will be a large group of HRA patients that are asymptomatic with normal initial cross-sectional imaging and low blood metal ion concentrations (<2 µg/l), who can safely be excluded from regular surveillance for five years. This will result in considerable financial savings (Chapter 5.12.4.), which can subsequently be focussed towards HRA patients that do require further surveillance. Although it is recommended that such patients would require a complete clinical assessment at five years if they remain asymptomatic, it is worth considering less intensive surveillance in male BHR patients who remain asymptomatic given the low risk of pseudotumour revision at up to 15 years follow-up (Chapter 3.). Therefore it is proposed that male BHR patients undergo blood metal ion sampling at five years following initial assessment, with the ion results interpreted in relation to the respective unilateral and bilateral BHR thresholds (Chapter 4.). However it is unclear whether this particular patient subgroup could safely be discharged from all forms of surveillance after five years if they remain asymptomatic.

It is worth considering some specific aspects of the proposed follow-up algorithm for HRA patients in more detail (Figure 7.1.). When an abnormality has been identified it must be interpreted in relation to the pseudotumour risk profile of the patient (Chapter 3.). For example a HRA with a large effusion on cross-sectional imaging needs more regular surveillance in female patients below 50 years of age (high pseudotumour risk) compared with male BHR patients (low pseudotumour risk). Blood metal ion samples must be measured at an accredited laboratory [269]. A blood metal ion threshold of 2 μg/l has been used to stratify patients into those with normal or abnormal assessments based on the longitudinal study presented in this thesis (Chapter 5.10.). However it is recommended that the newly devised BHR thresholds (Chapter 4.) are used when interpreting blood metal ion results in patients with abnormal investigations, given that these thresholds missed fewer patients with ARMD compared with currently proposed fixed regulatory authority thresholds. Combined cross-sectional imaging should be considered in more complex clinical cases where there remains a diagnostic dilemma following the initial assessment. Such cases include symptomatic patients with a normal initial ultrasound or MARS-MRI regardless of the blood metal ion concentrations, and in asymptomatic patients with blood metal ion concentrations above the newly devised BHR thresholds but who have normal initial cross-sectional imaging. Although some recommendations have been made regarding when to consider revision surgery for HRA patients it is important to acknowledge that these have been proposed based on the existing literature (Chapter 2.) given that this thesis did not formally investigate thresholds for performing revision surgery for ARMD.

The findings from the large observational studies involving MoMHA revision procedures recorded in the NJR have significant clinical importance (Chapter 6.).

Surgeons now have robust evidence available regarding the risks associated with MoMHA revision surgery. This information can be used to counsel patients prior to undergoing these procedures. For patients undergoing revision surgery for ARMD the findings provide some reassurance that these procedures are now relatively safe and successful compared with non-ARMD revision surgery, especially given the catastrophic outcomes initially reported following early ARMD revisions [12]. It is suspected that surgeons and worldwide regulatory authorities have positively influenced outcomes following ARMD revision by promoting regular patient surveillance and widely recommending a lower threshold for performing revision surgery. However now that the outcomes following ARMD revision are better than, or at the very least comparable with, non-ARMD revisions it is suggested that the threshold for performing ARMD revision surgery need not be lowered further. Lowering of the ARMD revision threshold any further introduces the potential for surgical risk to outweigh any clinical benefits. By contrast the high risk of re-revision surgery following MoMHA revisions performed for infection and dislocation is concerning and requires further investigation. Surgeons should carefully counsel MoMHA patients likely to require revision for infection and dislocation preoperatively about the expected outcomes. It is also recommended that these patients undergo regular surveillance following revision surgery to detect any complications early.

The outcomes following ARMD revision surgery in MoMHA patients were influenced by a number of modifiable factors within the surgeon's control (avoid modular component only revisions, ceramic-on-ceramic revision bearings, and possibly acetabular bone grafting). In light of these observations surgeons could potentially reduce failure rates further following ARMD revision procedures. Therefore this thesis

has provided surgeons with important information, which can be used to guide decisions about the type of reconstruction to perform when revising MoMHAs for ARMD.

7.4. Limitations

Specific limitations related to the work performed within this thesis have previously been detailed (Chapters 3. to 6.). However some limitations apply to a number of the studies presented within this thesis, which require further consideration. The study assessing the clinical impact of pseudotumours (Chapter 3.) and the studies investigating the roles of blood metal ions and hip imaging (Chapters 4. and 5.) were based on HRA patients. Many MoM THAs were implanted worldwide [2, 8, 11]. However during the time period in which this thesis was undertaken it was not possible to access the relevant clinical data for large cohorts of MoM THA patients and repeat all of the studies performed in HRA patients in MoM THAs. Given the obvious design differences between HRA patients and MoM THAs, namely the one or more additional modular THA junctions, the findings presented in this thesis relating to HRA patients cannot be assumed to directly apply to MoM THA patients.

Another important limitation relates to the terminology used for describing the unanticipated mode of failure observed in MoMHAs (Chapter 2.2.1.). The term ARMD has generally been accepted as the most inclusive for describing this heterogeneous clinical problem. However pseudotumour refers to a specific type of ARMD. Pseudotumour was a term coined by surgeons in Oxford [16], and is defined as a cystic, solid, or mixed mass communicating with the hip joint. By definition all hips with a pseudotumour can be classified as having ARMD, however the converse is not

necessarily true. Despite this it is reasonable on clinical grounds to separate hips with pseudotumours from the generic diagnosis of ARMD given that pseudotumours can be large, destructive, and invasive lesions. Such lesions are very different to hips with metallosis and synovitis alone that also fulfil a diagnosis of ARMD [12, 14, 16, 195, 203]. There is variability in the literature regarding the proportion of MoMHAs revised with ARMD who actually have a pseudotumour. In Oxford almost all such hip revisions have been performed for pseudotumour (Chapters 3. and 5.), whilst in other specialist arthroplasty centres the proportion of ARMD revisions with a pseudotumour has been less common [29, 195, 203]. This variance between centres may relate to different pathologies or variability in the way clinicians have chosen to classify these abnormalities. In the experimental chapters of this thesis the prospective classification of these abnormalities by surgeons at the various institutions has largely been respected. Therefore some chapters within this thesis have referred to ARMD whilst others referred to pseudotumour.

Finally, although the newly devised blood metal ion thresholds in BHR patients were appropriately validated (Chapter 4.) the other clinical studies presented in this thesis did not undergo any formal external validation. This must be taken into account when considering any changes to clinical practice.

7.5. Recommendations for future work

It is anticipated that the work presented within this thesis will make a substantial contribution to the literature regarding the investigation and management of MoMHA patients. However there are some important clinical questions that remain unanswered.

These specific areas are considered subsequently, and it is hoped that these questions can be addressed in future studies.

It is important to determine the clinical impact of ARMD and pseudotumour in MoM THA patients at extended follow-up given that this thesis reported on HRA patients (Chapter 3.). A recent study based on 489 Pinnacle MoM THAs, which represents one of the most commonly used MoM THA designs worldwide, observed a high revision rate at nine years (16.4%) with all but one revision performed for ARMD [102]. During the thesis a similar study was performed on 569 Pinnacle MoM THAs (with a Corail stem). This study also identified high revision rates with this prosthesis (20.3% at ten years) with most revisions performed for ARMD [135]. However this latter piece of work could not be included within this thesis due to space constraints. It would also be useful to perform similar studies in other commonly implanted MoM THA designs.

The newly devised and validated blood metal ion thresholds for identifying patients at low risk of ARMD are strictly only applicable to patients with unilateral and bilateral BHR implants (Chapter 4.). Furthermore a threshold for ASR HRAs has previously been devised [168]. Therefore it would be beneficial for future studies to devise and validate similar thresholds for patients with other commonly used HRA implants [2, 11, 66], such as Conserve Plus and Recap designs, given that these particular designs represent predictors of pseudotumour revision in male patients (Chapter 3.). Future studies should also devise and validate similar thresholds for MoM THA patients given the clear differences in design (one or more modular THA junctions) and failure mechanisms (corrosion versus bearing wear) compared with HRAs. During the thesis blood metal ion thresholds were devised and successfully validated for patients with a

unilateral Corail Pinnacle MoM THA (cobalt=3.57 µg/l) [324], with the findings from these MoM THAs paralleling the observations in BHR patients (Chapter 4.). However this work could not be included within this thesis due to space constraints.

The specific roles of various imaging modalities during patient follow-up were investigated in HRA patients (Chapter 5.). These studies should be repeated in MoM THA patients, particularly the longitudinal follow-up study (Chapter 5.9.) given that this would help inform surveillance in this group of patients. Although it is assumed that MARS-MRI could be substituted for ultrasound in the follow-up of asymptomatic HRA patients this should be formally assessed in a longitudinal study at different institutions. In the longitudinal follow-up study presented in this thesis patients did not undergo repeat pelvic radiographs and repeat blood metal ion sampling. Therefore on the basis of this thesis firm recommendations cannot be made about the roles of these investigations in subsequent patient surveillance. However a longitudinal study at extended follow-up is planned with ethical approval now granted. Patients participating in the longitudinal study presented in this thesis (Chapter 5.9.) will soon be invited to participate in this new study. Consenting HRA patients will undergo radiographs and blood metal ion testing in addition to ultrasound and completion of an OHS and UCLA score questionnaire at a minimum of ten years from the initial assessment. This study will therefore be able to provide further informative guidance regarding the surveillance of initially asymptomatic HRA patients at extended follow-up.

The large studies assessing outcomes following MoMHA revision surgery were based on observational data from the NJR (Chapter 6.). Therefore the outcomes reported in this thesis were limited by the data collected by the registry. A large multi-centre study

involving three European centres was commenced during the thesis time period to determine the outcomes following MoMHA revision surgery performed for ARMD in a minimum of 400 patients. Given the large scale of this study data collection is still underway. This multi-centre study will report on other important outcomes following ARMD revision surgery that are not available within the NJR, including patient reported outcome measures and the risk of further surgery excluding re-revision (such as closed reductions for dislocations, wound debridement for infections, and internal fixation for peri-prosthetic fractures). The other advantages of the multi-centre study over the NJR cohorts presented in this thesis include the longer follow-up period after ARMD revision (mean five years vs. three years) and the assessment of other potentially important predictors of re-revision surgery, such as specific MoMHA implant design and the pre-revision blood metal ions and cross-sectional imaging findings. It is hoped that the data from the multi-centre study could be used to develop a formal clinical decision-making tool for both the threshold (when to recommend revision) and type of revision surgery (which components require revision, and which bearing surface/fixation methods to use) required in MoMHA patients with ARMD. This tool would be developed using similar methods to that described in the thesis (Appendix 12.). It will also be possible to validate the predictors of re-revision surgery identified in the NJR study (Chapter 6.4.4.) using the multi-centre cohort.

External validation is an important process to confirm or refute the findings observed in research studies [288, 289]. It is therefore advisable for most of the findings from this thesis (Chapters 3., 5. and 6.) to undergo formal validation in external patient cohorts at other worldwide institutions before any significant changes to clinical practice are implemented. Furthermore as evidence becomes increasingly available regarding the

investigation and management of MoMHA patients, it is important to analyse the cost-effectiveness of different patient surveillance options, as well as for the proposed thresholds for revision surgery when these are established.

Finally it is worth acknowledging that more recently there have been reports of conventional THAs requiring revision surgery for ARMD [86, 87, 325]. The risk posed by this emerging problem was investigated during the thesis using data from the NJR. Although the risk of ARMD revision surgery in all primary conventional THAs recorded in the NJR was low (0.032%) this risk appears to be increasing steadily over time [326]. Furthermore the risk of ARMD revision was significantly higher in ceramic-on-ceramic THAs (compared with hard-on-soft bearing THAs), and in 36 mm metal-on-polyethylene THAs (compared with 28 mm and 32 mm metal-on-polyethylene THAs). It is therefore recommended that further research is undertaken to determine the aetiology and pathogenesis of ARMD associated with conventional THAs. It is also important to establish whether certain implant designs or subgroups of patients require more regular surveillance with blood metal ions and cross-sectional imaging, with some interim guidance for managing these patients currently available [327].

Appendix 1. Methods for blood metal ion sampling

Whole blood or serum samples are used for measuring metal ion concentrations. Samples are collected from peripheral veins into trace element tubes containing an anti-coagulant (usually sodium EDTA). Needles and collection tubes must have no evidence of trace metal contamination [268]. Samples should be collected, stored, transported, and processed according to laboratory guidance to prevent contamination. Analysis must be performed at an accredited laboratory with expertise in trace metal element analysis. Such laboratories take part in regular audits of their analytical methods, with excellent accuracy of measurement and reproducibility reported [269]. Serum samples require centrifuging, but whole blood samples do not, which represents a potential source for contamination [158]. Analysis of samples is performed using dynamic reaction cell or collision cell inductively coupled plasma mass spectrometry.

Blood metal ion concentrations have been reported with different units of measure. Micrograms per litre (μg/l) is commonly used, and is directly equivalent to parts per billion (ppb) and nanograms per millilitre (ng/ml). Nanomoles per litre (nmol/l) is based on the atomic weight of cobalt (58.93) and chromium (52.00). Thus nmol/l can be obtained by multiplying μg/l values by 16.97 (cobalt) and 19.23 (chromium), with μg/l obtained by multiplying nmol/l values by 0.0589 (cobalt) and 0.0520 (chromium).

Appendix 2. Outcomes following ARMD revision surgery: review methodology and study quality

A systematic review undertaken in 2013 by Matharu *et al*. identified six eligible studies reporting the outcomes following 216 MoMHA revisions performed for ARMD [28]. This systematic review was updated in December 2016 to identify any additional relevant publications reporting the outcomes of more than ten HRA and/or more than ten MoM THA revision procedures performed for ARMD. This criterion was used to eliminate small case series that may have been influenced by learning curves, and to minimise any selection and publication bias [28]. Outcomes of interest following ARMD revision surgery were the risk of all complications (intra-operative and post-revision), the risk of re-revision surgery, and post-revision functional outcomes.

Fifteen unique studies were eligible for inclusion, which comprised a total of 803 MoMHA revision procedures performed for ARMD, with a further study reporting the outcomes of an initial cohort at extended follow-up (Chapter 2.4.3.1.). All included studies were either case-control or cohort studies, with the majority being retrospective. Using the Oxford Centre for evidence-based medicine levels of evidence, all included studies were graded as level 4 evidence [244]. It is acknowledged that a number of registry reports of outcomes following MoMHA revision surgery exist [243, 263, 264]. However these studies were not considered subsequently as they did not stratify outcomes following MoMHA revision by the failure indication.

A number of issues were identified with the included studies, such as small sample size, data quality problems and missing data, which can all make synthesis of the available evidence problematic. Ten of the 15 unique studies involved small cohorts of less than

50 ARMD revisions. Apart from the extended follow-up report [256] all studies had a mean follow-up after ARMD revision of 5.1 years or shorter, with most having a mean follow-up of 3 years or less. A number of studies did not specifically provide patient demographics and outcomes for the ARMD revision procedures, but rather for the cohort of MoMHA revisions performed for all indications. Some studies reported on complications and re-revision surgery but not on post-revision functional outcomes, or vice-versa [251, 253]. Studies that did report functional outcomes used a number of different instruments. The Oxford Hip Score (OHS) [257, 258] was most commonly used, but others included the Harris Hip Score (HSS) [259], the Hip Disability and Osteoarthritis Outcome Score (HOOS) [260], and the Western Ontario and McMaster Universities Arthritis Index (WOMAC) [261].

Appendix 3. Methodology and limitations for financial analysis of regulatory authority patient surveillance

Methodology

Using MoMHA follow-up guidance from worldwide authorities (Table 2.5.) a financial analysis was performed for four common clinical scenarios. Patients with high-risk MoMHA devices and/or ARMD risk factors were not considered separately because follow-up regimens, and therefore costs, generally paralleled those for symptomatic MoMHAs (Scenario 1). Follow-up costs have been described previously [161]. The total cost for each scenario represents the annual follow-up of one patient and have been presented in the table below.

To more accurately estimate the financial impact of follow-up, NJR data was used which included 67,363 MoMHAs (35,470 HRA and 31,893 MoM THA) implanted in England and Wales [2]. This cohort was stratified by symptoms using data from Hart *et al*, where 597 unilateral ASR MoMHAs all underwent blood metal ion sampling and cross-sectional imaging, along with OHS questionnaires [168]. Within seven years of primary surgery 40% of MoM THAs and 16% of HRAs had failed [168]. These proportions for failed MoMHAs were applied to the NJR data, and along with the original follow-up costs per patient it was possible to estimate the financial implications of surveillance for all MoMHAs recorded within the NJR when different regulatory authority guidance was applied (Table 2.6.).

	MHRA UK	**European**	**FDA US**	**TGA Australia**	**Health Canada**
Scenario 1 Large-diameter MoM THA (≥36mm) Symptomatic	Clinical review, metal ions, MRI £330	Clinical review, x-ray, metal ions, MRI £494	Clinical review, x-ray, metal ions, MRI x 2* £988	Clinical review, x-ray, metal ions, MRI £494	Clinical review, x-ray, metal ions, MRI £494
Scenario 2 Large-diameter MoM THA (≥36mm) Asymptomatic	Clinical review, metal ions £114	Clinical review, x-ray, metal ions £278	Clinical review £84	Clinical review, x-ray, metal ions, MRI £494	Clinical review £84
Scenario 3 HRA Symptomatic	Clinical review, metal ions, MRI £330	Clinical review, x-ray, metal ions, MRI £494	Clinical review, x-ray, metal ions, MRI x 2* £988	Clinical review, x-ray, metal ions, MRI £494	Clinical review, x-ray, metal ions, MRI £494
Scenario 4 HRA Asymptomatic	Local protocol £0	Clinical review, x-ray, metal ions £278	Clinical review £84	Local protocol £0	Clinical review £84

Table A3: Annual follow up costs per patient using different surveillance protocols.

FDA = Food and Drug Administration; HRA = hip resurfacing arthroplasty; MRI = magnetic resonance imaging; MHRA = Medical and Healthcare products Regulatory Agency; MoM = metal-on-metal; TGA = Therapeutic Goods Administration; THA = total hip arthroplasty; UK = United Kingdom; US = United States.

All scenarios assume patients have unilateral metal-on-metal hip implants, the device has not been recalled or considered high-risk, and the implant has an acceptable radiographic appearance.

Costs were taken from a previous publication [161]: outpatient follow-up appointment = £84, radiographs (anteroposterior pelvis and lateral hip) = £164, magnetic resonance imaging = £216, blood metal ion sampling = £30.

Magnetic resonance imaging was chosen for this example rather than ultrasound (£49).

*** 6 monthly follow-up recommended for symptomatic patients in the US, therefore costs were doubled to provide annual costs.**

The cost of local protocol follow-up has been taken to be zero for the purposes of this analysis (at this institution all asymptomatic patients with low risk hip resurfacing devices and no adverse reaction to metal debris risk factors were discharged).

Limitations

The financial analysis performed has recognised limitations. It did not consider patients undergoing revision surgery who would have been removed from surveillance, as this would have been difficult to integrate into the model. To minimise this, costs for only one year of follow-up were assessed. Furthermore, geographical variations in cost exist for each aspect of follow-up. Although the estimates may change slightly, by comparing guidance using fixed costs the relationships highlighted between protocols are suspected to be accurate. Finally, it would have been preferable to extend the analysis to include implant types other than the ASR. Unfortunately this data was not available for large cohorts with all MoMHA designs, with registry data not stratifying patients according to symptoms [2]. Therefore this analysis represented a worst-case scenario given the ASR represents the poorest performing MoMHA design.

Appendix 4. Data collection form used at the time of hip resurfacing arthroplasty surgery

PRIMARY HIP RESURFACING

LOCATION OF OPERATION (Please tick appropriate hospital)

NUFFIELD ORTHOPAEDIC CENTRE		MANOR HOSPITAL	
OXFORD CLINIC:			

PATIENT DATA LABEL:

SIDE: Left ☐ Right ☐ Bilateral Simult ☐

DIAGNOSIS: OA ☐ Dysplasia ☐ AVN ☐
RA ☐ other_____

SURGEON GRADE
ASSIST GRADE

Date of Operation: (DD) ___ (MM) _____ (YY) _____ Funding: NHS ☐ PP ☐

Approach: Extended Posterior ☐ Posterior ☐ Antero-lateral ☐ Other _____

Suction Vent Used: Yes ☐ No ☐

Technique for guide wire insertion: Eyeball ☐ Amstuz ☐ McMinn ☐ Lollipop ☐ Claw ☐ Other ☐

PROSTHESIS LABEL ACETABULUM	PROSTHESIS LABEL FEMUR	CEMENT LABEL
If no stickers available please write type and size eg Recap/Conserve etc	If no stickers available please write type and size eg Recap/Conserve etc	

Weight Bearing Plan: 1st 6 weeks FWB ☐ 20 kg ☐ None ☐
2nd 6 weeks FWB ☐ 20 kg ☐ None ☐

Technical problems:	Intra-operative complications:

Comments:

1. RSA study: Yes ☐ No ☐

2. Blood Flow Study: Yes ☐ No ☐

Appendix 5. Postal questionnaire sent to hip resurfacing arthroplasty patients

PATIENT OUTCOME ASSESSMENT
METAL-ON-METAL-RESURFACING HIP REPLACEMENT –

	Date of birth: **Hospital number:** **Date of operation:** **Side:**

1. Have you had any operations on your **«Side»** hip since your **«Side»** hip resurfacing operation on **«Date_of_Surgery»**? (Please tick box)

Yes ❑ **No** ❑ **→If 'no' go to question 2 below**

If yes: Which operation(s) have you had on the **«Side»** hip?

(Please tick boxes)

a. Treatment of dislocation (hip came out of joint) ❑
b. Treatment of fracture (broken bone) ❑
c. Redo of «Side»hip resurfacing ❑
d. Another («Side») hip operation ❑

At which hospital(s) did you have this operation(s)?

NAME OF HOSPITAL(S)/TOWN	WHEN? (YEAR)	**«Side» hip** TYPE OF OPERATION

2. Do you have any other problems - apart from your **«Side»** hip - which make everyday activities (eg. walking, washing, dressing) difficult?

No ❑ **Yes** ❑ - If 'Yes', Please specify: .....

3. Overall, how pleased have you been with your **«Side»** resurfacing hip replacement?

Very pleased	Fairly pleased	Not very pleased	Very disappointed
❑	❑	❑	❑

4. Following your **«Side»** resurfacing hip replacement operation, for how many years were you pain free?

«Side» hip still pain free ❑ OR: **(please tick box)** **Never pain free** ❑

Number of years **.....pain free**

"Clicking" and your resurfacing hip replacement

Please circle the appropriate answer

1. Did your hip "click" before your hip resurfacing operation?

   Yes / No

2. Have you **ever** experienced a feeling of "clicking" in your resurfacing hip replacement?

   Yes / No

3. Does your hip replacement **currently** "click"? Yes / No

If you have answered YES to question 3, does this "clicking" cause you any problem?

Yes / No

If yes, please give details.

_____

_____

_____

_____

_____

Appendix 6. Oxford Hip Score questionnaire completed by hip resurfacing arthroplasty patients

OXFORD HIP QUESTIONNAIRE

Date: ☐☐/☐☐/☐☐

Op date: : ☐☐/☐☐/☐☐

✓tick one box for every question.

PROBLEMS WITH YOUR LEFT/RIGHT HIP

1. *During the past 4 weeks.....*

How would you describe the pain you usually have from your hip?

None	Very mild	Mild	Moderate	Severe
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

2. *During the past 4 weeks.....*

Have you had any trouble with washing and drying yourself (all over) because of your hip?

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

3. *During the past 4 weeks .....*

Have you had any trouble getting in and out of a car or using public transport because of your hip? *(whichever you tend to use)*

No trouble at all	Very little trouble	Moderate trouble	Extreme difficulty	Impossible to do
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

4. *During the past 4 weeks.....*

Have you been able to put on a pair of socks, stockings or tights?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

5. *During the past 4 weeks.....*

Could you do the household shopping on your own?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

6. *During the past 4 weeks.....*

For how long have you been able to walk before pain from your hip becomes severe? *(with or without a stick)*

No pain/ More than 30 minutes	16 to 30 minutes	5 to 15 minutes	Around the house only	Not at all -pain severe on walking
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

LEFT / RIGHT HIP — ✓tick one box for every question

7 *During the past 4 weeks.....*

Have you been able to climb a flight of stairs?

Yes, Easily	With little difficulty	With moderate difficulty	With extreme difficulty	No, Impossible
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

8 *During the past 4 weeks.....*

After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your hip?

Not at all painful	Slightly painful	Moderately painful	Very painful	Unbearable
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

9 *During the past 4 weeks.....*

Have you been limping when walking, because of your hip?

Rarely/ never	Sometimes, or just at first	Often, not just at first	Most of the time	All of the time
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

10 *During the past 4 weeks.....*

Have you had any sudden, severe pain - 'shooting', 'stabbing' or 'spasms' - from the affected hip?

No days	Only 1 or 2 days	Some days	Most days	Every day
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

11 *During the past 4 weeks.....*

How much has pain from your hip interfered with your usual work *(including housework)*?

Not at all	A little bit	Moderately	Greatly	Totally
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

12 *During the past 4 weeks.....*

Have you been troubled by pain from your hip in bed at night?

No nights	Only 1 or 2 nights	Some nights	Most nights	Every night
☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

Appendix 7. Patient and implant factors for non-revised hip resurfacing arthroplasty patients completing the follow-up questionnaire compared with those not completing the questionnaire

	All non-revised hips (n=1,249)	Hips completing questionnaire (n=1,185; 94.9%)	Hips not completing questionnaire (n=64; 5.1%)	p-value
Gender				
Male	780 (62.5)	750 (63.3)	30 (46.9)	**0.008**
Female	469 (37.5)	435 (36.7)	34 (53.1)	
Age (years)	53.9	54.0	52.4	0.219
Mean (range)	(16.5-85.5)	(16.5-85.5)	(29.7-69.7)	
Laterality				
Unilateral	870 (69.7)	830 (70.0)	40 (62.5)	0.201
Bilateral	379 (30.3)	355 (30.0)	24 (37.5)	
Hip diagnosis				
Primary OA	1,096 (87.8)	1,037 (87.5)	59 (92.2)	0.242
Dysplasia	107 (8.6)	105 (8.9)	2 (3.1)	
Other	46 (3.7)	43 (3.6)	3 (4.7)	
Femoral head size (mm)	48.7	48.7	48.2	0.355
Mean (range)	(38-58)	(38-58)	(42-58)	
Implant				
BHR	563 (45.1)	533 (45.0)	30 (46.0)	**<0.001**
Conserve Plus	551 (44.1)	525 (44.3)	26 (40.6)	
Recap	123 (9.9)	120 (10.1)	3 (4.7)	
Cormet	12 (0.96)	7 (0.59)	5 (7.8)	

BHR = Birmingham Hip Resurfacing; OA = Osteoarthritis. P-values <0.05 in bold text. Values in brackets are percentages unless stated otherwise in the left hand column. Patients not completing the questionnaire were significantly more likely to be female and have Cormet implants (however very few Cormet designs were used).

Appendix 8. NJR data collection form for primary hip arthroplasty

National Joint Registry www.njrcentre.org.uk	MDS VERSION 4.0 Hip Operation Form: MDSv4.0 H1 v1.0
H1 Hip Primary	Patient Addressograph
Important: Please tick relevant boxes. All component stickers should be affixed to the accompanying 'Minimum Dataset Form Component Labels Sheet'. Please ensure that all sheets are stapled together.	

All fields are Mandatory unless otherwise indicated

REMEMBER! MAKE A NOTE OF THE NJR REFERENCE NUMBER WHEN YOU ENTER THIS DATA	NJR REF:

PATIENT DETAILS				
Patient Consent Obtained	Yes ☐	No ☐	Not Recorded ☐	
Patient Hospital ID				
Body Mass Index (enter either H&W OR BMI OR tick Not Available box)	Height (IN CM) Weight (IN KG)	BMI	Not Available ☐	

PATIENT IDENTIFIERS				
Forename				
Surname				
Gender	Male ☐	Female ☐	Not Known ☐	Not Specified ☐
Date of Birth	DD/MM/YYYY			
Patient Postcode			Overseas Address ☐	
NHS Number (if available)				

OPERATION DETAILS					
Hospital					
Operation Date	DD/MM/YYYY				
Anaesthetic Types	General ☐ Regional - Epidural ☐		Regional – Nerve Block ☐ Regional – Spinal (Intrathecal) ☐		
Patient ASA Grade	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
Operation Funding	NHS ☐	Independent ☐			

SURGEON DETAILS					
Consultant in Charge					
Operating Surgeon					
Operating Surgeon Grade	Consultant ☐	SPR/ST3-8 ☐	F1-ST2 ☐	Specialty Doctor/SAS ☐	Other ☐
First Assistant Grade	Consultant ☐	Other ☐			

HIP PRIMARY PROCEDURE DETAILS						
Side	Left ☐	Right ☐				
Indications for Implantation (select all that apply)	Osteoarthritis	☐	Trauma – Chronic	☐		
	Inflammatory Arthropathy	☐	Previous Hip Surgery – non Trauma related	☐		
	Congenital Dislocation / Dysplasia of the Hip	☐				
	Avascular Necrosis	☐	Previous Arthrodesis	☐		
	Trauma – Acute (Neck of Femur)	☐	Previous Infection	☐		
	Failed Hemi-Arthroplasty	☐	Other	☐		

SURGICAL APPROACH						
Patient Procedure	Primary Total Prosthetic Replacement Using Cement ☐ Primary Total Prosthetic Replacement Not Using Cement ☐ Primary Resurfacing Arthroplasty of Joint ☐ Primary Total Prosthetic Replacement Not Classified Elsewhere (eg Hybrid) ☐					
Consultant in Charge – Default Technique used?	Yes ☐ No ☐ If Yes, ensure the relevant Surgeon Default Technique is recorded on the Data Entry system. The Surgeon's Default Technique is made up of several data fields.					
Patient Position	Lateral ☐	Supine ☐				
Approach	Hardinge	☐	Trochanteric Osteotomy	☐		
	Posterior	☐	Other	☐		
Minimally Invasive Technique Used?	Yes ☐	No ☐				
Computer Guided Surgery Used?	Yes ☐	No ☐				
THROMBOPROPHYLAXIS REGIME (intention to treat)						
Chemical	Aspirin	☐	Warfarin	☐	None	☐
	LMWH	☐	Direct Thrombin Inhibitor	☐		
	Pentasaccharide	☐	Other	☐		
Mechanical	Foot Pump	☐	Other	☐		
	Intermittent Calf Compression	☐	None	☐		
	TED Stockings	☐				
BONEGRAFT USED						
Femur	Yes ☐	No ☐				
Acetabulum	Yes ☐	No ☐				
SURGEON'S NOTES						

INTRA OPERATIVE EVENT						
Untoward Intra Operative Event	None	☐	Shaft Fracture	☐	Other	☐
	Calcar Crack	☐	Shaft Penetration	☐		
	Pelvic Penetration	☐	Trochanteric Fracture	☐		

Minimum Dataset Form - COMPONENT LABELS

1. Please affix any component labels to this sheet and ensure any extra component label sheets are attached to the main Minimum Dataset Form.
2. Ensure all component details are provided, including cement.
3. The NJR DOES NOT record the following: wire, mesh, cables, plates, screws, surgical tools, endoprostheses or bipolar heads.

Appendix 9. NJR data collection form for revision hip arthroplasty

National Joint Registry www.njrcentre.org.uk Working for patients, driving forward quality	MDS VERSION 6.0 Hip Operation Form: MDSv6.0 H2 v004
H2 **Hip Single Stage Revision** **Hip Stage 1 of 2 Stage Revision** **Hip Stage 2 of 2 Stage Revision** **Hip Excision Arthroplasty** **Insertion of PLAD/Stabiliser**	Patient Addressograph
Important: Please tick relevant boxes. All component stickers should be affixed to the accompanying 'Minimum Dataset Form Component Labels Sheet'. Please ensure that all sheets are stapled together.	

All fields are Mandatory unless otherwise indicated

REMEMBER! MAKE A NOTE OF THE NJR REFERENCE NUMBER WHEN YOU ENTER THIS DATA	NJR REF:

PATIENT DETAILS				
NJR Patient Consent Obtained	Yes ☐	No ☐	Not Recorded ☐	
Patient Hospital ID				
Body Mass Index (enter either H&W OR BMI OR tick Not Available box)	Height (IN M) Weight (IN KG)	BMI	Not Available ☐	

PATIENT IDENTIFIERS				
Forename				
Surname				
Gender	Male ☐	Female ☐	Not Known ☐	Not Specified ☐
Date of Birth	DD/MM/YYYY			
Patient Postcode			Overseas Address ☐	
NHS or National Patient Number (if available)				

OPERATION DETAILS				
Hospital				
Operation Date	DD/MM/YYYY			
Anaesthetic Types	General ☐ Regional - Epidural ☐		Regional – Nerve Block ☐ Regional – Spinal (Intrathecal) ☐	
Patient ASA Grade	1 ☐	2 ☐	3 ☐	4 ☐ 5 ☐
Operation Funding	NHS ☐	Independent ☐		

SURGEON DETAILS					
Consultant in Charge					
Operating Surgeon					
Operating Surgeon Grade	Consultant ☐	SPR/ST3-8 ☐	F1-ST2 ☐	Specialty Doctor/SAS ☐	Other ☐
First Assistant Grade	Consultant ☐		Other ☐		

HIP REVISION PROCEDURE DETAILS								
Procedure Type	Single Stage Revision ☐		Stage 1 of 2 Stage Revision ☐		Stage 2 of 2 Stage Revision ☐		Excision Arthroplasty ☐	
Revision of	Primary Total Arthroplasty ☐				Previous Revision Arthroplasty (excluding excision arthroplasty) ☐			
Side	Left ☐	Right ☐						
Indications For / Findings at Time of Revision		Stem	Socket	Head				
	Aseptic Loosening	☐	☐	-				
	Implant Fracture	☐	☐	☐				
	Head/Socket Mismatch	-	☐	☐				
	Lysis	☐	☐	-				
	Malalignment	☐	☐	-				
	Peri-Prosthetic Fracture	☐	☐	-				
	Dislocation/Subluxation ☐				Wear of Acetabular Component ☐			
	Infection ☐				Dissociation of Liner ☐			
	Unexplained Pain ☐				Adverse Soft Tissue Reaction to Particulate Debris ☐			
					Other ☐			

PRIMARY OPERATION DETAILS			
Primary Operation Date OR Year	DD/MM/YYYY Please enter Date if known		Not Available ☐
Primary Operation Hospital			Not Available ☐

COMPONENTS REMOVED (Do not complete for Stage 2 of 2 Stage Revision)				
Femoral Component Removed	Yes ☐	No ☐	Brand	Not Available ☐
Acetabular Component Removed	Yes ☐	No ☐	Brand	Not Available ☐

SURGICAL APPROACH (Used for Single Stage Revision & Stage 2 of 2 Stage Revision)				
Patient Procedure	Revision Using Cement ☐			
	Revision Not Using Cement ☐			
	Revision of and to Resurfacing Arthroplasty ☐			
	Revision Not Classified Elsewhere (eg Hybrid) ☐			
Patient Position	Lateral ☐	Supine ☐		
Approach	Hardinge ☐		Extended Trochanteric Osteotomy ☐	
	Posterior ☐		Other ☐	
	Trochanteric Osteotomy ☐			
THROMBOPROPHYLAXIS REGIME (intention to treat)				
Chemical (In Hospital)	Aspirin ☐		Direct Thrombin Inhibitor ☐	
	LMWH ☐		(eg Dabigatran)	
	Pentasaccharide (eg Fondaparinux) ☐		Factor Xa Inhibitor (eg Rivaroxaban/Apixaban) ☐	
	Warfarin ☐		Other ☐	
			None ☐	
Mechanical	Foot Pump ☐		Other ☐	
	Intermittent Calf Compression ☐		None ☐	
	TED Stockings ☐			
BONEGRAFT USED				
Femur	Yes ☐	No ☐		
Acetabulum	Yes ☐	No ☐		
SURGEON'S NOTES				

INTRA OPERATIVE EVENT						
Untoward Intra Operative Event	None	☐	Shaft Fracture	☐	Other	☐
	Calcar Crack	☐	Shaft Penetration	☐		
	Pelvic Penetration	☐	Trochanteric Fracture	☐		

Minimum Dataset Form - COMPONENT LABELS

1. Please affix any component labels to this sheet and ensure any extra component label sheets are attached to the main Minimum Dataset Form.
2. Ensure all component details are provided, including cement.
3. The NJR DOES NOT record the following: wire, mesh, cables, plates, screws, surgical tools, endoprostheses or bipolar heads.

Cup or Shell	Liner (if used)
Stem	Head
Cement (if used)	Accessories (not screws)

Appendix 10. Details of cross-sectional imaging protocols at different centres

The NOC, Oxford, UK [284]

MARS-MRI

All MRI images were obtained using a 1.5 tesla magnet scanner (Magnetom Symphony, Siemens, Erlangen, Germany) with an optimised MARS protocol. This included short tau inversion recovery (STIR) coronal (field of view (FOV) = 265 mm; time to resonate (TR) = 3500 ms; time to echo (TE) = 32 ms; bandwidth = 305 Hz/pixel) and STIR axial sequences (FOV = 240 mm; TR = 3810 ms; TE = 23 ms; bandwidth = 326 Hz/pixel), and T1-weighted coronal (FOV = 265 mm; TR = 692 ms; TE = 6.6 ms; bandwidth = 326 Hz/pixel) and axial sequences (FOV = 240 mm; TR = 634 ms; TE = 7.2 ms; bandwidth = 326 Hz/pixel). The thickness of the slices was 4 mm for all coronal and axial sequences.

Ultrasound

All ultrasound examinations were performed using a high linear array transducer (Sonoline Antares, Siemens, New York, US: 4-12 MHz and GE Logic E9 Wauwatosa, Wisconsin, US: 5-15 MHz). Probes of varying frequencies were used based on the size of the patient. The examinations were performed with the patient supine initially and then in the lateral decubitus position. All examinations were performed using a standard technique recommended by the European Society of Skeletal Radiology [204], which encompasses a systematic approach to assess the anterior, medial, lateral and posterior hip regions. Standard sagittal oblique images were acquired by positioning the probe parallel to the long axis of the femoral neck. The probe was placed anteriorly,

posteriorly, and directly lateral to the femoral neck to obtain dynamic views of these regions.

The ROH, Birmingham, UK [209]

MARS-MRI

All MRI images were obtained using a 1.5 tesla magnet scanner (Magnetom Avanto, Siemens, Erlangen, Germany) with an optimised MARS protocol. This included STIR coronal (FOV = 320 mm; TR = 4050 ms; TE = 35 ms; bandwidth = 400 Hz/pixel) and STIR axial sequences (FOV = 380 mm; TR = 5000 ms; TE = 35 ms; bandwidth = 400 Hz/pixel), and T1-weighted coronal (FOV = 320 mm; TR = 574 ms; TE = 19 ms; bandwidth = 400 Hz/pixel) and axial proton density sequences (FOV = 380 mm; TR = 3060 ms; TE = 28 ms; bandwidth = 400 Hz/pixel). The thickness of the slices was 3 mm for all coronal and axial sequences, with a slice gap of 0.6 mm. The coronal sequences each included 32 slices, with the axial sequences each including 40 slices. The MARS-MRI scan typically took 16 minutes to perform.

Ultrasound

All examinations were performed using an ACUSON S2000 ultrasound system employing a 9.4 MHz transducer (Siemens Medical Solutions, Malvern, PA, US). The examinations were performed with the patient supine initially and then in the lateral decubitus position. Examinations were performing using a systematic approach to assess the anterior, medial, lateral and posterior hip regions, as recommended by the European Society of Skeletal Radiology [204]. The ultrasound scan typically took 15 minutes to perform.

Coxa Hospital for Joint Replacement, Tampere, Finland [206, 220]

MARS-MRI

All MRI images were obtained using a 1.5 tesla magnet scanner (Magnetom Avanto, Siemens, Erlangen, Germany) with adjustments made to produce minimal metal artefact. Sequences used were coronal and axial T1-weighted fast spin echo (FSE) and coronal, axial, and sagittal STIR. The thickness of the slices was 4 mm for all sequences.

Ultrasound

All ultrasound examinations were performed using a Logiq E9 ultrasound machine (GE Healthcare, Barrington, Illinois, US). In most cases, an ML 6-15-D linear transducer was used (4.5–15.0 MHz, 13×58 mm footprint, FOV 50 mm, and depth of field (DOF) 8 cm). In larger patients visibility was poor, therefore a 9L-D linear transducer (2.4–10.0 MHz, 14×53 mm footprint, FOV 45 mm, and DOF 12 cm) or a C1-5 convex transducer (1.6–6.0 MHz, 17×75 mm footprint, FOV 65 degrees, and DOF 35 cm) was used. The examinations were performed with the patient supine initially and then in the lateral decubitus position. An anterior approach was used to evaluate the hip joint, the iliopsoas muscle, and the tendon region. A lateral approach was used to examine the greater trochanteric and deep fascia region. Posteriorly, the hip joint and the trochanteric and gluteal regions were assessed. Medial structures were also examined.

Appendix 11. Summary of the receiver operating characteristic analysis for optimal Birmingham Hip Resurfacing blood metal ion thresholds compared with fixed authority thresholds in the validation cohort

	Blood metal ion threshold (µg/l)	AUC (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Misclassification (%)	Patients with ARMD not found
Unilateral BHR (n= 323)	Co = 10	89.4 (82.8-96.0)	47.4 (31.3-64.0)	99.3 (97.2-99.9)	90.0 (66.9-98.2)	93.4 (89.8-95.8)	6.8	20
	Co = 7		55.3 (38.5-71.0)	98.2 (95.7-99.4)	80.8 (60.0-92.7)	94.3 (90.8-96.5)	6.8	17
	Co = 3		68.4 (51.2-82.0)	92.3 (88.4-95.0)	54.2 (39.3-68.4)	95.6 (92.3-97.6)	10.5	12
	Co = 2.15		78.9 (66.0-91.9)	86.7 (82.7-90.6)	44.1 (32.3-55.9)	96.9 (94.7-99.0)	14.2	8
Bilateral BHR (n=93)	Max Co or Cr = 10	89.2 (81.3-97.1)	41.2 (19.4-66.5)	94.7 (86.4-98.3)	63.6 (31.6-87.6)	87.8 (78.3-93.7)	15.1	10
	Max Co or Cr = 7		70.6 (44.0-88.6)	89.5 (79.8-95.0)	60.0 (36.4-80.0)	93.2 (84.1-97.5)	14.0	5
	Max Co or Cr = 3		88.2 (62.3-97.9)	75.0 (63.5-83.9)	44.1 (27.6-61.9)	96.6 (87.3-99.4)	22.6	2
	Max Co or Cr = 5.52		70.6 (48.9-92.2)	86.8 (79.2-94.4)	54.5 (33.7-75.4)	93.0 (87.0-98.9)	16.1	5

ARMD = adverse reactions to metal debris; AUC = area under the curve; BHR = Birmingham Hip Resurfacing; Co = cobalt; Cr = chromium; NPV = negative predictive value; PPV = positive predictive value. 95% confidence intervals are in brackets.

Appendix 12. Detailed statistical methodology and results for the clinical risk scoring tool

Statistical methods: Internal validation of the final multivariable model

Robust internal validation of the final multivariable model was performed using calibration, discrimination, and bootstrapping techniques [299-301]. Calibration measures how closely predicted risk agrees with observed risk. Calibration was assessed for each tenth of predicted risk by using ten equally sized groups, with a Hosmer-Lemeshow goodness-of-fit test performed to quantify calibration; a non-significant p-value indicates a well-calibrated model. Discrimination is the model's ability to differentiate between HRAs with and without a pseudotumour, which is determined by calculating the AUC from ROC curve analysis. The AUC for a useful prognostic model is between 0.60 and 0.85 [300]. The final multivariable regression model was further validated by using bootstrapping (random sampling with replacement) with backward stepwise variable deletion [299]. Two hundred bootstrap repetitions were performed to obtain a bias-corrected estimate of the AUC using a modified data set.

Statistical methods: Developing a clinical risk scoring tool

Patient and radiographic predictors included in the final multivariable logistic regression model were used to develop a clinical risk scoring, with each factor assigned a weighting on the basis of its respective regression coefficient [300-302]. The calculated overall score represents a patient's risk of having evidence of a pseudotumour, with higher scores associated with increased pseudotumour risk. The respective regression coefficient for each predictor from the final multivariable model was converted to an integer risk score. The reference group for all variables was

assigned a risk score of 0. For all variables other than inclination, the risk scores for the non-reference groups were calculated by multiplying the respective regression coefficient by ten and then rounding to the nearest whole number. Inclination data were grouped with the midpoint for each group taken; the midpoint for the 30° to 39° group was 34.5°. The midpoint for non-reference inclination groups were then multiplied by the respective regression coefficient, multiplied by ten, and rounded to the nearest whole number. The discriminatory ability of the final overall risk score was determined by calculating the AUC. The optimal risk score threshold for identifying HRAs with evidence of a pseudotumour was also calculated.

Results: Clinical risk scoring tool

The clinical risk scoring tool developed from the final multivariable model has been summarised in the table below. Validation of the overall risk score model demonstrated that it had good discriminatory ability (AUC=0.796; 95% CI 0.747-0.845). The optimal risk score threshold for identifying HRAs with evidence of a pseudotumour was 18 points or more (95% CI 11.9-24.1 points). This risk score threshold provided 80.8% sensitivity and 65.2% specificity for identifying HRAs with evidence of a pseudotumour.

Covariate	Regression coefficient from final multivariable model	Risk score*
Gender		
Male	Reference group	0
Female	1.14	11
Cup position		
Inclination	0.041	
20° to $<30^o$	Reference group	0
30° to $<40^o$		14
40° to $<50^o$		18
50° to $<60^o$		23
60° to $<70^o$		27
70° to $<80^o$		31
80° to $<90^o$		35
Anteversion		
$<5^o$	Reference group	0
5° to $<10^o$	-1.18	-12
$\geq 10^o$	-1.13	-11
Acetabular osteolysis		
No	Reference group	0
Yes	1.62	16
Femoral osteolysis		
No	Reference group	0
Yes	2.88	29
Acetabular loosening		
No	Reference group	0
Yes	1.21	12
Heterotopic ossification		
No	Reference group	0
Yes	-1.64	-16

Table A12: Clinical risk scoring tool for determining pseudotumour risk based on patient and radiographic factors identified from the final multivariable model.
***All reference groups for each category have a risk score of zero. Higher scores indicate an increased risk of pseudotumour.**

Appendix 13. Data collection form used by radiologist when assessing cross-sectional imaging

HIP ULTRASOUND OR HIP MARS-MRI (please circle one)

Patient ID **DOB**

DATE OF SCAN	
RADIOLOGIST	
SIDE	**LEFT** **RIGHT**
PATHOLOGY	**SCAN NORMAL** **SCAN ABNORMAL**
PSEUDOTUMOUR	**YES** **NO**
PSEUDOTUMOUR DETAILS	**CONSISTENCY** SOLID / CYSTIC / MIXED **LOCATION** ANTERIOR / POSTERIOR / LATERAL / OTHER …….. **DIMENSIONS IN 3 PLANES (in cm)**……. **VOLUME (in cm^3)**……..... **INVOLVES NEUROVASCULAR STRUCTURES**
JOINT EFFUSION	**YES** **NO** **If yes:** Location - Dimensions in 3 planes - Volume (in cm^3) - Other details -
MUSCLE ATROPHY	**YES** **NO** **If yes:** **Location -** Gluteus medius / gluteus minimus / iliopsoas / other **Details** -
TENDON ABNORMALITIES	**YES** **NO** **If yes:** **Location -** Gluteus medius / gluteus minimus / iliopsoas / other **Problem -** avulsion, atrophy, ossification, other **Details** -
BURSAL DISTENSION	**YES** **NO** **If yes:** Trochanteric / iliopsoas / other **Details** -
BURSAL THICKENING	**YES** **NO** **If yes:** Trochanteric / iliopsoas / other **Details** -
OTHER ABNORMALITIES	

Appendix 14. Detailed review of diagnostic test characteristics for cross-sectional imaging vs. intra-operative findings

All hip numbers referred to below represent the corresponding cases summarised in Table 5.6.

	Ultrasound findings		**Total hips**
True disease status	No pseudotumour	Pseudotumour	
No intra-operative pseudotumour	3	4	7
Intra-operative pseudotumour	3	30	33
Total hips	6	34	40

Table A14.1: Contingency table for ultrasound alone vs. intra-operative findings.

Ultrasound alone resulted in seven cases of disagreement with the intra-operative findings. There were four false positive cases (hip numbers 2, 5, 6, 7) and three false negative cases (hip numbers 3, 4, 8).

	MARS-MRI findings		**Total hips**
True disease status	No pseudotumour	Pseudotumour	
No intra-operative pseudotumour	4	3	7
Intra-operative pseudotumour	2	31	33
Total hips	6	34	40

Table A14.2: Contingency table for MARS-MRI alone vs. intra-operative findings.

MARS-MRI alone resulted in five cases of disagreement with the intra-operative findings. There were three false positive cases (hip numbers 2, 6, 7) and two false negative cases (hip numbers 1, 9).

	Combined imaging findings		**Total hips**
True disease status	No pseudotumour	Pseudotumour	
No intra-operative pseudotumour	4	3	7
Intra-operative pseudotumour	0	33	33
Total hips	4	36	40

Table A14.3: Contingency table for ultrasound and MARS-MRI combined vs. intra-operative findings.

Combined imaging with ultrasound and MARS-MRI resulted in three cases of disagreement with the intra-operative findings. There were three false positive cases (hip numbers 2, 6, 7) and no false negative cases. The false negative cases from ultrasound alone were all different hips compared with the false negative cases from MARS-MRI alone. Therefore combined imaging resulted in no false negative cases. This resulted in a marked improvement in the NPV when the results from both modalities were combined (100%), compared with when using ultrasound alone (50%) or MARS-MRI alone (66.7%) (Table 5.7.). As there were no false negative cases with combined imaging, the sensitivity also improved slightly to 100%. The three false positive cases with combined imaging that remained from using either MARS-MRI or ultrasound alone resulted in similar specificities and PPVs with combined imaging.

Appendix 15. University of California, Los Angeles (UCLA) activity score questionnaire completed by hip resurfacing arthroplasty patients

UCLA Activity-Level Rating[1]

Level	Activity	Please indicate
10	Regularly participate in *impact sports* such as jogging, tennis, skiing, acrobatics, ballet, heavy labour or backpacking	❑
9	Sometimes participate in *impact sports*	❑
8	Regularly participate in *very active* events, such as bowling or golf	❑
7	Regularly participate in *active* events, such as bicycling	❑
6	Regularly participate in *moderate activities*, such as swimming and unlimited housework or shopping	❑
5	Sometimes participate in *moderate activities*	❑
4	Regularly participate in *mild activities*	❑
3	Sometimes participate in *mild activities*	❑
2	Mostly *inactive*: restricted to minimal activities of daily living	❑
1	Wholly *inactive*: dependent on others; cannot leave residence	❑

[1] Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. J Arthroplasty. 1998 Dec;13(8):890-5.

Appendix 16. Patient demographics for the cohort undergoing initial assessment that were subsequently excluded from the analysis

Covariate		Revised before 2012 (n=16)	Died before 2012 (n=4)	Declined to participate (n=29)	Pseudotumour on initial assessment (n=12)
Gender	Male Female	2 (12%) 14 (88%)	4 (100%) 0 (0%)	20 (69%) 9 (31%)	4 (33%) 8 (67%)
Age at initial ultrasound	Mean (range) in years	59.2 (39.2-73.1)	66.4 (61.6-70.2)	60.6 (40.8-73.8)	61.6 (52.9-71.4)
MoM hip laterality	Unilateral Bilateral	10 (63%) 6 (37%)	2 (50%) 2 (50%)	11 (38%) 18 (62%)	8 (67%) 4 (33%)
Implant design	BHR Conserve Plus Recap	10 (63%) 4 (25%) 2 (12%)	1 (25%) 3 (75%) 0 (%)	18 (62%) 11 (38%) 0 (0%)	3 (25%) 9 (75%) 0 (0%)
Time between hip resurfacing arthroplasty and initial ultrasound	Mean (range) in years	4.1 (3.0-5.8)	3.9 (3.4-5.2)	4.7 (3.0-7.1)	3.1 (3.0-6.8)
Acetabular inclination	Mean (range) in ^{o}	47.0 (34.7-55.3)	39.4 (37.9-40.8)	45.5 (30.1-64.0)	48.3 (26.0-62.6)
Acetabular anteversion	Mean (range) in ^{o}	16.5 (5.1-34.3)	16.2 (14.9-17.5)	13.9 (7.2-20.8)	17.6 (8.2-32.0)
Blood cobalt	Median (IQR) in µg/l	6.3 (4.7-13.3)*	0.9 (0.9-1.0)	1.6 (1.3-2.7)	3.8 (1.8-4.8)
Blood chromium	Median (IQR) in µg/l	7.1 (3.8-19.6)*	1.3 (1.1-1.5)	1.5 (1.2-2.4)	3.0 (1.2-7.2)
Initial ultrasound grade	Grade 1 Grade 2 Grade 3 Grade 4	2 (13%) 0 (0%) 1 (6%) 13 (81%)	2 (50%) 2 (50%) 0 (0%) 0 (0%)	26 (90%) 3 (10%) 0 (0%) 0 (0%)	0 (0%) 0 (0%) 0 (0%) 12 (100%)
Initial pseudotumour volume	Median (IQR) in cm^3	36.0 (22.4-48.0)	NA	NA	21.3 (16.9-28.5)
Initial pseudotumour consistency	Cystic Mixed Solid	6 (46%) 5 (39%) 2 (15%)	NA	NA	9 (75%) 3 (25%) 0 (0%)

CONTINUED /...

Initial OHS	Median (IQR)	41.0 (35.0-46.5)	48.0 (48.0-48.0)	47.0 (44.0-48.0)	47.0 (41.0-48.0)
Initial UCLA score	Mean (range)	5.6 (3-8)	8.5 (7-10)	7.3 (1-10)	6.9 (3-9)
Hips undergoing repeat assessment					
Time between repeat ultrasound	Mean (range) in years	NA	NA	NA	4.1 (3.2-4.6)
Repeat ultrasound grade	Grade 1 Grade 2 Grade 3 Grade 4	NA	NA	NA	0 (0%) 0 (0%) 0 (0%) 12 (100%)
Repeat pseudotumour volume	Median (IQR) in cm^3	NA	NA	NA	15.9 (8.1-44.7)
Repeat pseudotumour consistency	Cystic Mixed Solid	NA	NA	NA	9 (75%) 3 (25%) 0 (0%)
OHS					
Repeat	Median (IQR)	NA	NA	NA	46.0 (39.5-48.0)
Change	Mean (range)	NA	NA	NA	0.9 (-18-12)
UCLA score					
Repeat	Mean (range)	NA	NA	NA	6.4 (4-8)
Change	Mean (range)	NA	NA	NA	-0.7 (-2-0)

BHR = Birmingham Hip Resurfacing; IQR = interquartile range; MoM = metal-on-metal; NA = not applicable; OHS = Oxford hip score; UCLA = University of California, Los Angeles.

***For the 14 hip revisions performed for symptomatic pseudotumour the range of blood metal ion concentrations at initial assessment were 2.9 µg/l to 64.1 µg/l for cobalt and 2.1 µg/l to 45.2 µg/l for chromium.**

Appendix 17. Detailed review of diagnostic test characteristics for initial assessment findings vs. the development of new pseudotumours

Diagnostic test characteristics were assessed for three different clinical scenarios based on the results of the initial assessment (Table 5.10.). The gold standard diagnosis of interest for all scenarios was HRAs that did not develop new pseudotumours by repeat assessment. A positive test result was having a normal initial assessment, which was defined differently for each scenario. As most HRAs (n=125) did not develop new pseudotumours (Table 5.8. and 5.9.), the prevalence of the gold standard diagnosis of interest was high (89.3% for all three scenarios; 95% CI 82.9%-93.9%). The contingency tables used to calculate the diagnostic test characteristics for each scenario have been summarised below.

	Initial assessment		**Total hips**
True disease status	Abnormal	Normal	
Developed new pseudotumour	15	0	15
Did not develop new pseudotumour	75	50	125
Total hips	90	50	140

Table A17.1: Contingency table with normal initial assessment defined as both a normal ultrasound scan and blood cobalt and chromium concentrations under 2 µg/l.

	Initial assessment		**Total hips**
True disease status	Abnormal	Normal	
Developed new pseudotumour	11	4	15
Did not develop new pseudotumour	19	106	125
Total hips	30	110	140

Table A17.2: Contingency table with normal initial assessment defined as a normal ultrasound scan alone.

	Initial assessment		**Total hips**
True disease status	Abnormal	Normal	
Developed new pseudotumour	13	2	15
Did not develop new pseudotumour	68	57	125
Total hips	81	59	140

Table A17.3: Contingency table with normal initial assessment defined as blood cobalt and chromium concentrations below 2 µg/l alone.

Appendix 18. Letter of study approval from the National Joint Registry Research Sub-committee

National Joint Registry
www.njrcentre.org.uk

HQIP
Healthcare Quality Improvement Partnershi

Eve Riley
Research Officer- NJR
Healthcare Quality Improvement Partnership (HQIP)
6th Floor, Tenter House
45 Moorfields | London | EC2Y 9AE
T: 020 7997 7354 **W:** www.hqip.org.uk

2 January 2015
To: Professor David Murray
Email: david.murray@ndorms.ox.ac.uk

Dear Professor Murray

Re. NJR research application: ***'Outcomes following revision of MoM hip replacements'***

Thank you for your application to the NJR Research Sub-committee (RSC) for support of the above named project.

The RSC has recently reviewed your research application and we are pleased to inform you that the Committee was able to recommend **approval** of this study, which has been classified as an external, independent research project supported by the NJR (an 'NJR Supported Project').

Please note that this provisional approval is subject to the following technical points being noted:

- NJR will give anonymised patient ID only (i.e. no consultant anonymised ID or unit anonymised ID will be given).
- The Committee recommends the formal involvement of a statistician in the project team.
- Confounder -ALVAL (use of option ALVAL came in later (c.2009/10)). Thus genuine ALVAL might be confused with infection or unexplained pain in earlier years.
- The Committee also advised the applicants to note that this study is unlikely to detect differences in peri-operative mortality as this is so rare.

Conditions of Approval:

A condition of this approval is that you agree to provide 6 monthly progress reports to the NJR. Please submit these to me electronically within 30 days of the review date. These progress reports will be circulated to the NJR RSC and the final report summary will be published on the NJR website.

Please also read carefully the accompanying 'Terms and Conditions of approval' and 'Data Sharing Agreement' and return two signed copies of each document to me at the address above. Once received, subject to the above points being addressed to the satisfaction of the Committee, the Principal Investigator will receive an authorisation letter on behalf of the NJR Research Sub-committee stating that it is happy for the project to proceed, and instructing Northgate to prepare and release the data.

National Joint Registry
www.njrcentre.org.uk

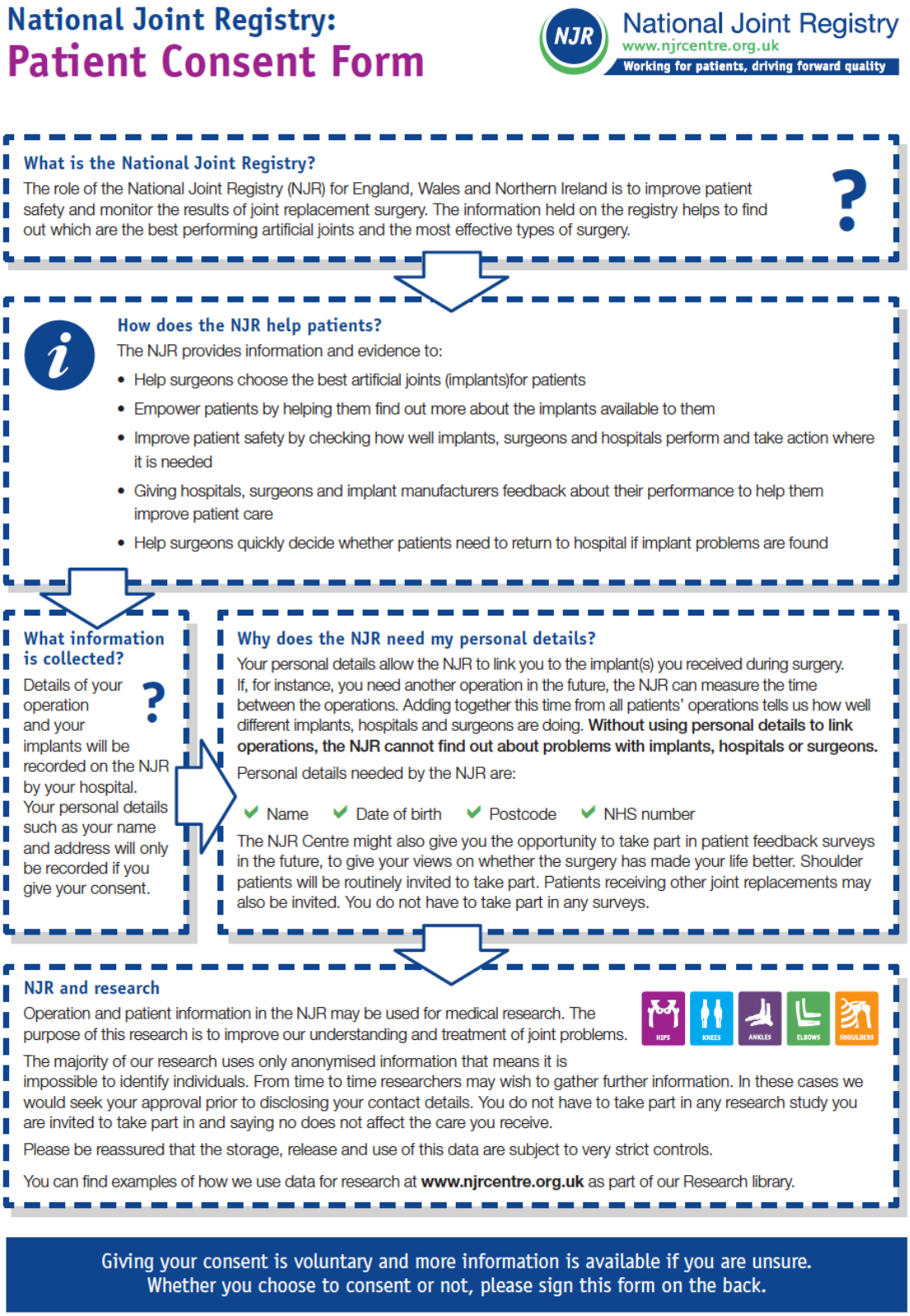
HQIP
Healthcare Quality Improvement Partnership

Yours sincerely

E. Riley

Eve Riley, Research Officer
On behalf of the Research Sub-committee

Appendix 19. National Joint Registry patient consent form



National Joint Registry:
Patient Consent Form

National Joint Registry
www.njrcentre.org.uk
Working for patients, driving forward quality

What is the National Joint Registry?

The role of the National Joint Registry (NJR) for England, Wales and Northern Ireland is to improve patient safety and monitor the results of joint replacement surgery. The information held on the registry helps to find out which are the best performing artificial joints and the most effective types of surgery.

How does the NJR help patients?

The NJR provides information and evidence to:

- Help surgeons choose the best artificial joints (implants)for patients
- Empower patients by helping them find out more about the implants available to them
- Improve patient safety by checking how well implants, surgeons and hospitals perform and take action where it is needed
- Giving hospitals, surgeons and implant manufacturers feedback about their performance to help them improve patient care
- Help surgeons quickly decide whether patients need to return to hospital if implant problems are found

What information is collected?

Details of your operation and your implants will be recorded on the NJR by your hospital. Your personal details such as your name and address will only be recorded if you give your consent.

Why does the NJR need my personal details?

Your personal details allow the NJR to link you to the implant(s) you received during surgery. If, for instance, you need another operation in the future, the NJR can measure the time between the operations. Adding together this time from all patients' operations tells us how well different implants, hospitals and surgeons are doing. **Without using personal details to link operations, the NJR cannot find out about problems with implants, hospitals or surgeons.**

Personal details needed by the NJR are:

✔ Name ✔ Date of birth ✔ Postcode ✔ NHS number

The NJR Centre might also give you the opportunity to take part in patient feedback surveys in the future, to give your views on whether the surgery has made your life better. Shoulder patients will be routinely invited to take part. Patients receiving other joint replacements may also be invited. You do not have to take part in any surveys.

NJR and research

Operation and patient information in the NJR may be used for medical research. The purpose of this research is to improve our understanding and treatment of joint problems.

HIPS KNEES ANKLES ELBOWS SHOULDERS

The majority of our research uses only anonymised information that means it is impossible to identify individuals. From time to time researchers may wish to gather further information. In these cases we would seek your approval prior to disclosing your contact details. You do not have to take part in any research study you are invited to take part in and saying no does not affect the care you receive.

Please be reassured that the storage, release and use of this data are subject to very strict controls.

You can find examples of how we use data for research at **www.njrcentre.org.uk** as part of our Research library.

Giving your consent is voluntary and more information is available if you are unsure.
Whether you choose to consent or not, please sign this form on the back.

NJR data and other healthcare information

Operation and patient information in the NJR is used to link to other healthcare information. It is also analysed and made available to surgeons, hospitals and manufacturers - without any of your personal details - as part of our feedback services.

Doing this improves the NJR's ability to monitor patient safety and patient outcomes. It also means that people and organisations involved in improving joint replacement surgery can better understand and develop improved or more cost-effective medical treatments.

Do I have to give my consent?

No. If you do not consent, your operation details will be stored on the NJR without any personal details, so you cannot be identified. If you change your mind about consenting to the NJR holding your personal details, please contact the NJR Centre – see 'Finding out more' below.

Is my information safe?

Your personal details are kept confidential at all times. Very strict rules and secure procedures are in place to ensure that your information is kept safe. You can ask for a copy of your personal data held on the NJR by writing to the NJR Centre.

Finding out more

NJR Website
www.njrcentre.org.uk

E-mail
enquiries@njrcentre.org.uk

National Joint Registry
www.njrcentre.org.uk
Working for patients, driving forward quality

NJR Helpline
Tel: 0845 345 9991
Mon-Fri, 9am to 5pm (excluding public holidays)

NJR Centre
Peoplebuilding 2, Peoplebuilding Estate, Maylands Avenue, Hemel Hempstead HP2 4NQ

This form should be kept as part of the patient record. DO NOT send this form to the NJR Centre

Patient Details (all mandatory)	To be completed by the hospital
Surname	Hospital
Forename	NHS or national patient number
Date of birth	Height m & Weight kg
Postcode (home)	or BMI

I CONSENT to my personal details being recorded within the NJR. I understand that the NJR will not release my personalised data unless required by law or where there is a clear overriding public interest in disclosure. Where possible, I will be told if any disclosure is to take place.

Signature Date

I DO NOT CONSENT to my personal details being recorded within the NJR.

Signature Date

To be completed by the hospital (person accepting patient consent)

Name Signature

Position Date

Giving your consent is voluntary and more information is available if you are unsure.

Version 6.0 August 2014

Appendix 20. Methodology for multiple imputation of missing data

In total, 50 copies of the entire dataset were produced using the MI procedure in Stata for multiple imputation by chained equations [310, 311]. Data were assumed to be missing at random, although this cannot be proven definitively. In each copy of the dataset the missing data values were imputed. These imputed values were estimated using the data that was available for the missing variables, the data available for the other covariates with no missing data, and the study outcome of re-revision surgery (which included the Nelson-Aalen estimate of survival), given that all of these factors carried information about the missing data values. A degree of uncertainty is built into the imputed values based on the data available and the proportion of missing values for each variable.

The adequacy of the imputation process was formally assessed and was deemed satisfactory. The Monte Carlo error was examined separately for each covariate with missing data to ensure the amount of variation introduced by the imputed data values was minimal. Descriptive statistics (mean and SD for revision BMI, and the percentage for the three categorical covariates) were assessed before and after imputation to confirm that they remained the same. Similarly crude univariable competing-risk regression models were assessed separately for the four covariates with imputed data to ensure the respective sub-hazard ratio (SHR) and 95% CI was the same before and after imputation.

The regression analyses performed in the subgroup of patients with complete data available were subsequently repeated using the 50 copies of the dataset, which now contained complete data for all variables. The regression models were fitted to the

imputed datasets, with standard errors calculated using Rubin's Rules. An overall estimate of the regression model was calculated from the imputed datasets, which was produced by averaging the results obtained for each of the 50 datasets.

Appendix 21. Competing-risk regression analysis to identify predictors of re-revision surgery when using the imputed dataset

Covariate	Univariable sub-hazard ratio (95% CI)	p-value	Multivariable sub-hazard ratio (95% CI)	p-value
Gender				
Female *vs. male*	1.13 (0.84-1.52)	0.411	^	
Age at revision (per year)	0.99 (0.98-1.01)	0.232	^	
Revision BMI (per kg/m^2)	1.04 (1.01-1.07)	**0.008**	1.05 (1.02-1.08)	**0.003**
Bilateral revisions for ARMD	1.04 (0.64-1.70)	0.863	^	
Primary arthroplasty				
Hip resurfacing arthroplasty	1.00	Ref	^	
MoM THA	1.51 (1.09-2.10)	**0.013**		
Time from primary to revision (per year)	0.89 (0.82-0.97)	**0.006**	0.91 (0.84-0.99)	**0.038**
Revision ASA grade				
1	1.00	Ref	^	
2	0.98 (0.69-1.41)	0.926		
3 or above	0.89 (0.49-1.61)	0.691		
VTE – chemical				
None	1.00	Ref	^	
LMWH (+/-other)	0.69 (0.44-1.06)	0.093		
Aspirin only	0.83 (0.34-2.02)	0.679		
Other	0.76 (0.48-1.19)	0.230		
VTE – mechanical				
Any *vs. none*	0.88 (0.43-1.78)	0.715	^	
Number of ARMD revisions performed per centre (per ten cases)	1.02 (0.99-1.04)	0.074	1.03 (1.00-1.06)	**0.011**
Revision surgeon grade				
Consultant *vs. other*	0.87 (0.45-1.70)	0.684	^	
Surgical approach				
Posterior *vs. other*	0.90 (0.63-1.29)	0.559	^	
Revision indications				
ARMD only	1.00	Ref	^	
2 to 6 indications	0.81 (0.59-1.09)	0.160		
Revision procedure performed				
Cup only	1.00	Ref	1.00	Ref
All components	0.77 (0.56-1.06)	0.105	*	0.338
Stem only	0.88 (0.36-2.16)	0.776	*	0.712
Modular component only	1.61 (1.07-2.44)	**0.023**	2.04 (1.33-3.13)	**0.001**
Revision femoral head size (mm)				
<36	1.00	Ref	^	
36	1.10 (0.81-1.50)	0.530		
>36	1.07 (0.56-2.03)	0.835		

CONTINUED /...

Covariate	Univariable sub-hazard ratio (95% CI)	p-value	Multivariable sub-hazard ratio (95% CI)	p-value
Revision bearing surface				
CoP	1.00	Ref	1.00	Ref
CoC	1.51 (1.08-2.12)	**0.015**	1.57 (1.10-2.24)	**0.013**
MoP	1.42 (0.97-2.07)	0.069	1.45 (0.97-2.17)	0.068
CoM, MoM or MoC	**	**	**	**
Revision cup fixation				
Uncemented	1.00	Ref	^	
Cemented	0.89 (0.54-1.47)	0.656		
Revision stem fixation				
Uncemented	1.00	Ref	^	
Cemented	0.76 (0.44-1.30)	0.316		
Bone graft (femoral)	0.91 (0.37-2.20)	0.827	^	
Bone graft (acetabular)	1.51 (1.10-2.07)	**0.011**	1.85 (1.33-2.58)	**<0.001**

ARMD = adverse reactions to metal debris; ASA = American Society of Anesthesiologists; BMI = body mass index; CoC = ceramic-on-ceramic; CoM = ceramic-on-metal; CoP = ceramic-on-polyethylene; CI = confidence interval; LMWH = low molecular weight heparin; MoC = metal-on-ceramic; MoM = metal-on-metal; MoP = metal-on-polyethylene; Ref = reference group; THA = total hip arthroplasty; VTE = venous thrombo-embolism prophylaxis. P-values <0.05 in bold text.
^ Covariate was not eligible for inclusion in the final multivariable model
*** Specific subgroup did not meet inclusion criteria for final multivariable model (p>0.20)**
**** Unable to calculate value as no hips in this subgroup underwent re-revision surgery**
All univariable and multivariable analyses were based on a cohort of 2,535 hips with complete data available for all variables following multiple imputation (192 hips undergoing re-revision surgery and 41 deaths).

Appendix 22. Indications for re-revision surgery in matched metal-on-metal hip arthroplasty revision patients

Indication for re-revision surgery	Cohort of interest			Hazard ratio for Cox regression (95%CI) comparison between ARMD vs. non-ARMD revisions	p-value
	Whole cohort	ARMD revisions (% of all re-revisions for specified indication)	Non-ARMD revisions (% of all re-revisions for specified indication)		
All causes	134	43 (32.1)	91 (67.9)	0.52 (0.36-0.75)	**<0.001**
ARMD	26	14 (53.8)	12 (46.2)	1.32 (0.61-2.85)	0.486
Infection	23	12 (52.2)	11 (47.8)	1.12 (0.49-2.54)	0.784
Peri-prosthetic fracture *	9	2 (22.2)	7 (77.8)	0.32 (0.07-1.53)	0.153
Aseptic loosening and/or osteolysis *	23	5 (21.7)	18 (78.3)	0.32 (0.12-0.87)	**0.026**
Dislocation / subluxation	28	6 (21.4)	22 (78.6)	0.29 (0.12-0.72)	**0.007**
Other **	15	4 (26.7)	11 (73.3)	0.41 (0.13-1.30)	0.129
Unexplained pain	10	0 (0.0)	10 (100)	***	***

ARMD = adverse reactions to metal debris; CI = confidence interval. P-values <0.05 in bold text.

*** Re-revision indication included failures on the acetabular and/or femoral side.**

**** Other indications included implant malalignment, head/socket mismatch, acetabular component wear, implant fracture, and liner dissociation.**

***** Unable to calculate as no re-revisions for unexplained pain in the ARMD revision group**

Hazard ratios greater than 1 represent an increased risk of implant re-revision surgery for the listed re-revision indication in the ARMD revision group.

Apart from unexplained pain, any listed revision indication may include anything below it but nothing above it.

Appendix 23. Sensitivity analysis using the complete dataset with revision body mass index excluded

In the unmatched cohort of 4,896 primary MoMHAs undergoing revision surgery for any indication there were three covariates with missing data: revision BMI (n=1,731); revision femoral head size (n=111); and revision bearing surface (n=320). The analyses presented (Chapter 6.7.) were based on a complete dataset for all covariates with any missing values estimated using multiple imputation before the ARMD and non-ARMD revision groups were matched. A sensitivity analysis was performed to ensure the findings obtained from this imputed dataset were similar to those obtained from the subgroup where complete data was available.

The effect of revision BMI on each outcome of interest was assessed for the 3,165 hips where revision BMI data was available.

	Intra-operative complications (95% confidence interval)	**Implant survival (95% confidence interval)**	**Patient survival (95% confidence interval)**
Regression analysis assessing the effect of revision BMI on outcome of interest	Crude odds ratio = 1.02 (0.97-1.06) p = 0.490	Crude hazard ratio = 1.02 (0.99-1.04) p = 0.094	Crude hazard ratio = 0.96 (0.91-1.01) p = 0.084

Table A23.1: Univariable regression analyses for revision body mass index vs. outcomes in the unmatched cohort with complete revision body mass index data available.

Given the large proportion of missing data for revision BMI (35%) and the findings that revision BMI had no statistically significant effect on any outcome of interest, a sensitivity analysis was performed using the complete dataset with revision BMI excluded. The unmatched cohort with complete data available for all covariates

(excluding revision BMI) consisted of 4,576 hips. After propensity score matching the cohort consisted of 2,394 hips (1,197 ARMD vs. 1,197 non-ARMD). When the regression analyses were repeated in this subgroup, the outcomes and effect sizes were similar to those obtained when using the complete imputed dataset of matched patients (Table 6.5.).

	Number of hips	**Risk of intra-operative complications**	**5 year implant survival (95% CI)**	**5 year patient survival (95% CI)**
Overall cohort	2,394	2.5% (n=59)	92.1% (90.0%-93.7%)	95.6% (93.2%-97.2%)
ARMD group	1,197	2.5% (n=30)	94.6% (91.7%-96.5%)	96.4% (90.8%-98.6%)
Non-ARMD group	1,197	2.4% (n=29)	89.9% (86.8%-92.4%)	95.2% (92.3%-97.0%)
Regression analysis (95% CI)		OR = 1.04 (0.62-1.74) p = 0.895	HR = 0.48 (0.33-0.70) **p < 0.001**	HR = 0.61 (0.32-1.16) p = 0.133

Table A23.2: Univariable regression analyses in the matched cohort with complete data available (excluding revision body mass index).

ARMD = adverse reactions to metal debris; CI = confidence interval; HR = hazard ratio; OR = odds ratio. P-values <0.05 in bold text.

Odds ratios and hazard ratios greater than 1 represent an increased risk of the specified outcome in the ARMD revision group.

Appendix 24. Analysis of outcomes following revision surgery in the unmatched cohort

Outcomes following revision surgery were compared between the unmatched ARMD and non-ARMD groups using univariable logistic (intra-operative complications) and Cox (mortality and re-revision surgery) regression models. The findings from the unmatched cohort were different compared with those obtained from the complete imputed dataset of matched patients (Table 6.5.). These observations support the use of the propensity score matching methods to reduce any potential confounding by revision indication.

	Number of hips	**Risk of intra-operative complications**	**5 year implant survival (95% CI)**	**5 year patient survival (95% CI)**
Overall cohort	4,896	2.1% (n=101)	88.1% (86.7%-89.3%)	94.8% (93.6%-95.7%)
ARMD group	2,535	1.7% (n=42)	89.6% (87.4%-91.4%)	95.8% (92.2%-97.7%)
Non-ARMD group	2,361	2.5% (n=59)	87.3% (85.5%-88.8%)	93.9% (92.4%-95.1%)
Regression analysis (95% CI)		Crude OR = 0.65 (0.44-0.98) **p = 0.037**	Crude HR = 0.89 (0.73-1.07) p = 0.217	Crude HR = 0.62 (0.43-0.89) **p = 0.010**

Table A24.1: Univariable regression analyses in the unmatched cohort.

ARMD = adverse reactions to metal debris; CI = confidence interval; HR = hazard ratio; OR = odds ratio. P-values <0.05 in bold text.

Odds ratios and hazard ratios greater than 1 represent an increased risk of the specified outcome in the ARMD revision group.

Appendix 25. Publications associated with this thesis

The Journal of Arthroplasty 30 (2015) 1317–1323

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Follow-Up of Metal-on-Metal Hip Arthroplasty Patients Is Currently Not Evidence Based or Cost Effective

Gulraj S. Matharu, BSc (Hons), MBChB, MRCS, Stephen J. Mellon, PhD, David W. Murray, MD, FRCS (Orth), Hemant G. Pandit, DPhil, FRCS (Orth)

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Oxford, United Kingdom, OX3 7LD

ARTICLE INFO

Article history:
Received 4 December 2014
Accepted 7 March 2015

Keywords:
metal-on-metal
hip resurfacing
total hip arthroplasty
follow-up
finance
guidelines

ABSTRACT

Over one-million patients worldwide have received metal-on-metal (MoM) hip arthroplasties with a significant proportion requiring revision surgery in the short-term for adverse reaction to metal debris (ARMD). Worldwide authorities have subsequently issued follow-up guidance for MoM hip patients. This article compares follow-up guidelines for MoM hips published by five worldwide authorities, analyses these protocols in relation to published evidence, and assesses the financial implications of these guidelines. A number of major differences exist between authorities regarding patient follow-up, with vast cost differences between protocols (£84 to £988/patient/year for stemmed MoM hips and £0 to £988/patient/year for hip resurfacing). Current worldwide guidance is neither evidence-based nor financially sustainable with most protocols lacking the sensitivity to detect asymptomatic ARMD lesions.

© 2015 Elsevier Inc. All rights reserved.

Over one-million patients worldwide have received large-diameter metal-on-metal (MoM) hip arthroplasties (hip resurfacing (HR) and total hip arthroplasty (THA)) [1]. Recently high short-term failure rates due to adverse reaction to metal debris (ARMD) have been observed with MoM hips [2–6]. ARMD is the sequelae of metal debris released from MoM articulations due to wear and corrosion [7], which can result in destructive soft-tissue masses often requiring revision [8]. In the United Kingdom (UK) the prevalence of ARMD revision surgery is increasing, accounting for 13% of all revisions performed in 2012 [9]. As ARMD can cause significant bone loss and muscle damage, short-term outcomes following revision are often poor [10,11]. These unsatisfactory outcomes are concerning given that most patients are young and active [12–14]. Early revision surgery for ARMD is currently recommended which may improve outcomes [15,16]. To identify patients with ARMD early, regulatory authorities worldwide have published follow-up guidance for MoM hip patients [16–20].

The authors would like to thank The Royal College of Surgeons of England and The Arthritis Research Trust, which have provided one of the authors with funding in the form of a Surgical Research Fellowship.

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to http://dx.doi.org/10.1016/j.arth.2015.03.009.

Reprint requests: Gulraj S Matharu, BSc (Hons), MBChB, MRCS, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Oxford, United Kingdom, OX3 7LD.

http://dx.doi.org/10.1016/j.arth.2015.03.009
0883-5403/© 2015 Elsevier Inc. All rights reserved.

This article compares follow-up guidelines for MoM hip patients issued by five worldwide authorities, analyses these protocols in relation to published evidence, and assesses the financial implications.

Follow-Up Guidelines

Guidance for MoM hip follow-up has been issued by the: UK Medical and Healthcare Products Regulatory Agency (MHRA) [16], European Federation of National Associations of Orthopaedics and Traumatology (EFORT) [17], United States of America (USA) Food and Drug Administration (FDA) [18], Therapeutic Goods Administration of Australia [19], and Health Canada [20] (Table 1). Major differences in follow-up guidance exist between authorities.

HR and THA

The USA and Canadian authorities do not distinguish between the follow-up of HRs and THAs. Other authorities make this distinction and follow-up large-diameter (≥36 mm) THAs regardless of symptoms [16,17,19]. Furthermore, these three authorities stratify HR follow-up by symptoms, with some also using ARMD risk factors. Small HR femoral head sizes are a high-risk group requiring at least annual surveillance in European (<50 mm) and Australian (≤45 mm) guidance, but not by the MHRA. Although USA and Canadian guidance stratifies follow-up according to symptoms, both advise closer follow-up (review intervals not stated) for patients with ARMD risk factors.

COPYRIGHT © 2016 BY THE JOURNAL OF BONE AND JOINT SURGERY, INCORPORATED

Prevalence of and Risk Factors for Hip Resurfacing Revision

A Cohort Study Into the Second Decade After the Operation

Gulraj S. Matharu, BSc(Hons), MRCS, MRes, Andrew Judge, BSc, MSc, PhD, David W. Murray, MD, FRCS(Orth), and Hemant G. Pandit, FRCS(Tr & Orth), DPhil

Investigation performed at the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Nuffield Orthopaedic Centre, University of Oxford, Oxford, United Kingdom

Background: Most metal-on-metal hip resurfacing (MoMHR) designs have experienced high short-term failure rates because of pseudotumors. The impact of this complication into the second decade after the procedure is unknown. We investigated (1) the prevalence of, and risk factors for, all-cause and pseudotumor-related revision at up to 15 years following MoMHR and (2) whether risk factors were sex-specific.

Methods: This single-center prospective cohort study included 1,429 MoMHRs (1216 patients; 40% female) implanted between 1999 and 2009. Patients were contacted in 2010 and 2012 as per national recommendations. Patients with symptoms related to the hip and/or suboptimal Oxford Hip Scores (≤41 of 48 points) underwent cross-sectional imaging and blood metal-ion sampling. Revision diagnoses were established using operative and histopathological findings. Multivariate Cox proportional hazard models were used to assess the association of predictor variables with the time to all-cause and pseudotumor-related revisions.

Results: One hundred and eighty MoMHRs (12.6%) were revised for all causes, and 111 (7.8% of the series and 61.7 % of all revisions) were revised because of pseudotumor. Survival analysis showed the 15-year cumulative revision rate for all causes to be 19.5% (95% confidence interval [CI] = 16.2% to 23.2%) and the 15-year rate of revision due to pseudotumor to be 14.0% (95% CI = 11.0% to 17.7%). Small femoral head size (hazard ratio [HR] per 2 mm = 0.92, 95% CI = 0.88 to 0.97; p = 0.003) and certain implant designs (HR = 1.55 to 3.01; p ≤ 0.029) significantly increased the all-cause revision risk. Female sex (HR = 2.03, 95% CI = 1.19 to 3.44; p = 0.009) and young age (HR per year = 0.98, 95% CI = 0.96 to 1.00; p = 0.020) significantly increased the pseudotumor-related revision risk but not the all-cause revision risk. Risk factors for all-cause and pseudotumor-related revision were sex-specific. In females, small femoral head size (p = 0.014) increased the all-cause revision risk, and young age was the only predictor of pseudotumor-related revision (p = 0.019). In males, implant design was the only predictor of all-cause revision (p ≤ 0.015) and pseudotumor-related revision (p = 0.001).

Conclusions: The prevalence and rates of revision for all causes and pseudotumor were high at up to 15 years following MoMHR. Predictors of revision differed between all-cause and pseudotumor-related revisions and were sex-specific. These factors must be appropriately weighted when risk-stratifying patients with MoMHRs for surveillance.

Level of Evidence: Prognostic Level II. See Instructions for Authors for a complete description of levels of evidence.

Peer review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. It was also reviewed by an expert in methodology and statistics. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

High short-term failure rates (9.7% to 12.6% at 10 years) have been reported for most metal-on-metal hip resurfacing (MoMHR) designs[1,2]. Revisions have commonly been due to pseudotumors[3,4], with the prevalence of this complication increasing annually[1]. Pseudotumor-related revision rates of 4% at 8 years have been reported[5], with rates

Disclosure: Arthritis Research U.K. provided one of the authors with funding in the form of a Clinical Research Fellowship. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work.

J Bone Joint Surg Am. 2016;98:1444-52 • http://dx.doi.org/10.2106/JBJS.15.01234

COPYRIGHT © 2016 BY THE JOURNAL OF BONE AND JOINT SURGERY, INCORPORATED

A commentary by Lawrence D. Dorr, MD, is linked to the online version of this article at jbjs.org.

The Effectiveness of Blood Metal Ions in Identifying Patients with Unilateral Birmingham Hip Resurfacing and Corail-Pinnacle Metal-on-Metal Hip Implants at Risk of Adverse Reactions to Metal Debris

Gulraj S. Matharu, BSc(Hons), MBChB, MRCS, MRes, Fiona Berryman, BSc(Hons), PhD, Lesley Brash, MSc, RN, Paul B. Pynsent, PhD, Ronan B.C. Treacy, FRCS(Tr&Orth), and David J. Dunlop, FRCS(Tr&Orth)

Investigation performed at the Royal Orthopaedic Hospital, Birmingham, United Kingdom

Background: We investigated whether blood metal ions could effectively identify patients with metal-on-metal hip implants with two common designs (Birmingham Hip Resurfacing [BHR] and Corail-Pinnacle) who were at risk of adverse reactions to metal debris.

Methods: This single-center, prospective study involved 598 patients with unilateral hip implants (309 patients with the BHR implant and 289 patients with the Corail-Pinnacle implant) undergoing whole blood metal ion sampling at a mean time of 6.9 years. Patients were classified into two groups, one that had adverse reactions to metal debris (those who had to undergo revision for adverse reactions to metal debris or those with adverse reactions to metal debris on imaging; n = 46) and one that did not (n = 552). Three metal ion parameters (cobalt, chromium, and cobalt-chromium ratio) were compared between groups. Optimal metal ion thresholds for identifying patients with adverse reactions to metal debris were determined using receiver operating characteristic analysis.

Results: All ion parameters were significantly higher ($p < 0.0001$) in the patients who had adverse reactions to metal debris compared with those who did not. Cobalt maximized the area under the curve for patients with the BHR implant (90.5%) and those with the Corail-Pinnacle implant (79.6%). For patients with the BHR implant, the area under the curve for cobalt was significantly greater than that for the cobalt-chromium ratio ($p = 0.0005$), but it was not significantly greater than that for chromium ($p = 0.8483$). For the patients with the Corail-Pinnacle implant, the area under the curve for cobalt was significantly greater than that for chromium ($p = 0.0004$), but it was similar to that for the cobalt-chromium ratio ($p = 0.8139$). Optimal blood metal ion thresholds for identifying adverse reactions to metal debris varied between the two different implants. When using cobalt, the optimal threshold for identifying adverse reactions to metal debris was 2.15 μg/L for the BHR group and 3.57 μg/L for the Corail-Pinnacle group. These thresholds had good sensitivities (88.5% for the BHR group and 80.0% for the Corail-Pinnacle group) and specificities (84.5% for the BHR group and 76.2% for the Corail-Pinnacle group), high negative predictive values (98.8% for the BHR group and 98.1% for the Corail-Pinnacle group), and low positive predictive values (34.3% for the BHR group and 20.0% for the Corail-Pinnacle group). The authority thresholds proposed by the United States

continued

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. It was also reviewed by an expert in methodology and statistics. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

Disclosure: Four authors of this study (F.B., P.B.P., R.B.C.T., and D.J.D.) received, through their institutions, a grant from the manufacturer of the device that is the subject of this study. One author of this study (G.S.M.) received a grant from the Royal College of Surgeons of England and the Arthritis Research Trust, which provided this author's salary in the form of a surgical research fellowship. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work and "yes" to indicate that the author had other relationships or activities that could be perceived to influence, or have the potential to influence, what was written in this work.

J Bone Joint Surg Am. 2016;98:617-26 • http://dx.doi.org/10.2106/JBJS.15.00340

THE BONE & JOINT JOURNAL

■ HIP

Can blood metal ion levels be used to identify patients with bilateral Birmingham Hip Resurfacings who are at risk of adverse reactions to metal debris?

G. S. Matharu,
F. Berryman,
L. Brash,
P. B. Pynsent,
D. J. Dunlop,
R. B. C. Treacy

From The Royal Orthopaedic Hospital, Birmingham, United Kingdom

■ G. S. Matharu, BSc (Hons), MRCS, MRes, Clinical Research Fellow and Specialist Registrar in Trauma and Orthopaedics
The Royal Orthopaedic Hospital, Birmingham B31 2AP, UK and Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK.

■ F. Berryman, BSc (Hons), PhD, Clinical Scientist
■ L. Brash, MSc, RN, Advanced Nurse Practitioner in Arthroplasty
■ D. J. Dunlop, FRCS (Tr&Orth), Consultant Orthopaedic Surgeon
■ R. B. C. Treacy, FRCS (Tr&Orth), Consultant Orthopaedic Surgeon
The Royal Orthopaedic Hospital, Birmingham B31 2AP, UK.

■ P. B. Pynsent, PhD, Honorary Senior Research Fellow, School of Clinical and Experimental Medicine
University of Birmingham, Birmingham B15 2TT, UK.

Correspondence should be sent to G. S. Matharu; e-mail: gsm@doctors.org.uk

©2016 Matharu et al
doi:10.1302/0301-620X.98B11.
38042 $2.00

Bone Joint J
2016;98-B:1455–62.
Received 12 February 2016;
Accepted after revision 29 July 2016

Aims

We investigated whether blood metal ion levels could effectively identify patients with bilateral Birmingham Hip Resurfacing (BHR) implants who have adverse reactions to metal debris (ARMD).

Patients and Methods

Metal ion levels in whole blood were measured in 185 patients with bilateral BHRs. Patients were divided into those with ARMD who either had undergone a revision for ARMD or had ARMD on imaging (n = 30), and those without ARMD (n = 155). Receiver operating characteristic analysis was used to determine the optimal thresholds of blood metal ion levels for identifying patients with ARMD.

Results

The maximum level of cobalt or chromium ions in the blood was the parameter which produced the highest area under the curve (91.0%). The optimal threshold for distinguishing between patients with and without ARMD was 5.5 µg/l (83.3% sensitivity, 88.4% specificity, 58.1% positive and 96.5% negative predictive values). Similar results were obtained in a subgroup of 111 patients who all underwent cross-sectional imaging. Between 3.2% and 4.3% of patients with ARMD were missed if United Kingdom (7 µg/l) and United States (10 µg/l) authority thresholds were used respectively, compared with 2.7% if our implant specific threshold was used, though these differences did not reach statistical significance ($p \geq 0.248$).

Conclusion

Patients with bilateral BHRs who have blood metal ion levels below our implant specific threshold were at low-risk of having ARMD.

Cite this article: ***Bone Joint J*** **2016;98-B:1455–62.**

Adverse reactions to metal debris (ARMD) have contributed to high rates of failure for most metal-on-metal (MoM) hip resurfacing designs.[1,2] As outcomes following revision arthroplasty for ARMD are reportedly poor,[3] worldwide regulatory authorities recommend regular follow-up and monitoring of these patients to identify ARMD early.[4-6] Monitoring currently includes measurement of the levels of cobalt and chromium ions in the blood, which reflects *in vivo* wear.[7] The United Kingdom Medical and Healthcare Products Regulatory Agency (MHRA) published thresholds for these levels which should be of concern in 2010, recommending cross-sectional imaging if the levels were > 7 µg/l (µg/l are equivalent to parts per billion).[8] The thresholds used to identify poorly functioning unilateral MoM hip resurfacings have ranged from 3.5 µg/l to 7 µg/l in subsequent studies, with thresholds having higher specificity than sensitivity.[9-13]

Patients with bilateral MoM hip arthroplasties are considered to be at increased risk of ARMD by some authorities.[6,14] However, little is known in these bilateral patients about the thresholds which might be suggestive of ARMD.[15] One study has assessed the thresholds for identifying patients with poorly functioning bilateral hip resurfacings and concluded that optimal threshold levels were 5.0 µg/l for cobalt and 7.4 µg/l for chromium.[12] This study included 139 patients, but with heterogenous combinations of implants, and only short-term follow-up (a mean follow-up period of 4.3 years). A recent consensus statement from the United States suggested that blood metal ion concentrations of > 10 µg/l represented a high-risk group, which could be

COPYRIGHT © 2017 BY THE JOURNAL OF BONE AND JOINT SURGERY, INCORPORATED

Patient and Radiographic Factors Help to Predict Metal-on-Metal Hip Resurfacings with Evidence of a Pseudotumor

Gulraj S. Matharu, BSc(Hons), MRCS, MRes, Oliver Blanshard, BA, Kawaljit Dhaliwal, BSc, MRCS, Andrew Judge, BSc, MSc, PhD, David W. Murray, MD, FRCS(Orth), and Hemant G. Pandit, FRCS(Tr & Orth), DPhil

Investigation performed at the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Nuffield Orthopaedic Centre, University of Oxford, Oxford, United Kingdom

Background: The role of radiographs in the follow-up of patients with metal-on-metal hip resurfacing (MoMHR) implants is unclear. We investigated whether a combination of patient and radiographic factors predicted MoMHRs with evidence of a pseudotumor.

Methods: We performed a retrospective single-center case-control study of 384 MoMHRs. The pseudotumor group of 130 hips all had evidence of a symptomatic pseudotumor on cross-sectional imaging, with the diagnosis confirmed at revision. The nonpseudotumor group of 254 hips (a subgroup of these hips were previously reported on) all had normal findings on cross-sectional imaging. Radiographs taken immediately prior to revision were assessed in the pseudotumor group and were compared with radiographs taken at the time of normal cross-sectional imaging in the nonpseudotumor group. Two blinded independent observers analyzed the radiographs for signs of failure, with excellent interobserver agreement. Logistic regression modeling identified the patient and radiographic predictors of revision for pseudotumor.

Results: Hips with a pseudotumor more commonly had abnormal findings on radiographs compared to hips without a pseudotumor (80.0% compared with 63.4%; p = 0.001). Patient and radiographic factors predictive of revision for pseudotumor in the multivariable model were female sex (odds ratio [OR], 3.14; 95% confidence interval [CI], 1.85 to 5.35; p < 0.001), high inclination (OR, 1.04 per degree; 95% CI, 1.01 to 1.07 per degree; p = 0.006), acetabular osteolysis (OR, 5.06; 95% CI, 2.14 to 12.0; p < 0.001), femoral osteolysis (OR, 17.8; 95% CI, 5.09 to 62.2; p < 0.001), and acetabular loosening (OR, 3.35; 95% CI, 1.34 to 8.35; p = 0.009). Factors predictive of not having a pseudotumor were anteversion of ≥5° (5° to <10°: OR, 0.31; 95% CI, 0.12 to 0.77; p = 0.012; and ≥10°: OR, 0.32; 95% CI, 0.15 to 0.70; p = 0.004) and heterotopic ossification (OR, 0.19; 95% CI, 0.05 to 0.72; p = 0.015). The final multivariable model was well calibrated (p = 0.589), with good discriminatory ability (area under the curve = 0.801; sensitivity = 74.4%; specificity = 71.7%).

Conclusions: A combination of patient and radiographic factors provided useful information for distinguishing between MoMHRs with and without evidence of a pseudotumor. Surgeons may wish to consider these specific patient and radiographic factors before proceeding with cross-sectional imaging. Radiographs are important when assessing patients with MoMHR implants and should be included in all follow-up protocols.

Level of Evidence: Diagnostic Level III. See Instructions for Authors for a complete description of levels of evidence.

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

High short-term failure rates are reported for most metal-on-metal hip resurfacing (MoMHR) designs[1-3], with many revisions performed following pseudotumor formation[4,5]. In an attempt to identify pseudotumors early, worldwide medical regulatory authorities recommend regular surveillance for most patients with MoMHR implants[6-8].

Disclosure: One of the authors received funding from Arthritis Research UK and the Royal College of Surgeons of England in support of this research. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work.

J Bone Joint Surg Am. 2017;99:214-22 • http://dx.doi.org/10.2106/JBJS.16.00212

THE BONE & JOINT JOURNAL

■ HIP

Which imaging modality is most effective for identifying pseudotumours in metal-on-metal hip resurfacings requiring revision

ULTRASOUND OR MARS-MRI OR BOTH?

G. S. Matharu,
R. Mansour,
O. Dada,
S. Ostlere,
H. G. Pandit,
D. W. Murray

From NDORMS, University of Oxford, Oxford, United Kingdom

Aims

The aims of this study were to compare the diagnostic test characteristics of ultrasound alone, metal artefact reduction sequence MRI (MARS-MRI) alone, and ultrasound combined with MARS-MRI for identifying intra-operative pseudotumours in metal-on-metal hip resurfacing (MoMHR) patients undergoing revision surgery.

Methods

This retrospective diagnostic accuracy study involved 39 patients (40 MoMHRs). The time between imaging modalities was a mean of 14.6 days (0 to 90), with imaging performed at a mean of 5.3 months (0.06 to 12) before revision. The prevalence of intra-operative pseudotumours was 82.5% (n = 33).

Results

Agreement with the intra-operative findings was 82.5% (n = 33) for ultrasound alone, 87.5% (n = 35) for MARS-MRI alone, and 92.5% (n = 37) for ultrasound and MARS-MRI combined. The diagnostic characteristics for ultrasound alone and MARS-MRI alone reached similar sensitivities (90.9% *vs* 93.9%) and positive predictive values (PPVs; 88.2% *vs* 91.2%), but higher specificities (57.1% *vs* 42.9%) and negative predictive values (NPVs; 66.7% *vs* 50.0%) were achieved with MARS-MRI. Ultrasound and MARS-MRI combined produced 100% sensitivity and 100% NPV, whilst maintaining both specificity (57.1%) and PPV (91.7%).

For the identification of a pseudotumour, which was confirmed at revision surgery, agreement was substantial for ultrasound and MARS-MRI combined (κ = 0.69), moderate for MARS-MRI alone (κ = 0.54), and fair for ultrasound alone (κ = 0.36).

Discussion

These findings suggest that ultrasound and/or MARS-MRI have a role when assessing patients with a MoMHR, with the choice dependent on local financial constraints and the availability of ultrasound expertise. However in patients with a MoMHR who require revision, combined imaging was most effective.

Take home message: Combined imaging with ultrasound and MARS-MRI always identified intra-operative pseudotumours if present. Furthermore, if neither imaging modality showed a pseudotumour, one was not found intra-operatively.

Cite this article: *Bone Joint J* **2016;98-B:40–8.**

More than one million patients worldwide have received large-diameter metal-on-metal (MoM) arthroplasties of the hip.[1] High short-term rates of failure owing to pseudotumour formation have been reported with many MoM hip designs.[2,3] Furthermore, poor outcomes from revision surgery following the development of a pseudotumour have been reported,[4,5] which is a concern given most patients with these articulations are young and active.[6,7]

Currently most patients with MoM hip arthroplasties are reviewed regularly as recommended by worldwide regulatory authorities, as it is perceived that early detection and revision surgery for pseudotumour may improve the outcome.[8-10] All authorities recommend either ultrasound or metal artefact reduction sequence magnetic resonance imaging (MARS-MRI) for assessing the periprosthetic soft-tissues around MoM hips.[8-10] However, no preference between these modalities has been

■ G. S. Matharu, BSc (Hons), MRCS, MRes, Clinical Research Fellow in Orthopaedics, NDORMS
■ R. Mansour, FRCR, Consultant Musculoskeletal Radiologist, Nuffield Orthopaedic Centre
■ O. Dada, BSc, Medical Student, NDORMS
■ S. Ostlere, FRCP, FRCR, Consultant Musculoskeletal Radiologist, Nuffield Orthopaedic Centre
■ H. G. Pandit, DPhil, FRCS (Orth), Orthopaedic Surgeon and Associate Professor, NDORMS and Nuffield Orthopaedic Centre
■ D. W. Murray, MD, FRCS (Orth), Consultant Orthopaedic Surgeon and Professor of Orthopaedic Surgery NDORMS and Nuffield Orthopaedic Centre, Oxford, UK.

Correspondence should be sent to Mr G. S. Matharu; e-mail: gsm@doctors.org.uk

©2016 The British Editorial Society of Bone & Joint Surgery
doi:10.1302/0301-620X.98B1.36746 $2.00

Bone Joint J 2016;98-B:40–8.
Received 2 July 2015; Accepted after revision 3 September 2015

The Journal of Arthroplasty 31 (2016) 146–151

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

How Should We Follow-Up Asymptomatic Metal-on-Metal Hip Resurfacing Patients? A Prospective Longitudinal Cohort Study

Adrian K. Low, MBBS, PhD, FRACS, Gulraj S. Matharu, BSc (Hons), MRCS, MRes, Simon J. Ostlere, FRCP, FRCR, David W. Murray, MD, FRCS (Orth), Hemant G. Pandit, DPhil, FRCS (Orth)

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Oxford, United Kingdom, OX3 7LD

ARTICLE INFO

Article history:
Received 25 May 2015
Accepted 11 August 2015

Keywords:
asymptomatic
blood metal ions
follow-up
hip resurfacing
metal-on-metal
ultrasound

ABSTRACT

Current surveillance for metal-on-metal hip resurfacing (MoMHR) patients is not evidence based. This study established changes that occurred in 152 asymptomatic MoMHRs using repeat ultrasound and patient-reported outcomes. Factors associated with (1) ultrasound progression and (2) developing new pseudotumors were analyzed. Patients underwent repeat assessments 4.3 years later. Ultrasound progression was observed in 19% (n = 29), with 10% (n = 15) developing new pseudotumors. Key predictors of ultrasound progression included high blood cobalt ($P = .00013$) and chromium ($P = .00065$), and high initial ultrasound grade ($P = .003$) and volume ($P = .036$). No asymptomatic MoMHRs with initially normal metal ions (<2 μg/L) and normal ultrasounds (33% of cohort) developed new pseudotumors. This patient subgroup does not require repeat follow-up within 5 years.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Numerous metal-on-metal (MoM) hip arthroplasty designs have experienced high failure rates due to pseudotumors [1,2]. Patients developing these problems often require revision surgery. As lesions can be destructive with significant bone and muscle damage, outcomes after revision can be poor [3,4]. To identify patients with pseudotumors early, regulatory authorities have published guidance regarding the regular follow-up of MoM hip patients [5–7].

Presently, there is no consensus on how to follow up asymptomatic MoM hip resurfacing (MoMHR) patients, with this patient subgroup being the most difficult to manage clinically [8,9]. European guidance recommends annual follow-up with radiographs and blood metal ions in these patients. However, the US Food and Drug Administration guidance recommends annual clinical review, whereas the Medical and Healthcare Products Regulatory Agency in the United Kingdom recommend reviewing asymptomatic MoMHR patients according to local protocol. Recent work has demonstrated that worldwide MoM hip follow-up guidance is neither evidence based nor financially sustainable, with most protocols lacking the sensitivity to detect asymptomatic pseudotumors [10].

Most MoMHR patients remain asymptomatic and do not develop pseudotumors [11,12]. However, a small but significant number of asymptomatic patients do develop pseudotumors [8,9,13], and it is important to identify these individuals early in order to prevent bone and soft tissue damage. At present, there is no clear guidance as to which parameters should be used to distinguish asymptomatic MoMHR patients who can be safely discharged and need not be subjected to repeated investigations from those asymptomatic MoMHR patients who need monitoring. Decisions regarding which asymptomatic MoMHR patients require monitoring, at what intervals, and how frequently such follow-up should be repeated require well-designed prospective longitudinal cohort studies. At present, very few such studies exist involving MoMHR patients, with most reporting on serial magnetic resonance imaging (MRI) at short-term follow-up in small cohorts [14,15]. Ultrasound is another commonly used and recommended modality for cross-sectional imaging in MoM hip patients [5,16]. It has many advantages over MRI: it is cheaper, faster to perform, and not affected by metal artifact. Furthermore, ultrasound is sensitive when screening for pseudotumors [16] with results comparable to MRI [17], and recent work suggests that ultrasound has a role in the surveillance of asymptomatic pseudotumors [18].

This article reports on a prospective cohort of 152 asymptomatic MoMHRs who were comprehensively assessed between 2007 and 2008 [8]. Ultrasound and patient-reported outcomes were repeated at a mean of 4.3 years since initial assessment and at a mean of 8.2 years since primary MoMHR. The present study aimed to assess factors associated with (1) ultrasound finding progression and (2) developing new pseudotumors. This information will assist in risk stratifying patients,

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to http://dx.doi.org/10.1016/j.arth.2015.08.007.

Reprint requests: Gulraj S. Matharu, BSc (Hons), MRCS, MRes, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Oxford, United Kingdom, OX3 7LD.

http://dx.doi.org/10.1016/j.arth.2015.08.007
0883-5403/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

BIBLIOGRAPHY

1. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. *Lancet*. 2007;**370**(9597):1508-1519.

2. National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. 13th Annual Report. 2016. http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Reports/13th Annual Report/07950%07920NJR%07920Annual%07920Report 2016 2020ONLINE 2020REPORT.pdf.

3. Culliford D, Maskell J, Judge A, Cooper C, Prieto-Alhambra D, Arden NK; COASt Study Group. Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink. *Osteoarthritis Cartilage*. 2015;**23**(4):594-600.

4. Callaghan JJ, Forest EE, Olejniczak JP, Goetz DD, Johnston RC. Charnley total hip arthroplasty in patients less than fifty years old. A twenty to twenty-five-year follow-up note. *J Bone Joint Surg Am*. 1998;**80**(5):704-714.

5. Makela KT, Eskelinen A, Pulkkinen P, Paavolainen P, Remes V. Results of 3,668 primary total hip replacements for primary osteoarthritis in patients under the age of 55 years. *Acta Orthop*. 2011;**82**(5):521-529.

6. Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, Cooper C, Carr AJ, Arden NK, Beard DJ, Price AJ. The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *Lancet*. 2017;**389**(10077):1424-1430.

7. Fisher J, Jin Z, Tipper J, Stone M, Ingham E. Tribology of alternative bearings. *Clin Orthop Relat Res*. 2006;**453**:25-34.

8. Bozic KJ, Kurtz S, Lau E, Ong K, Chiu V, Vail TP, Rubash HE, Berry DJ. The epidemiology of bearing surface usage in total hip arthroplasty in the United States. *J Bone Joint Surg Am*. 2009;**91**(7):1614-1620.

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

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

11. Australian Orthopaedic Association National Joint Replacement Registry. Hip, Knee & Shoulder Arthroplasty. Annual Report 2016. https://aoanjrr.sahmri.com/annual-reports-2016.

12. Grammatopoulos G, Pandit H, Kwon YM, Gundle R, McLardy-Smith P, Beard DJ, Murray DW, Gill HS. Hip resurfacings revised for inflammatory pseudotumour have a poor outcome. *J Bone Joint Surg Br*. 2009;**91**(8):1019-1024.

13. Daniel J, Pradhan C, Ziaee H, Pynsent PB, McMinn DJ. Results of Birmingham hip resurfacing at 12 to 15 years: a single-surgeon series. *Bone Joint J*. 2014;**96-B**(10):1298-1306.

14. Langton DJ, Jameson SS, Joyce TJ, Hallab NJ, Natu S, Nargol AV. 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.

15. Langton DJ, Jameson SS, Joyce TJ, Gandhi JN, Sidaginamale R, Mereddy P, Lord J, Nargol AV. Accelerating failure rate of the ASR total hip replacement. *J Bone Joint Surg Br*. 2011;**93**(8):1011-1016.

16. Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gibbons CL, Ostlere S, Athanasou N, Gill HS, Murray DW. Pseudotumours associated with metal-on-metal hip resurfacings. *J Bone Joint Surg Br*. 2008;**90**(7):847-851.

17. Kwon YM, Glyn-Jones S, Simpson DJ, Kamali A, McLardy-Smith P, Gill HS, Murray DW. Analysis of wear of retrieved metal-on-metal hip resurfacing implants revised due to pseudotumours. *J Bone Joint Surg Br*. 2010;**92**(3):356-361.

18. Matthies A, Underwood R, Cann P, Ilo K, Nawaz Z, Skinner J, Hart AJ. Retrieval analysis of 240 metal-on-metal hip components, comparing modular total hip replacement with hip resurfacing. *J Bone Joint Surg Br*. 2011;**93**(3):307-314.

19. Langton DJ, Sidaginamale R, Lord JK, Nargol AV, Joyce TJ. Taper junction failure in large-diameter metal-on-metal bearings. *Bone Joint Res*. 2012;**1**(4):56-63.

20. Langton DJ, Joyce TJ, Jameson SS, Lord J, Van Orsouw M, Holland JP, Nargol AV, De Smet KA. Adverse reaction to metal debris following hip resurfacing: the influence of component type, orientation and volumetric wear. *J Bone Joint Surg Br*. 2011;**93**(2):164-171.

21. Jacobs JJ, Cooper HJ, Urban RM, Wixson RL, Della Valle CJ. What do we know about taper corrosion in total hip arthroplasty? *J Arthroplasty*. 2014;**29**(4):668-669.

22. Hannemann F, Hartmann A, Schmitt J, Lutzner J, Seidler A, Campbell P, Delaunay CP, Drexler H, Ettema HB, Garcia-Cimbrelo E, Huberti H, Knahr K, Kunze J, Langton DJ, Lauer W, Learmonth I, Lohmann CH, Morlock M, Wimmer MA, Zagra L, Gunther KP. European multidisciplinary consensus statement on the use and monitoring of metal-on-metal bearings for total hip replacement and hip resurfacing. *Orthop Traumatol Surg Res*. 2013;**99**(3):263-271.

23. MHRA. Medical Device Alert: all metal-on-metal (MoM) hip replacements. MDA/2012/036. 2012. https://www.gov.uk/drug-device-alerts/medical-device-alert-metal-on-metal-mom-hip-replacements-updated-advice-with-patient-follow-ups.

24. U.S. Food and Drug Administration. Medical Devices: Metal-on-Metal Hip Implants. Information for Orthopaedic Surgeons. 2013. http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/MetalonMetalHipImplants/ucm241667.htm.

25. Hart AJ, Satchithananda K, Liddle AD, Sabah SA, McRobbie D, Henckel J, Cobb JP, Skinner JA, Mitchell AW. Pseudotumors in association with well-functioning metal-on-metal hip prostheses: a case-control study using three-dimensional computed tomography and magnetic resonance imaging. *J Bone Joint Surg Am*. 2012;**94**(4):317-325.

26. Fehring TK, Odum S, Sproul R, Weathersbee J. High frequency of adverse local tissue reactions in asymptomatic patients with metal-on-metal THA. *Clin Orthop Relat Res*. 2014;**472**(2):517-522.

27. Matharu GS, Mellon SJ, Murray DW, Pandit HG. Follow-Up of Metal-on-Metal Hip Arthroplasty Patients Is Currently Not Evidence Based or Cost Effective. *J Arthroplasty*. 2015;**30**(8):1317-1323.

28. Matharu GS, Pynsent PB, Dunlop DJ. Revision of metal-on-metal hip replacements and resurfacings for adverse reaction to metal debris: a systematic review of outcomes. *Hip Int*. 2014;**24**(4):311-320.

29. De Smet KA, Van Der Straeten C, Van Orsouw M, Doubi R, Backers K, Grammatopoulos G. Revisions of metal-on-metal hip resurfacing: lessons learned and improved outcome. *Orthop Clin North Am*. 2011;**42**(2):259-269, ix.

30. Revell PA. The combined role of wear particles, macrophages and lymphocytes in the loosening of total joint prostheses. *J R Soc Interface*. 2008;**5**(28):1263-1278.

31. Smith-Petersen MN. Evolution of mould arthroplasty of the hip joint. *J Bone Joint Surg Br*. 1948;**30B**(1):59-75.

32. Charnley J. Arthroplasty of the hip. A new operation. *Lancet*. 1961;**1**(7187):1129-1132.

33. Haboush EJ. A new operation for arthroplasty of the hip based on biomechanics, photoelasticity, fast-setting dental acrylic, and other considerations. *Bull Hosp Joint Dis*. 1953;**14**(2):242-277.

34. Muller ME. Lessons of 30 years of total hip arthroplasty. *Clin Orthop Relat Res*. 1992;(274):12-21.

35. Freeman MA, Bradley GW. ICLH surface replacement of the hip. An analysis of the first 10 years. *J Bone Joint Surg Br*. 1983;**65**(4):405-411.

36. Wagner H. Surface replacement arthroplasty of the hip. *Clin Orthop Relat Res*. 1978;(134):102-130.

37. Amstutz HC, Clarke IC, Christie J, Graff-Radford A. Total hip articular replacement by internal eccentric shells: the "tharies" approach to total surface replacement arthroplasty. *Clin Orthop Relat Res*. 1977;(128):261-284.

38. Grigoris P, Roberts P, Panousis K, Jin Z. Hip resurfacing arthroplasty: the evolution of contemporary designs. *Proc Inst Mech Eng H*. 2006;**220**(2):95-105.

39. Muller ME. The benefits of metal-on-metal total hip replacements. *Clin Orthop Relat Res*. 1995;(311):54-59.

40. Andrew TA, Berridge D, Thomas A, Duke RN. Long-term review of ring total hip arthroplasty. *Clin Orthop Relat Res*. 1985;(201):111-122.

41. August AC, Aldam CH, Pynsent PB. The McKee-Farrar hip arthroplasty. A long-term study. *J Bone Joint Surg Br*. 1986;**68**(4):520-527.

42. Wilson JN, Scales JT. The Stanmore metal on metal total hip prosthesis using a three pin type cup. A follow-up of 100 arthroplasties over nine years. *Clin Orthop Relat Res*. 1973;(95):239-249.

43. McMinn DJW. Modern Hip Resurfacing. 1st edition. London, UK: Springer; 2009.

44. McMinn D, Treacy R, Lin K, Pynsent P. Metal on metal surface replacement of the hip. Experience of the McMinn prothesis. *Clin Orthop Relat Res*. 1996;(329 Suppl):S89-98.

45. Wagner M, Wagner H. Preliminary results of uncemented metal on metal stemmed and resurfacing hip replacement arthroplasty. *Clin Orthop Relat Res*. 1996;(329 Suppl):S78-88.

46. Amstutz HC. Hip Resurfacing: Principles, Indications, Technique and Results. 1st edition. Philadelphia, USA: Saunders; 2008.

47. Jones LC, Haggard WO, Greenwald AS. Metal-on-Metal Bearings. A Clinical Practicum. 1st edition. New York, USA: Springer; 2014.

48. Haddad FS, Thakrar RR, Hart AJ, Skinner JA, Nargol AV, Nolan JF, Gill HS, Murray DW, Blom AW, Case CP. Metal-on-metal bearings: the evidence so far. *J Bone Joint Surg Br*. 2011;**93**(5):572-579.

49. Girard J, Lavigne M, Vendittoli PA, Roy AG. Biomechanical reconstruction of the hip: a randomised study comparing total hip resurfacing and total hip arthroplasty. *J Bone Joint Surg Br*. 2006;**88**(6):721-726.

50. Little JP, Taddei F, Viceconti M, Murray DW, Gill HS. Changes in femur stress after hip resurfacing arthroplasty: response to physiological loads. *Clin Biomech (Bristol, Avon)*. 2007;**22**(4):440-448.

51. Kishida Y, Sugano N, Nishii T, Miki H, Yamaguchi K, Yoshikawa H. Preservation of the bone mineral density of the femur after surface replacement of the hip. *J Bone Joint Surg Br*. 2004;**86(**2):185-189.

52. Silva M, Lee KH, Heisel C, Dela Rosa MA, Schmalzried TP. The biomechanical results of total hip resurfacing arthroplasty. *J Bone Joint Surg Am*. 2004;**86-A**(1):40-46.

53. Ball ST, Le Duff MJ, Amstutz HC. Early results of conversion of a failed femoral component in hip resurfacing arthroplasty. *J Bone Joint Surg Am*. 2007;**89**(4):735-741.

54. Amstutz HC, Beaule PE, Dorey FJ, Le Duff MJ, Campbell PA, Gruen TA. Metal-on-metal hybrid surface arthroplasty: two to six-year follow-up study. *J Bone Joint Surg Am*. 2004;**86-A**(1):28-39.

55. 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**(2):177-184.

56. Treacy RB, McBryde CW, Pynsent PB. Birmingham hip resurfacing arthroplasty. A minimum follow-up of five years. *J Bone Joint Surg Br*. 2005;**87**(2):167-170.

57. Hing CB, Back DL, Bailey M, Young DA, Dalziel RE, Shimmin AJ. The results of primary Birmingham hip resurfacings at a mean of five years. An independent prospective review of the first 230 hips. *J Bone Joint Surg Br*. 2007;**89(**11):1431-1438.

58. Steffen RT, Pandit HP, Palan J, Beard DJ, Gundle R, McLardy-Smith P, Murray DW, Gill HS. The five-year results of the Birmingham Hip Resurfacing arthroplasty: an independent series. *J Bone Joint Surg Br*. 2008;**90**(4):436-441.

59. Heilpern GN, Shah NN, Fordyce MJ. 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):1137-1142.

60. Khan M, Kuiper JH, Edwards D, Robinson E, Richardson JB. Birmingham hip arthroplasty: five to eight years of prospective multicenter results. *J Arthroplasty*. 2009;**24**(7):1044-1050.

61. Ollivere B, Duckett S, August A, Porteous M. The Birmingham Hip Resurfacing: 5-year clinical and radiographic results from a District General Hospital. *Int Orthop*. 2010;**34**(5):631-634.

62. Bolland BJ, Culliford DJ, Langton DJ, Millington JP, Arden NK, Latham JM. High failure rates with a large-diameter hybrid metal-on-metal total hip replacement: clinical, radiological and retrieval analysis. *J Bone Joint Surg Br*. 2011;**93**(5):608-615.

63. Matharu GS, McBryde CW, Revell MP, Pynsent PB. Femoral neck fracture after Birmingham Hip Resurfacing Arthroplasty: prevalence, time to fracture, and outcome after revision. *J Arthroplasty*. 2013;**28**(1):147-153.

64. Nunley RM, Zhu J, Brooks PJ, Engh CA, Jr., Raterman SJ, Rogerson JS, Barrack RL. The learning curve for adopting hip resurfacing among hip specialists. *Clin Orthop Relat Res*. 2010;**468**(2):382-391.

65. National Institute for Health and Care Excellence. Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip. Technology appraisal guidance [TA304]. 2014. https://www.nice.org.uk/guidance/ta304.

66. Seppanen M, Karvonen M, Virolainen P, Remes V, Pulkkinen P, Eskelinen A, Liukas A, Makela KT. Poor 10-year survivorship of hip resurfacing arthroplasty. *Acta Orthop*. 2016;**87**(6):554-559.

67. MHRA. Medical Device Alert: ASR hip replacement implant manufactured by DePuy International Ltd. 2010. http:// www.mhra.gov.uk.

68. Matharu GS, McBryde CW, Pynsent WB, Pynsent PB, Treacy RB. The outcome of the Birmingham Hip Resurfacing in patients aged < 50 years up to 14 years post-operatively. *Bone Joint J*. 2013;**95-B**(9):1172-1177.

69. Murray DW, Grammatopoulos G, Pandit H, Gundle R, Gill HS, McLardy-Smith P. The ten-year survival of the Birmingham hip resurfacing: an independent series. *J Bone Joint Surg Br*. 2012;**94**(9):1180-1186.

70. Coulter G, Young DA, Dalziel RE, Shimmin AJ. Birmingham hip resurfacing at a mean of ten years: results from an independent centre. *J Bone Joint Surg Br*. 2012;**94**(3):315-321.

71. Holland JP, Langton DJ, Hashmi M. Ten-year clinical, radiological and metal ion analysis of the Birmingham Hip Resurfacing: from a single, non-designer surgeon. *J Bone Joint Surg Br*. 2012;**94**(4):471-476.

72. Van Der Straeten C, Van Quickenborne D, De Roest B, Calistri A, Victor J, De Smet K. Metal ion levels from well-functioning Birmingham Hip Resurfacings decline significantly at ten years. *Bone Joint J*. 2013;**95-B**(10):1332-1338.

73. Reito A, Puolakka T, Elo P, Pajamaki J, Eskelinen A. Outcome of Birmingham hip resurfacing at ten years: role of routine whole blood metal ion measurements in screening for pseudotumours. *Int Orthop*. 2014;**38**(11):2251-2257.

74. Mehra A, Berryman F, Matharu GS, Pynsent PB, Isbister ES. Birmingham Hip Resurfacing: A Single Surgeon Series Reported at a Minimum of 10 Years Follow-Up. *J Arthroplasty*. 2015;**30**(7):1160-1166.

75. Azam MQ, McMahon S, Hawdon G, Sankineani SR. Survivorship and clinical outcome of Birmingham hip resurfacing: a minimum ten years' follow-up. *Int Orthop*. 2016;**40**(1):1-7.

76. Carrothers AD, Gilbert RE, Jaiswal A, Richardson JB. Birmingham hip resurfacing: the prevalence of failure. *J Bone Joint Surg Br*. 2010;**92**(10):1344-1350.

77. Shimmin AJ, Back D. Femoral neck fractures following Birmingham hip resurfacing: a national review of 50 cases. *J Bone Joint Surg Br*. 2005;**87**(4):463-464.

78. Zustin J, Krause M, Breer S, Hahn M, von Domarus C, Ruther W, Sauter G, Morlock MM, Amling M. Morphologic analysis of periprosthetic fractures after hip resurfacing arthroplasty. *J Bone Joint Surg Am*. 2010;**92**(2):404-410.

79. Steffen RT, Athanasou NA, Gill HS, Murray DW. Avascular necrosis associated with fracture of the femoral neck after hip resurfacing: histological assessment of femoral bone from retrieval specimens. *J Bone Joint Surg Br*. 2010;**92**(6):787-793.

80. Davis ET, Olsen M, Zdero R, Papini M, Waddell JP, Schemitsch EH. A biomechanical and finite element analysis of femoral neck notching during hip resurfacing. *J Biomech Eng*. 2009;**131**(4):041002.

81. Gilbert RE, Cheung G, Carrothers AD, Meyer C, Richardson JB. Functional results of isolated femoral revision of hip resurfacing arthroplasty. *J Bone Joint Surg Am*. 2010;**92**(7):1600-1604.

82. Willert HG, Buchhorn GH, Fayyazi A, Flury R, Windler M, Koster G, Lohmann CH. 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.

83. 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.

84. Gruber FW, Bock 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):923-927.

85. Matthies AK, Racasan R, Bills P, Blunt L, Cro S, Panagiotidou A, Blunn G, Skinner J, Hart AJ. Material loss at the taper junction of retrieved large head metal-on-metal total hip replacements. *J Orthop Res*. 2013;**31**(11):1677-1685.

86. Cooper HJ, Della Valle CJ, Berger RA, Tetreault M, Paprosky WG, Sporer SM, Jacobs JJ. Corrosion at the head-neck taper as a cause for adverse local tissue reactions after total hip arthroplasty. *J Bone Joint Surg Am*. 2012;**94**(18):1655-1661.

87. Whitehouse MR, Endo M, Zachara S, Nielsen TO, Greidanus NV, Masri BA, Garbuz DS, Duncan CP. Adverse local tissue reactions in metal-on-polyethylene total hip arthroplasty due to trunnion corrosion: the risk of misdiagnosis. *Bone Joint J*. 2015;**97-B**(8):1024-1030.

88. Berend KR, Morris MJ, Adams JB, Lombardi AV, Jr. Metal-on-metal hip arthroplasty: going, going, gone... - affirms. *J Bone Joint Surg Br*. 2012;**94**(11 Suppl A):75-77.

89. 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**(8):1025-1030.

90. Ebreo D, Khan A, El-Meligy M, Armstrong C, Peter V. Metal ion levels decrease after revision for metallosis arising from large-diameter metal-on-metal hip arthroplasty. *Acta Orthop Belg*. 2011;**77**(6):777-781.

91. Engh CA, Jr., Ho H, Engh CA. Metal-on-metal hip arthroplasty: does early clinical outcome justify the chance of an adverse local tissue reaction? *Clin Orthop Relat Res*. 2010;**468**(2):406-412.

92. Schmalzried TP. The future of hip resurfacing. *Orthop Clin North Am*. 2011;**42**(2):271-273, ix-x.

93. Plummer DR, Berger RA, Paprosky WG, Sporer SM, Jacobs JJ, Della Valle CJ. Diagnosis and Management of Adverse Local Tissue Reactions Secondary to Corrosion at the Head-Neck Junction in Patients With Metal on Polyethylene Bearings. *J Arthroplasty*. 2016;**31**(1):264-268.

94. Mittal S, Revell M, Barone F, Hardie DL, Matharu GS, Davenport AJ, Martin RA, Grant M, Mosselmans F, Pynsent P, Sumathi VP, Addison O, Revell PA, Buckley CD. Lymphoid aggregates that resemble tertiary lymphoid organs define a specific pathological subset in metal-on-metal hip replacements. *PLoS One*. 2013;**8**(5):e63470.

95. Natu S, Sidaginamale RP, Gandhi J, Langton DJ, Nargol AV. Adverse reactions to metal debris: histopathological features of periprosthetic soft tissue reactions seen in association with failed metal on metal hip arthroplasties. *J Clin Pathol*. 2012;**65**(5):409-418.

96. Sabah SA, Henckel J, Cook E, Whittaker R, Hothi H, Pappas Y, Blunn G, Skinner JA, Hart AJ. Validation of primary metal-on-metal hip arthroplasties on the National Joint Registry for England, Wales and Northern Ireland using data from the London Implant Retrieval Centre: a study using the NJR dataset. *Bone Joint J*. 2015;**97-B**(1):10-18.

97. Langton D, Ahmed I, Avery P, Bone M, Cooke N, Deehan D, Duffy P, Foguet P, Green S, Holland J, Jafri A, Longstaff L, Lord J, Loughead J, Meek RM, Murray H, Nanu A, Nargol AV, Scholes SC, Sidaginamale RP, Waller S, Joyce T. Investigation of Taper Failure in a Contemporary Metal-on-Metal Hip Arthroplasty System Through Examination of Unused and Explanted Prostheses. *J Bone Joint Surg Am*. 2017;**99**(5):427-436.

98. Campbell P, Ebramzadeh E, Nelson S, Takamura K, De Smet K, Amstutz HC. Histological features of pseudotumor-like tissues from metal-on-metal hips. *Clin Orthop Relat Res*. 2010;**468**(9):2321-2327.

99. Grammatopoulos G, Pandit H, Kamali A, Maggiani F, Glyn-Jones S, Gill HS, Murray DW, Athanasou N. The correlation of wear with histological features after failed hip resurfacing arthroplasty. *J Bone Joint Surg Am*. 2013;**95**(12):e81.

100. Kwon YM, Thomas P, Summer B, Pandit H, Taylor A, Beard D, Murray DW, Gill HS. Lymphocyte proliferation responses in patients with pseudotumors following metal-on-metal hip resurfacing arthroplasty. *J Orthop Res*. 2010;**28**(4):444-450.

101. Glyn-Jones S, Roques A, Taylor A, Kwon YM, McLardy-Smith P, Gill HS, Walter W, Tuke M, Murray D. The in vivo linear and volumetric wear of hip resurfacing implants revised for pseudotumor. *J Bone Joint Surg Am*. 2011;**93**(23):2180-2188.

102. Langton DJ, Sidaginamale RP, Avery P, Waller S, Tank G, Lord J, Joyce T, Cooke N, Logishetty R, Nargol AV. Retrospective cohort study of the performance of the Pinnacle metal on metal (MoM) total hip replacement: a single-centre investigation in combination with the findings of a national retrieval centre. *BMJ Open*. 2016;**6**(4):e007847.

103. Rodrigues DC, Urban RM, Jacobs JJ, Gilbert JL. In vivo severe corrosion and hydrogen embrittlement of retrieved modular body titanium alloy hip-implants. *J Biomed Mater Res B Appl Biomater*. 2009;**88**(1):206-219.

104. Lucas LC, Buchanan RA, Lemons JE. Investigations on the galvanic corrosion of multialloy total hip prostheses. *J Biomed Mater Res*. 1981;**15**(5):731-747.

105. Athanasou NA. The pathobiology and pathology of aseptic implant failure. *Bone Joint Res*. 2016;**5**(5):162-168.

106. Neyt K, Perros F, GeurtsvanKessel CH, Hammad H, Lambrecht BN. Tertiary lymphoid organs in infection and autoimmunity. *Trends Immunol*. 2012;**33**(6):297-305.

107. Aloisi F, Pujol-Borrell R. Lymphoid neogenesis in chronic inflammatory diseases. *Nat Rev Immunol*. 2006;**6**(3):205-217.

108. Sidaginamale RP, Joyce TJ, Bowsher JG, Lord JK, Avery PJ, Natu S, Nargol AV, Langton DJ. The clinical implications of metal debris release from the taper junctions and bearing surfaces of metal-on-metal hip arthroplasty: joint fluid and blood metal ion concentrations. *Bone Joint J*. 2016;**98-B**(7):925-933.

109. Almousa SA, Greidanus NV, Masri BA, Duncan CP, Garbuz DS. The natural history of inflammatory pseudotumors in asymptomatic patients after metal-on-metal hip arthroplasty. *Clin Orthop Relat Res*. 2013;**471**(12):3814-3821.

110. Matharu GS, Ostlere SJ, Pandit HG, Murray DW. What is the natural history of asymptomatic pseudotumours in metal-on-metal hip resurfacing patients? *Hip Int*. 2016;**26**(6):522-530.

111. Bosker BH, Ettema HB, van Rossum M, Boomsma MF, Kollen BJ, Maas M, Verheyen CC. Pseudotumor formation and serum ions after large head metal-on-metal stemmed total hip replacement. Risk factors, time course and revisions in 706 hips. *Arch Orthop Trauma Surg*. 2015;**135**(3):417-425.

112. Kwon YM, Lombardi AV, Jacobs JJ, Fehring TK, Lewis CG, Cabanela ME. Risk stratification algorithm for management of patients with metal-on-metal hip arthroplasty: consensus statement of the American Association of Hip and Knee

Surgeons, the American Academy of Orthopaedic Surgeons, and the Hip Society. *J Bone Joint Surg Am*. 2014;**96**(1):e4.

113. De Haan R, Campbell PA, Su EP, De Smet KA. Revision of metal-on-metal resurfacing arthroplasty of the hip: the influence of malpositioning of the components. *J Bone Joint Surg Br*. 2008;**90**(9):1158-1163.

114. Memon AR, Galbraith JG, Harty JA, Gul R. Inflammatory pseudotumor causing deep vein thrombosis after metal-on-metal hip resurfacing arthroplasty. *J Arthroplasty*. 2013;**28**(1):197.e9-12.

115. Parfitt DJ, Wood SN, Chick CM, Lewis P, Rashid MH, Evans AR. Common femoral vein thrombosis caused by a metal-on-metal hip arthroplasty-related pseudotumor. *J Arthroplasty*. 2012;**27**(8):1581. e9-1581. e11.

116. Zywiel MG, Cherian JJ, Banerjee S, Cheung AC, Wong F, Butany J, Gilbert C, Overgaard C, Syed K, Jacobs JJ, Mont MA. Systemic cobalt toxicity from total hip arthroplasties: review of a rare condition Part 2. measurement, risk factors, and step-wise approach to treatment. *Bone Joint J*. 2016;**98-B**(1):14-20.

117. Cheung AC, Banerjee S, Cherian JJ, Wong F, Butany J, Gilbert C, Overgaard C, Syed K, Zywiel MG, Jacobs JJ, Mont MA. Systemic cobalt toxicity from total hip arthroplasties: review of a rare condition Part 1 - history, mechanism, measurements, and pathophysiology. *Bone Joint J*. 2016;**98-B**(1):6-13.

118. Tower SS. Arthroprosthetic cobaltism: neurological and cardiac manifestations in two patients with metal-on-metal arthroplasty: a case report. *J Bone Joint Surg Am*. 2010;**92**(17):2847-2851.

119. Bradberry SM, Wilkinson JM, Ferner RE. Systemic toxicity related to metal hip prostheses. *Clin Toxicol (Phila)*. 2014;**52**(8):837-847.

120. Martin JR, Spencer-Gardner L, Camp CL, Stulak JM, Sierra RJ. Cardiac cobaltism: a rare complication after bilateral metal-on-metal total hip arthroplasty. *Arthroplast Today*. 2015;**1**(4):99-102.

121. Gilbert CJ, Cheung A, Butany J, Zywiel MG, Syed K, McDonald M, Wong F, Overgaard C. Hip pain and heart failure: the missing link. *Can J Cardiol*. 2013;**29**(5):639 e631-632.

122. Gillam MH, Pratt NL, Inacio MC, Roughead EE, Shakib S, Nicholls SJ, Graves SE. Heart failure after conventional metal-on-metal hip replacements. *Acta Orthop*. 2017;**88**(1):2-9.

123. Berber R, Abdel-Gadir A, Palla L, Moon J, Mainsty C, Skinner J, Hart A. Cardiac toxicity from metal hip implants. *British Hip Society Annual Scientific Meeting*. London, UK. 2017. Podium presentation.

124. Keegan GM, Learmonth ID, Case CP. Orthopaedic metals and their potential toxicity in the arthroplasty patient: A review of current knowledge and future strategies. *J Bone Joint Surg Br*. 2007;**89**(5):567-573.

125. Smith AJ, Dieppe P, Porter M, Blom AW, National Joint Registry of England and Wales. Risk of cancer in first seven years after metal-on-metal hip replacement compared with other bearings and general population: linkage study between the National Joint Registry of England and Wales and hospital episode statistics. *BMJ*. 2012;**344**:e2383.

126. Makela KT, Visuri T, Pulkkinen P, Eskelinen A, Remes V, Virolainen P, Junnila M, Pukkala E. Risk of cancer with metal-on-metal hip replacements: population based study. *BMJ*. 2012;**345**:e4646.

127. Makela KT, Visuri T, Pulkkinen P, Eskelinen A, Remes V, Virolainen P, Junnila M, Pukkala E. Cancer incidence and cause-specific mortality in patients with metal-on-metal hip replacements in Finland. *Acta Orthop*. 2014;**85**(1):32-38.

128. Kendal AR, Prieto-Alhambra D, Arden NK, Carr A, Judge A. Mortality rates at 10 years after metal-on-metal hip resurfacing compared with total hip replacement in England: retrospective cohort analysis of hospital episode statistics. *BMJ*. 2013;**347**:f6549.

129. McMinn DJ, Snell KI, Daniel J, Treacy RB, Pynsent PB, Riley RD. Mortality and implant revision rates of hip arthroplasty in patients with osteoarthritis: registry based cohort study. *BMJ*. 2012;**344**:e3319.

130. Lombardi AV, Jr., Barrack RL, Berend KR, Cuckler JM, Jacobs JJ, Mont MA, Schmalzried TP. The Hip Society: algorithmic approach to diagnosis and management of metal-on-metal arthroplasty. *J Bone Joint Surg Br*. 2012;**94**(11 Suppl A):14-18.

131. Mikhael MM, Hanssen AD, Sierra RJ. Failure of metal-on-metal total hip arthroplasty mimicking hip infection. A report of two cases. *J Bone Joint Surg Am*. 2009;**91**(2):443-446.

132. Glyn-Jones S, Pandit H, Kwon YM, Doll H, Gill HS, Murray DW. Risk factors for inflammatory pseudotumour formation following hip resurfacing. *J Bone Joint Surg Br*. 2009;**91**(12):1566-1574.

133. Reito A, Elo P, Puolakka T, Pajamaki J, Eskelinen A. Femoral diameter and stem type are independent risk factors for ARMD in the large-headed ASR THR group. *BMC Musculoskelet Disord*. 2015;**16**:118.

134. Reito A, Puolakka T, Elo P, Pajamaki J, Eskelinen A. High prevalence of adverse reactions to metal debris in small-headed ASR hips. *Clin Orthop Relat Res*. 2013;**471**(9):2954-2961.

135. Matharu GS, Nandra R, Berryman F, Judge A, Pynsent PB, Dunlop DJ. Risk factors for failure of the 36 mm metal-on-metal Pinnacle total hip replacement system: a retrospective single-centre cohort study. *Bone Joint J*. 2017; In Press (accepted 30th December 2016).

136. Reito A, Lainiala O, Elo P, Eskelinen A. Prevalence of Failure due to Adverse Reaction to Metal Debris in Modern, Medium and Large Diameter Metal-on-Metal Hip Replacements--The Effect of Novel Screening Methods: Systematic Review and Metaregression Analysis. *PLoS One*. 2016;**11**(3):e0147872.

137. Sabah SA, Henckel J, Koutsouris S, Rajani R, Hothi H, Skinner JA, Hart AJ. Are all metal-on-metal hip revision operations contributing to the National Joint Registry implant survival curves? : a study comparing the London Implant Retrieval Centre and National Joint Registry datasets. *Bone Joint J*. 2016;**98-B**(1):33-39.

138. Matharu GS, Revell MP, Pynsent PB, Treacy RB. A review of hip resurfacings revised for unexplained pain. *Hip Int*. 2012;**22**(6):633-640.

139. Canadian Arthroplasty Society. The Canadian Arthroplasty Society's experience with hip resurfacing arthroplasty. An analysis of 2773 hips. *Bone Joint J*. 2013;**95-B**(8):1045-1051.

140. Jameson SS, Baker PN, Mason J, Porter ML, Deehan DJ, Reed MR. Independent predictors of revision following metal-on-metal hip resurfacing: a retrospective cohort study using National Joint Registry data. *J Bone Joint Surg Br*. 2012;**94**(6):746-754.

141. Health Canada, Government of Canada. Metal-on-Metal Hip Implants - Information for Orthopaedic Surgeons Regarding Patient Management Following Surgery - For Health Professionals. 2012. http://healthycanadians.gc.ca/recall-alert-rappel-avis/hc-sc/2012/15835a-eng.php.

142. Therapeutic Goods Administration, Department of Health, Australian Government. Metal-on-metal hip replacement implants — information for general practitioners, orthopaedic surgeons and other health professionals. 2012. http://www.tga.gov.au/hp/information-devices-mom-hip-implants.htm.

143. Hallab N, Merritt K, Jacobs JJ. Metal sensitivity in patients with orthopaedic implants. *J Bone Joint Surg Am*. 2001;**83-A**(3):428-436.

144. McBryde CW, Theivendran K, Thomas AM, Treacy RB, Pynsent PB. The influence of head size and sex on the outcome of Birmingham hip resurfacing. *J Bone Joint Surg Am*. 2010;**92**(1):105-112.

145. Van Der Straeten C, Grammatopoulos G, Gill HS, Calistri A, Campbell P, De Smet KA. The 2012 Otto Aufranc Award: The interpretation of metal ion levels in unilateral and bilateral hip resurfacing. *Clin Orthop Relat Res*. 2013;**471**(2):377-385.

146. Pandit H, Vlychou M, Whitwell D, Crook D, Luqmani R, Ostlere S, Murray DW, Athanasou NA. Necrotic granulomatous pseudotumours in bilateral resurfacing hip arthoplasties: evidence for a type IV immune response. *Virchows Arch*. 2008;**453**(5):529-534.

147. Underwood R, Matthies A, Cann P, Skinner JA, Hart AJ. A comparison of explanted Articular Surface Replacement and Birmingham Hip Resurfacing components. *J Bone Joint Surg Br*. 2011;**93**(9):1169-1177.

148. Matthies AK, Henckel J, Cro S, Suarez A, Noble PC, Skinner J, Hart AJ. Predicting wear and blood metal ion levels in metal-on-metal hip resurfacing. *J Orthop Res*. 2014;**32**(1):167-174.

149. Hothi HS, Whittaker RK, Meswania JM, Blunn GW, Skinner JA, Hart AJ. Influence of stem type on material loss at the metal-on-metal pinnacle taper junction. *Proc Inst Mech Eng H*. 2015;**229**(1):91-97.

150. Grammatopoulos G, Pandit H, Glyn-Jones S, McLardy-Smith P, Gundle R, Whitwell D, Gill HS, Murray DW. Optimal acetabular orientation for hip resurfacing. *J Bone Joint Surg Br*. 2010;**92**(8):1072-1078.

151. Mellon SJ, Grammatopoulos G, Andersen MS, Pegg EC, Pandit HG, Murray DW, Gill HS. Individual motion patterns during gait and sit-to-stand contribute to edge-loading risk in metal-on-metal hip resurfacing. *Proc Inst Mech Eng H*. 2013;**227**(7):799-810.

152. Langton DJ, Sprowson AP, Joyce TJ, Reed M, Carluke I, Partington P, Nargol AV. 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):1287-1295.

153. Grammatopoulos G, Pandit H, Oxford Hip and Knee Group, Murray DW, Gill HS. The relationship between head-neck ratio and pseudotumour formation in metal-on-metal resurfacing arthroplasty of the hip. *J Bone Joint Surg Br*. 2010;**92**(11):1527-1534.

154. Lainiala O, Eskelinen A, Elo P, Puolakka T, Korhonen J, Moilanen T. Adverse reaction to metal debris is more common in patients following MoM total hip replacement with a 36 mm femoral head than previously thought: results from a modern MoM follow-up programme. *Bone Joint J*. 2014;**96-B**(12):1610-1617.

155. De Smet K, De Haan R, Calistri A, Campbell PA, Ebramzadeh E, Pattyn C, Gill HS. Metal ion measurement as a diagnostic tool to identify problems with metal-on-metal hip resurfacing. *J Bone Joint Surg Am*. 2008;**90**(Suppl 4):202-208.

156. Daniel J, Ziaee H, Pradhan C, Pynsent PB, McMinn DJ. Renal clearance of cobalt in relation to the use of metal-on-metal bearings in hip arthroplasty. *J Bone Joint Surg Am*. 2010;**92**(4):840-845.

157. Rahme M, Lavigne M, Barry J, Cirtiu CM, Belanger P, Vendittoli PA. Whole blood metal ion measurement reproducibility between different laboratories. *J Arthroplasty*. 2014;**29**(11):2214-2218.

158. Daniel J, Ziaee H, Pynsent PB, McMinn DJ. The validity of serum levels as a surrogate measure of systemic exposure to metal ions in hip replacement. *J Bone Joint Surg Br*. 2007;**89**(6):736-741.

159. Engh CA, MacDonald SJ, Sritulanondha S, Korczak A, Naudie D, Engh C. Metal ion levels after metal-on-metal total hip arthroplasty: a five-year, prospective randomized trial. *J Bone Joint Surg Am*. 2014;**96**(6):448-455.

160. 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**(4):441-448.

161. Lloyd J, Starks I, Wainwright, Middleton R. Metal-on-Metal Resurfacing and the Cost to the Nation: A Conservative Estimate of the Unexpected Costs Required to Implement the New Metal-on-Metal Follow-Up Programme in the UK. In: Knahr K, editor. Total Hip Arthroplasty Tribological Considerations and Clinical Consequences. Berlin Heidelberg: Springer; 2013. p. 45-51.

162. Sidaginamale RP, Joyce TJ, Lord JK, Jefferson R, Blain PG, Nargol AV, Langton DJ. Blood metal ion testing is an effectivescreening tool to identify poorly performing metal-on-metal bearingsurfaces. *Bone Joint Res*. 2013;**2**(5):84-95.

163. Heisel C, Streich N, Krachler M, Jakubowitz E, Kretzer JP. Characterization of the running-in period in total hip resurfacing arthroplasty: an in vivo and in vitro metal ion analysis. *J Bone Joint Surg Am*. 2008;**90**(Suppl 3):125-133.

164. Daniel J, Ziaee H, Pradhan C, McMinn DJ. Six-year results of a prospective study of metal ion levels in young patients with metal-on-metal hip resurfacings. *J Bone Joint Surg Br*. 2009;**91**(2):176-179.

165. Amstutz HC, Campbell PA, Dorey FJ, Johnson AJ, Skipor AK, Jacobs JJ. Do ion concentrations after metal-on-metal hip resurfacing increase over time? A prospective study. *J Arthroplasty*. 2013;**28**(4):695-700.

166. Bernstein M, Desy NM, Petit A, Zukor DJ, Huk OL, Antoniou J. Long-term follow-up and metal ion trend of patients with metal-on-metal total hip arthroplasty. *Int Orthop*. 2012;**36**(9):1807-1812.

167. Hart AJ, Sabah SA, Bandi AS, Maggiore P, Tarassoli P, Sampson B, J AS. Sensitivity and specificity of blood cobalt and chromium metal ions for predicting failure of metal-on-metal hip replacement. *J Bone Joint Surg Br*. 2011;**93**(10):1308-1313.

168. Hart AJ, Sabah SA, Sampson B, Skinner JA, Powell JJ, Palla L, Pajamaki KJ, Puolakka T, Reito A, Eskelinen A. Surveillance of Patients with Metal-on-Metal Hip Resurfacing and Total Hip Prostheses: A Prospective Cohort Study to Investigate the Relationship Between Blood Metal Ion Levels and Implant Failure. *J Bone Joint Surg Am*. 2014;**96**(13):1091-1099.

169. Hart AJ, Skinner JA, Henckel J, Sampson B, Gordon F. Insufficient acetabular version increases blood metal ion levels after metal-on-metal hip resurfacing. *Clin Orthop Relat Res*. 2011;**469**(9):2590-2597.

170. Matharu GS, Berryman F, Brash L, Pynsent PB, Treacy RB, Dunlop DJ. Predicting high blood metal ion concentrations following hip resurfacing. *Hip Int*. 2015;**25**(6):510-519.

171. Matharu GS, Berryman F, Brash L, Pynsent PB, Treacy RB, Dunlop DJ. Influence of implant design on blood metal ion concentrations in metal-on-metal total hip replacement patients. *Int Orthop*. 2015;**39**(9):1803-1811.

172. Lainiala OS, Moilanen TP, Hart AJ, Huhtala HS, Sabah SA, Eskelinen AP. Higher Blood Cobalt and Chromium Levels in Patients With Unilateral Metal-on-Metal Total Hip Arthroplasties Compared to Hip Resurfacings. *J Arthroplasty*. 2016;**31**(6):1261-1266.

173. De Haan R, Pattyn C, Gill HS, Murray DW, Campbell PA, De Smet K. 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**(10):1291-1297.

174. Langton DJ, Joyce TJ, Mangat N, Lord J, Van Orsouw M, De Smet K, Nargol AV. Reducing metal ion release following hip resurfacing arthroplasty. *Orthop Clin North Am*. 2011;**42**(2):169-180, viii.

175. Emmanuel AR, Bergin KM, Kelly GE, McCoy GF, Wozniak AP, Quinlan JF. The effect of acetabular inclination on metal ion levels following metal-on-metal hip arthroplasty. *J Arthroplasty*. 2014;**29**(1):186-191.

176. Bernstein M, Walsh A, Petit A, Zukor DJ, Huk OL, Antoniou J. Femoral head size does not affect ion values in metal-on-metal total hips. *Clin Orthop Relat Res*. 2011;**469**(6):1642-1650.

177. Lavigne M, Belzile EL, Roy A, Morin F, Amzica T, Vendittoli PA. Comparison of whole-blood metal ion levels in four types of metal-on-metal large-diameter femoral

head total hip arthroplasty: the potential influence of the adapter sleeve. *J Bone Joint Surg Am*. 2011;**93**(Suppl 2):128-136.

178. Chang EY, McAnally JL, Van Horne JR, Van Horne JG, Wolfson T, Gamst A, Chung CB. Relationship of plasma metal ions and clinical and imaging findings in patients with ASR XL metal-on-metal total hip replacements. *J Bone Joint Surg Am*. 2013;**95**(22):2015-2020.

179. Garbuz DS, Tanzer M, Greidanus NV, Masri BA, Duncan CP. The John Charnley Award: Metal-on-metal hip resurfacing versus large-diameter head metal-on-metal total hip arthroplasty: a randomized clinical trial. *Clin Orthop Relat Res*. 2010;**468**(2):318-325.

180. Beaule PE, Kim PR, Hamdi A, Fazekas A. A prospective metal ion study of large-head metal-on-metal bearing: a matched-pair analysis of hip resurfacing versus total hip replacement. *Orthop Clin North Am*. 2011;**42**(2):251-257, ix.

181. Fehring TK, Carter JL, Fehring KA, Odum SM, Griffin WL. Cobalt to Chromium Ratio is Not a Key Marker for Adverse Local Tissue Reaction (ALTR) in Metal on Metal Hips. *J Arthroplasty*. 2015;**30**(9 Suppl):107-109.

182. Matharu GS, Pandit HG. The Cobalt-To-Chromium Ratio "May Be" a Key Marker for Adverse Local Tissue Reactions in Metal-On-Metal Hips. *J Arthroplasty*. 2016;**31**(6):1374-1375.

183. Reito A, Lainiala O, Eskelinen A. Letter to Editor: Cobalt to Chromium Ratio Is Not a Key Marker for Adverse Local Tissue Reaction in Metal-On-Metal Hips. *J Arthroplasty*. 2016;**31**(2):556-557.

184. MHRA. Medical Device Alert: metal-on-metal (MoM) hip replacements: Birmingham Hip Resurfacing (BHR) system (Smith & Nephew Orthopaedics). 2015. https://www.gov.uk/drug-device-alerts/metal-on-metal-mom-hip-replacements-guidance-on-implantation-and-patient-management.

185. Malek IA, King A, Sharma H, Malek S, Lyons K, Jones S, John A. The sensitivity, specificity and predictive values of raised plasma metal ion levels in the diagnosis of adverse reaction to metal debris in symptomatic patients with a metal-on-metal arthroplasty of the hip. *J Bone Joint Surg Br*. 2012;**94**(8):1045-1050.

186. Langton DJ, Sidaginamale RP, Joyce TJ, Natu S, Blain P, Jefferson RD, Rushton S, Nargol AV. The clinical implications of elevated blood metal ion concentrations in asymptomatic patients with MoM hip resurfacings: a cohort study. *BMJ Open*. 2013;**3**(3).

187. Reito A, Moilanen T, Puolakka T, Pajamaki J, Eskelinen A. Repeated metal ion measurements in patients with high risk metal-on-metal hip replacement. *Int Orthop*. 2014;**38**(7):1353-1361.

188. Reito A, Lainiala O, Nieminen J, Eskelinen A. Repeated metal ion measurement in patients with bilateral metal on metal (ASR) hip replacements. *Orthop Traumatol Surg Res*. 2016;**102**(2):167-173.

189. Ball ST, Severns D, Linn M, Meyer RS, Swenson FC. What happens to serum metal ion levels after a metal-on-metal bearing is removed? *J Arthroplasty*. 2013;**28**(8 Suppl):53-55.

190. Durrani SK, Noble PC, Sampson B, Panetta T, Liddle AD, Sabah SA, Chan NK, Skinner JA, Hart AJ. Changes in blood ion levels after removal of metal-on-metal hip replacements: 16 patients followed for 0-12 months. *Acta Orthop*. 2014;**85**(3):259-265.

191. Lainiala O, Reito A, Elo P, Pajamaki J, Puolakka T, Eskelinen A. Revision of Metal-on-metal Hip Prostheses Results in Marked Reduction of Blood Cobalt and Chromium Ion Concentrations. *Clin Orthop Relat Res*. 2015;**473**(7):2305-2313.

192. Chen Z, Pandit H, Taylor A, Gill H, Murray D, Ostlere S. Metal-on-metal hip resurfacings--a radiological perspective. *Eur Radiol*. 2011;**21**(3):485-491.

193. 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**(10):955-961.

194. Toms AP, Marshall TJ, Cahir J, Darrah C, Nolan J, Donell ST, Barker T, Tucker JK. 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.

195. Matharu GS, Pynsent PB, Sumathi VP, Mittal S, Buckley CD, Dunlop DJ, Revell PA, Revell MP. Predictors of time to revision and clinical outcomes following revision of metal-on-metal hip replacements for adverse reaction to metal debris. *Bone Joint J*. 2014;**96-B**(12):1600-1609.

196. Madanat R, Rolfson O, Donahue GS, Hussey DK, Potter HG, Wallace R, Muratoglu OK, Malchau H. Medial Calcar Erosion Is Associated With Synovial Thickness in Patients With ASR XL Total Hip Arthroplasty. *J Arthroplasty*. 2016;**31**(11):2588-2592.

197. 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**(8):1019-1024.

198. 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**(8):715-722.

199. Nishii T, Sakai T, Takao M, Yoshikawa H, Sugano N. Ultrasound screening of periarticular soft tissue abnormality around metal-on-metal bearings. *J Arthroplasty*. 2012;**27**(6):895-900.

200. Douis H, Dunlop DJ, Pearson AM, O'Hara JN, James SL. The role of ultrasound in the assessment of post-operative complications following hip arthroplasty. *Skeletal Radiol*. 2012;**41**(9):1035-1046.

201. Siddiqui IA, Sabah SA, Satchithananda K, Lim AK, Cro S, Henckel J, Skinner JA, Hart AJ. A comparison of the diagnostic accuracy of MARS MRI and ultrasound of the painful metal-on-metal hip arthroplasty. *Acta Orthop*. 2014;**85**(4):375-382.

202. Kwon YM, Ostlere SJ, McLardy-Smith P, Athanasou NA, Gill HS, Murray DW. "Asymptomatic" pseudotumors after metal-on-metal hip resurfacing arthroplasty: prevalence and metal ion study. *J Arthroplasty*. 2011;**26**(4):511-518.

203. Liddle AD, Satchithananda K, Henckel J, Sabah SA, Vipulendran KV, Lewis A, Skinner JA, Mitchell AW, Hart AJ. Revision of metal-on-metal hip arthroplasty in a tertiary center: a prospective study of 39 hips with between 1 and 4 years of follow-up. *Acta Orthop*. 2013;**84**(3):237-245.

204. Beggs I, Bianchi S, Bueno A. European Society of Musculoskeletal Radiology. Musculoskeletal Ultrasound Technical Guidelines IV. Hip. 2016. https://essr.org/content-essr/uploads/2016/2010/hip.pdf.

205. Williams DH, Greidanus NV, Masri BA, Duncan CP, Garbuz DS. Prevalence of pseudotumor in asymptomatic patients after metal-on-metal hip arthroplasty. *J Bone Joint Surg Am*. 2011;**93**(23):2164-2171.

206. Lainiala O, Elo P, Reito A, Pajamaki J, Puolakka T, Eskelinen A. Good sensitivity and specificity of ultrasound for detecting pseudotumors in 83 failed metal-on-metal hip replacements. *Acta Orthop*. 2015;**86**(3):339-344.

207. Sofka CM, Haddad ZK, Adler RS. Detection of muscle atrophy on routine sonography of the shoulder. *J Ultrasound Med*. 2004;**23**(8):1031-1034.

208. Garcia FL, Picado CH, Nogueira-Barbosa MH. Sonographic evaluation of the abductor mechanism after total hip arthroplasty. *J Ultrasound Med*. 2010;**29**(3):465-471.

209. Matharu GS, Janardhan S, Brash L, Pynsent PB, Dunlop DJ, James SL. The utility of repeat ultrasound imaging in the follow-up of metal-on-metal hip arthroplasty patients. *Ann R Coll Surg Engl*. 2016;**98**(2):143-149.

210. Siddiqui IA, Sabah SA, Satchithananda K, Lim AK, Henckel J, Skinner JA, Hart AJ. Cross-sectional imaging of the metal-on-metal hip prosthesis: the London ultrasound protocol. *Clin Radiol*. 2013;**68**(8):e472-478.

211. Bal BS, Lowe JA. Muscle damage in minimally invasive total hip arthroplasty: MRI evidence that it is not significant. *Instr Course Lect*. 2008;**57**:223-229.

212. Bisseling P, de Wit BW, Hol AM, van Gorp MJ, van Kampen A, van Susante JL. Similar incidence of periprosthetic fluid collections after ceramic-on-polyethylene total hip arthroplasties and metal-on-metal resurfacing arthroplasties: results of a screening metal artefact reduction sequence-MRI study. *Bone Joint J*. 2015;**97-B**(9):1175-1182.

213. Das DH, van der Weegen W, Wullems JA, Brakel K, Sijbesma T, Nelissen RG. Periprosthetic pathology in 'at risk' ceramic-on-polyethylene total hip arthroplasty: a clinical study using MARS-MRI in 50 patients. *Hip Int*. 2016;**26**(3):244-248.

214. van der Veen HC, Reininga IH, Zijlstra WP, Boomsma MF, Bulstra SK, van Raay JJ. Pseudotumour incidence, cobalt levels and clinical outcome after large head metal-on-metal and conventional metal-on-polyethylene total hip arthroplasty: mid-term results of a randomised controlled trial. *Bone Joint J*. 2015;**97-B**(11):1481-1487.

215. 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):303-307.

216. Hauptfleisch J, Pandit H, Grammatopoulos G, Gill HS, Murray DW, Ostlere S. A MRI classification of periprosthetic soft tissue masses (pseudotumours) associated with metal-on-metal resurfacing hip arthroplasty. *Skeletal Radiol*. 2012;**41**(2):149-155.

217. Wynn-Jones H, Macnair R, Wimhurst J, Chirodian N, Derbyshire B, Toms A, Cahir J. Silent soft tissue pathology is common with a modern metal-on-metal hip arthroplasty. *Acta Orthop*. 2011;**82**(3):301-307.

218. van der Weegen W, Brakel K, Horn RJ, Wullems JA, Das HP, Pilot P, Nelissen RG. Comparison of different pseudotumor grading systems in a single cohort of metal-on-metal hip arthroplasty patients. *Skeletal Radiol*. 2014;**43**(2):149-155.

219. Briant-Evans TW, Lyle N, Barbur S, Hauptfleisch J, Amess R, Pearce AR, Conn KS, Stranks GJ, Britton JM. A longitudinal study of MARS MRI scanning of soft-tissue lesions around metal-on-metal total hip arthroplasties and disease progression. *Bone Joint J*. 2015;**97-B**(10):1328-1337.

220. Lainiala O, Elo P, Reito A, Pajamaki J, Puolakka T, Eskelinen A. Comparison of extracapsular pseudotumors seen in magnetic resonance imaging and in revision surgery of 167 failed metal-on-metal hip replacements. *Acta Orthop*. 2014;**85**(5):474-479.

221. Hayter CL, Koff MF, Potter HG. Magnetic resonance imaging of the postoperative hip. *J Magn Reson Imaging*. 2012;3**5**(5):1013-1025.

222. Berber R, Khoo M, Cook E, Guppy A, Hua J, Miles J, Carrington R, Skinner J, Hart A. Muscle atrophy and metal-on-metal hip implants: a serial MRI study of 74 hips. *Acta Orthop*. 2015;**86**(3):351-357.

223. van der Weegen W, Brakel K, Horn RJ, Hoekstra HJ, Sijbesma T, Pilot P, Nelissen RG. Asymptomatic pseudotumours after metal-on-metal hip resurfacing show little change within one year. *Bone Joint J*. 2013;**95-B**(12):1626-1631.

224. Reito A, Elo P, Puolakka T, Pajamaki J, Nieminen J, Eskelinen A. Repeated magnetic resonance imaging in 154 hips with large-diameter metal-on-metal hip replacement. *Acta Orthop*. 2014;**85**(6):570-576.

225. Ebreo D, Bell PJ, Arshad H, Donell ST, Toms A, Nolan JF. Serial magnetic resonance imaging of metal-on-metal total hip replacements. Follow-up of a cohort of 28 mm Ultima TPS THRs. *Bone Joint J*. 2013;**95-B**(8):1035-1039.

226. Goldstein JM, Fehring TK, Fehring KA. Cystic Adverse Local Tissue Reactions in Asymptomatic Modular Metal-on-Metal Total Hips May Decrease Over Time. *J Arthroplasty*. 2016;**31**(7):1589-1594.

227. Kwon YM, Liow MH, Dimitriou D, Tsai TY, Freiberg AA, Rubash HE. What Is the Natural History of "Asymptomatic" Pseudotumours in Metal-on-Metal Hip Arthroplasty? Minimum 4-Year Metal Artifact Reduction Sequence Magnetic Resonance Imaging Longitudinal Study. *J Arthroplasty*. 2016;**31**(9 Suppl):121-126.

228. Hasegawa M, Miyamoto N, Miyazaki S, Wakabayashi H, Sudo A. Longitudinal magnetic resonance imaging of pseudotumors following metal-on-metal total hip arthroplasty. *J Arthroplasty*. 2014;**29**(12):2236-2238.

229. Garbuz DS, Hargreaves BA, Duncan CP, Masri BA, Wilson DR, Forster BB. The John Charnley Award: Diagnostic accuracy of MRI versus ultrasound for detecting pseudotumors in asymptomatic metal-on-metal THA. *Clin Orthop Relat Res*. 2014;**472**(2):417-423.

230. Nishii T, Sakai T, Takao M, Yoshikawa H, Sugano N. Is ultrasound screening reliable for adverse local tissue reaction after hip arthroplasty? *J Arthroplasty*. 2014;**29**(12):2239-2244.

231. Muraoka K, Naito M, Nakamura Y, Hagio T, Takano K. Usefulness of ultrasonography for detection of pseudotumors after metal-on-metal total hip arthroplasty. *J Arthroplasty*. 2015;**30**(5):879-884.

232. Kwon YM, Dimitriou D, Liow MH, Tsai TY, Li G. Is Ultrasound As Useful As Metal Artifact Reduction Sequence Magnetic Resonance Imaging in Longitudinal Surveillance of Metal-on-Metal Hip Arthroplasty Patients? *J Arthroplasty*. 2016;**31**(8):1821-1827.

233. Hart AJ, Sabah S, Henckel J, Lewis A, Cobb J, Sampson B, Mitchell A, Skinner JA. The painful metal-on-metal hip resurfacing. *J Bone Joint Surg Br*. 2009;**91**(6):738-744.

234. Roth TD, Maertz NA, Parr JA, Buckwalter KA, Choplin RH. CT of the hip prosthesis: appearance of components, fixation, and complications. *Radiographics*. 2012;**32**(4):1089-1107.

235. Robinson E, Henckel J, Sabah S, Satchithananda K, Skinner J, Hart A. Cross-sectional imaging of metal-on-metal hip arthroplasties. Can we substitute MARS MRI with CT? *Acta Orthop*. 2014;**85**(6):577-584.

236. Yi PH, Cross MB, Moric M, Levine BR, Sporer SM, Paprosky WG, Jacobs JJ, Della Valle CJ. Do serologic and synovial tests help diagnose infection in revision hip arthroplasty with metal-on-metal bearings or corrosion? *Clin Orthop Relat Res*. 2015;**473**(2):498-505.

237. Judd KT, Noiseux N. Concomitant infection and local metal reaction in patients undergoing revision of metal on metal total hip arthroplasty. *Iowa Orthop J*. 2011;**31**:59-63.

238. Watters TS, Eward WC, Hallows RK, Dodd LG, Wellman SS, Bolognesi MP. Pseudotumor with superimposed periprosthetic infection following metal-on-metal total hip arthroplasty: a case report. *J Bone Joint Surg Am*. 2010;**92**(7):1666-1669.

239. Davda K, Lali FV, Sampson B, Skinner JA, Hart AJ. An analysis of metal ion levels in the joint fluid of symptomatic patients with metal-on-metal hip replacements. *J Bone Joint Surg Br*. 2011;**93**(6):738-745.

240. Tischler EH, Plummer DR, Chen AF, Della Valle CJ, Parvizi J. Leukocyte Esterase: Metal-on-Metal Failure and Periprosthetic Joint Infection. *J Arthroplasty*. 2016;**31**(10):2260-2263.

241. Chambers AW, Lacy KW, Liow MH, Manalo JP, Freiberg AA, Kwon YM. Multiple Hip Intra-Articular Steroid Injections Increase Risk of Periprosthetic Joint Infection Compared With Single Injections. *J Arthroplasty*. 2017.

242. Berber R, Henckel J, Khoo M, Wan S, Hua J, Skinner J, Hart A. Clinical Usefulness of SPECT-CT in Patients with an Unexplained Pain in Metal on Metal (MOM) Total Hip Arthroplasty. *J Arthroplasty*. 2015;**30**(4):687-694.

243. de Steiger RN, Miller LN, Prosser GH, Graves SE, Davidson DC, Stanford TE. Poor outcome of revised resurfacing hip arthroplasty. *Acta Orthop*. 2010;**81**(1):72-76.

244. OCEBM Levels of Evidence Working Group. Oxford Centre for Evidence-Based Medicine Levels of Evidence. 2011. http://www.cebm.net/wp-content/uploads/2014/2006/CEBM-Levels-of-Evidence-2012.2011.pdf.

245. Berber R, Pappas Y, Khoo M, Miles J, Carrington R, Skinner J, Hart A. A new approach to managing patients with problematic metal hip implants: the use of an Internet-enhanced multidisciplinary team meeting: AAOS exhibit selection. *J Bone Joint Surg Am*. 2015;**97**(4):e20.

246. Berber R, Skinner J, Board T, Kendoff D, Eskelinen A, Kwon YM, Padgett DE, Hart A, IsccoMh. International metal-on-metal multidisciplinary teams: do we manage patients with metal-on-metal hip arthroplasty in the same way? An analysis from the International Specialist Centre Collaboration on MOM Hips (ISCCoMH). *Bone Joint J*. 2016;**98-B**(2):179-186.

247. Munro JT, Masri BA, Duncan CP, Garbuz DS. High complication rate after revision of large-head metal-on-metal total hip arthroplasty. *Clin Orthop Relat Res*. 2014;**472**(2):523-528.

248. Pritchett JW. One-component revision of failed hip resurfacing from adverse reaction to metal wear debris. *J Arthroplasty*. 2014;**29**(1):219-224.

249. Rajpura A, Porter ML, Gambhir AK, Freemont AJ, Board TN. Clinical experience of revision of metal on metal hip arthroplasty for aseptic lymphocyte dominated vasculitis associated lesions (ALVAL). *Hip Int*. 2011;**21**(1):43-51.

250. Su EP, Su SL. Surface replacement conversion: results depend upon reason for revision. *Bone Joint J*. 2013;**95-B**(11 Suppl A):88-91.

251. Norris RJ, McArthur J, Parsons H, Smith NA, Sprowson AP, Foguet P. A case series of 35 hip revisions for adverse reactions to metal debris following Cormet hip resurfacing. *Hip Int*. 2014;**24**(6):575-581.

252. Cip J, Bach C, Widemschek M, Luegmair M, Martin A. Revision of Articular Surface Replacement (ASR) Total Hip Arthroplasty: Correlation of Perioperative Data and Early Post-Revision Outcome Results. *J Arthroplasty*. 2015;**30**(9):1607-1617.

253. Stryker LS, Odum SM, Fehring TK, Springer BD. Revisions of monoblock metal-on-metal THAs have high early complication rates. *Clin Orthop Relat Res*. 2015;**473**(2):469-474.

254. van Lingen CP, Ettema HB, Bosker BH, Verheyen CC. Revision of a single type of large metal head metal-on-metal hip prosthesis. *Hip Int*. 2015;**25**(3):221-226.

255. Liow MH, Dimitriou D, Tsai TY, Kwon YM. Preoperative Risk Factors Associated With Poor Outcomes of Revision Surgery for "Pseudotumors" in Patients With Metal-on-Metal Hip Arthroplasty. *J Arthroplasty*. 2016;**31**(12):2835-2842.

256. Matharu GS, Pandit HG, Murray DW. Poor Survivorship and Frequent Complications at a Median of 10 Years After Metal-on-Metal Hip Resurfacing Revision. *Clin Orthop Relat Res*. 2017;**475**(2):304-314.

257. 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):185-190.

258. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, Dawson J. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br*. 2007;**89**(8):1010-1014.

259. Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am*. 1969;**51**(4):737-755.

260. Klassbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol*. 2003;**32**(1):46-51.

261. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;**15**(12):1833-1840.

262. Kalairajah Y, Azurza K, Hulme C, Molloy S, Drabu KJ. Health outcome measures in the evaluation of total hip arthroplasties--a comparison between the Harris hip score and the Oxford hip score. *J Arthroplasty*. 2005;**20**(8):1037-1041.

263. Wong JM, Liu YL, Graves S, de Steiger R. What Is the Rerevision Rate After Revising a Hip Resurfacing Arthroplasty? Analysis From the AOANJRR. *Clin Orthop Relat Res*. 2015;**473**(11):3458-3464.

264. Penrose CT, Seyler TM, Wellman SS, Bolognesi MP, Lachiewicz PF. Complications Are Not Increased With Acetabular Revision of Metal-on-metal Total Hip Arthroplasty. *Clin Orthop Relat Res*. 2016;**474**(10):2134-2142.

265. Scientific Committee of Emergng Newly Identified Health Risks. Final Opinion on the safety of Metal-on-Metal joint replacements with a particular focus on hip implants. 2014. https://ec.europa.eu/health/scientific_committees/consultations/public_consultations/scenihr_consultation_20_en.

266. Cohen D. How safe are metal-on-metal hip implants? *BMJ*. 2012;**344**:e1410.

267. Prosser GH, Yates PJ, Wood DJ, Graves SE, de Steiger RN, Miller LN. Outcome of primary resurfacing hip replacement: evaluation of risk factors for early revision. *Acta Orthop*. 2010;**81**:66-71.

268. Daniel J, Ziaee H, Salama A, Pradhan C, McMinn DJ. The effect of the diameter of metal-on-metal bearings on systemic exposure to cobalt and chromium. *J Bone Joint Surg Br*. 2006;**88**(4):443-448.

269. Harrington CF, Taylor A. Metal-on-metal hip implants. UK quality assurance of blood cobalt and chromium after hip implants. *BMJ*. 2012;**344**:e4017.

270. Lewthwaite SC, Squires B, Gie GA, Timperley AJ, Ling RS. The Exeter Universal hip in patients 50 years or younger at 10-17 years' followup. *Clin Orthop Relat Res*. 2008;**466**(2):324-331.

271. Abdel MP, Roth PV, Harmsen WS, Berry DJ. What is the lifetime risk of revision for patients undergoing total hip arthroplasty? a 40-year observational study of patients treated with the Charnley cemented total hip arthroplasty. *Bone Joint J*. 2016;**98-B**(11):1436-1440.

272. Eskelinen A, Paavolainen P, Helenius I, Pulkkinen P, Remes V. Total hip arthroplasty for rheumatoid arthritis in younger patients: 2,557 replacements in the Finnish Arthroplasty Register followed for 0-24 years. *Acta Orthop*. 2006;**77**(6):853-865.

273. Judge A, Wallace G, Prieto-Alhambra D, Arden NK, Edwards CJ. Can the publication of guidelines change the management of early rheumatoid arthritis? An interrupted time series analysis from the United Kingdom. *Rheumatology (Oxford).* 2015;**54**(12):2244-2248.

274. Youdas JW, Krause DA, Hollman JH, Harmsen WS, Laskowski E. The influence of gender and age on hamstring muscle length in healthy adults. *J Orthop Sports Phys Ther*. 2005;**35**(4):246-252.

275. Murray DW, Grammatopoulos G, Gundle R, Gibbons CL, Whitwell D, Taylor A, Glyn-Jones S, Pandit HG, Ostlere S, Gill HS, Athanasou N, McLardy-Smith P. Hip resurfacing and pseudotumour. *Hip Int*. 2011;**21**(3):279-283.

276. Orthopaedic Data Evaluation Panel (ODEP). http://www.odep.org.uk.

277. Beyond Compliance. http://www.beyondcompliance.org.uk.

278. Dorr LD. Metal-on-Metal Testing to Salvage a Device Disaster: Commentary on an article by Gulraj S. Matharu, BSc(Hons), MBChB, MRCS, MRes, et al.: "The Effectiveness of Blood Metal Ions in Identifying Patients with Unilateral Birmingham Hip Resurfacing and Corail-Pinnacle Metal-on-Metal Hip Implants at Risk of Adverse Reactions to Metal Debris". *J Bone Joint Surg Am*. 2016;**98**(8):e31.

279. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, Altman DG, Moons KG. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med*. 2012;**9**(5):1-12.

280. National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. NJR Surgeon and Hospital Profile. 2017. http://www.njrsurgeonhospitalprofile.org.uk.

281. McMinn DJ, Daniel J, Ziaee H, Pradhan C. Indications and results of hip resurfacing. *Int Orthop*. 2011;**35**(2):231-237.

282. Treacy RB, McBryde CW, Shears E, Pynsent PB. Birmingham hip resurfacing: a minimum follow-up of ten years. *J Bone Joint Surg Br*. 2011;**93**(1):27-33.

283. Reito A, Puolakka T, Pajamaki J. Birmingham hip resurfacing: five to eight year results. *Int Orthop*. 2011;**35**(8):1119-1124.

284. Matharu GS, Mansour R, Dada O, Ostlere S, Pandit HG, Murray DW. Which imaging modality is most effective for identifying pseudotumours in metal-on-metal hip resurfacings requiring revision: ultrasound or MARS-MRI or both? *Bone Joint J*. 2016;**98-B**(1):40-48.

285. Winship PD, Faria N, Skinner JA, Hart AJ, Powerll JJ. The determination of cobalt and chromium in whole blood sampled from metal-metal and metal-polyethylene hip articulation patients by DRC-ICP-MS. In: Collery P, Maymard I, Theophanides T, Khassanova L, Collery T, editors. Metal ions in biology and medicine. Volume 10. Paris: John Libbey Eurotext; 2008. p. 27-33.

286. Eng J. Receiver operating characteristic analysis: a primer. *Acad Radiol*. 2005;**12**(7):909-916.

287. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;**12**:77.

288. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ*. 2009;**338**:b605.

289. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med*. 2006;**144**(3):201-209.

290. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;**338**:b606.

291. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med*. 1999;**130**(6):515-524.

292. Hussey DK, Madanat R, Donahue GS, Rolfson O, Bragdon CR, Muratoglu OK, Malchau H. Scoring the Current Risk Stratification Guidelines in Follow-up Evaluation of Patients After Metal-on-Metal Hip Arthroplasty: A Proposal for a Metal-on-Metal Risk Score Supporting Clinical Decision-Making. *J Bone Joint Surg Am*. 2016;**98**(22):1905-1912.

293. DeLee JG, Charnley J. Radiological demarcation of cemented sockets in total hip replacement. *Clin Orthop Relat Res*. 1976;(121):20-32.

294. Brooker AF, Bowerman JW, Robinson RA, Riley LH, Jr. Ectopic ossification following total hip replacement. Incidence and a method of classification. *J Bone Joint Surg Am*. 1973;**55**(8):1629-1632.

295. Rashband WS. ImageJ. 2016. http://imagej.nih.gov/ij/.

296. Lewinnek GE, Lewis JL, Tarr R, Compere CL, Zimmerman JR. Dislocations after total hip-replacement arthroplasties. *J Bone Joint Surg Am*. 1978;**60**(2):217-220.

297. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;**33**(1):159-174.

298. Chen X, Ender P, Mitchell M, Wells C. Stata web books: logistic regression with Stata. Los Angeles, USA: UCLA Academic Technology Services, Statistical Consulting Group; 2009.

299. Harrell FE Jr. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. New York, USA: Springer; 2001.

300. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ*. 2009;**338**:b604.

301. Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ*. 2009;**339**:b2584.

302. van Staa TP, Geusens P, Kanis JA, Leufkens HG, Gehlbach S, Cooper C. A simple clinical score for estimating the long-term risk of fracture in post-menopausal women. *QJM*. 2006;**99**(10):673-682.

303. Amstutz HC, Thomas BJ, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip. A comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am*. 1984;**66**(2):228-241.

304. Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. *J Arthroplasty*. 1998;**13**(8):890-895.

305. Back DL, Smith JD, Dalziel RE, Young DA, Shimmin A. Incidence of heterotopic ossification after hip resurfacing. *ANZ J Surg*. 2007;**77**(8):642-647.

306. National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. NJR mission statement. 2017. http://www.njrcentre.org.uk/njrcentre/AbouttheNJR/tabid/73/Default.aspx.

307. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*. 2011;**30**(10):1105-1117.

308. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;**338**:b2393.

309. Mackinnon A. The use and reporting of multiple imputation in medical research - a review. *J Intern Med*. 2010;**268**(6):586-593.

310. Royston, P. Multiple imputation of missing values. *The Stata Journal*. 2004;**4**(3): 227-241.

311. Royston, P. Multiple imputation of missing values: update. *The Stata Journal*. 2005;**5**(2); 188-201.

312. Austin PC. Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. *Biom J*. 2009;**51**(1):171-184.

313. Glynn RJ, Schneeweiss S, Sturmer T. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol*. 2006;**98**(3):253-259.

314. Jameson SS, Baker PN, Mason J, Rymaszewska M, Gregg PJ, Deehan DJ, Reed MR. Independent predictors of failure up to 7.5 years after 35 386 single-brand cementless total hip replacements: a retrospective cohort study using National Joint Registry data. *Bone Joint J*. 2013;**95-B**(6):747-757.

315. Abdel MP, Houdek MT, Watts CD, Lewallen DG, Berry DJ. Epidemiology of periprosthetic femoral fractures in 5417 revision total hip arthroplasties: a 40-year experience. *Bone Joint J*. 2016;**98-B**(4):468-474.

316. Ibrahim MS, Raja S, Khan MA, Haddad FS. A multidisciplinary team approach to two-stage revision for the infected hip replacement: a minimum five-year follow-up study. *Bone Joint J*. 2014;**96-B**(10):1312-1318.

317. Berend KR, Lombardi AV, Jr., Morris MJ, Bergeson AG, Adams JB, Sneller MA. Two-stage treatment of hip periprosthetic joint infection is associated with a high rate of infection control but high mortality. *Clin Orthop Relat Res*. 2013;**471**(2):510-518.

318. Bhattacharyya T, Chang D, Meigs JB, Estok DM, 2nd, Malchau H. Mortality after periprosthetic fracture of the femur. *J Bone Joint Surg Am*. 2007;**89**(12):2658-2662.

319. Parvizi J, Haddad FS. Periprosthetic joint infection: the last frontier. *Bone Joint J*. 2015;**97-B**(9):1157-1158.

320. Houdek MT, Wagner ER, Watts CD, Osmon DR, Hanssen AD, Lewallen DG, Mabry TM. Morbid obesity: a significant risk factor for failure of two-stage revision total hip arthroplasty for infection. *J Bone Joint Surg Am*. 2015;**97**(4):326-332.

321. Pulos N, McGraw MH, Courtney PM, Lee GC. Revision THA in obese patients is associated with high re-operation rates at short-term follow-up. *J Arthroplasty*. 2014;**29**(9 Suppl):209-213.

322. Lomas R, Chandrasekar A, Board TN. Bone allograft in the U.K.: perceptions and realities. *Hip Int*. 2013;**23**(5):427-433.

323. Grammatopoulos G, Thomas GE, Pandit H, Beard DJ, Gill HS, Murray DW. The effect of orientation of the acetabular component on outcome following total hip arthroplasty with small diameter hard-on-soft bearings. *Bone Joint J*. 2015;**97-B**(2):164-172.

324. Matharu GS, Berryman F, Brash L, Pynsent PB, Treacy RB, Dunlop DJ. The Effectiveness of Blood Metal Ions in Identifying Patients with Unilateral Birmingham Hip Resurfacing and Corail-Pinnacle Metal-on-Metal Hip Implants at Risk of Adverse Reactions to Metal Debris. *J Bone Joint Surg Am*. 2016;**98**(8):617-626.

325. Cooper HJ, Urban RM, Wixson RL, Meneghini RM, Jacobs JJ. Adverse local tissue reaction arising from corrosion at the femoral neck-body junction in a dual-taper stem with a cobalt-chromium modular neck. *J Bone Joint Surg Am*. 2013;**95**(10):865-872.

326. Matharu GS, Pandit HG, Murray DW, Judge A. Adverse reactions to metal debris occur with all types of hip replacement not just metal-on-metal hips: a retrospective observational study of 3340 revisions for adverse reactions to metal debris from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. *BMC Musculoskelet Disord*. 2016;**17**(1):495.

327. Kwon YM. Evaluation of the Painful Dual Taper Modular Neck Stem Total Hip Arthroplasty: Do They All Require Revision? *J Arthroplasty*. 2016;**31**(7):1385-1389.