The handle http://hdl.handle.net/1887/25896 holds various files of this Leiden University dissertation

**Author:** Weegen, Walter van der  
**Title:** Metal-on-metal hip arthroplasty: local tissue reactions and clinical outcome  
**Issue Date:** 2014-06-11
Metal-on-Metal Hip Arthroplasty

Local tissue reactions and clinical outcome

Walter van der Weegen

2014
Metal-on-Metal Hip Arthroplasty

Local tissue reactions and clinical outcome

Walter van der Weegen
Metal-on-Metal Hip Arthroplasty

Local tissue reactions and clinical outcome

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit van Leiden, op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker, volgens besluit van het College voor Promoties te verdedigen op woensdag 11 juni 2014 klokke 16:15 uur

door

Walter van der Weegen geboren te Hoogerheide in 1965
Promotiecommissie

Promotor: Prof. dr. R.G.H.H. Nelissen

Copromotores: Dr. P. Pilot (Reinier de Graaf Gasthuis, Delft)
Dr. T. Sijbesma (St. Anna ziekenhuis, Geldrop)

Overige Leden: Prof. dr. J.L. Bloem
Prof. dr. J. Verhaar (Erasmus Universiteit, Rotterdam)
Dr. C.C.P.M. Verheyen (Isala Klinieken, Zwolle)
Prof. dr. J.M. Wilkinson (Sheffield University, Sheffield, UK)
Voor Sylvie
Voor mijn kinderen
Voor mijn ouders
Contents

Chapter 1
Introduction and research questions._________________________________________________ 9

Chapter 2
The development and introduction of metal-on-metal bearing in total
hip arthroplasty______________________________________________________________ 17

Chapter 3
Polyethylene wear in metal-backed cups. A retrospective analysis of 200
uncemented prostheses.
Walter van der Weegen, Shennah Austen, Thea Sijbesma and Henk J. Hoekstra.
Karl Knahr, editor. Springer Heidelberg 2012______________________________________ 31

Chapter 4
Survival of metal-on-metal hip resurfacing arthroplasty. A systematic review
of the literature.
Walter van der Weegen, Henk J. Hoekstra, Thea. Sijbesma, Ellis Bos
Emil H. Schemitsch and Rudolf W. Poolman.

Chapter 5
Hip resurfacing in a district general hospital: 6-year clinical results using
the ReCap hip resurfacing system.
Walter van der Weegen, Henk J. Hoekstra, Thea Sijbesma, Shennah Austen
and Rudolf W. Poolman.
BMC Musculoskeletal Disorders 2012;13;13:247 _____________________________ 81

Chapter 6
High incidence of pseudotumors after hip resurfacing even in low risk
patients; results from an intensified MRI screening protocol.
Walter van der Weegen, Jose M.H. Smolders, Thea Sijbesma, Henk J. Hoekstra,
Koen Brakel and Job L.C. van Susante.
Hip International 2013;23(3):243-9 _____________________________________________ 101

Chapter 7
Comparison of different pseudotumor grading systems in a single cohort
of metal-on-metal hip arthroplasty patients.
Walter van der Weegen, Koen Brakel, Roelof J. Horn, Jorgen A. Wullems,
Dirk Das, Peter Pilot and Rob G.H.H. Nelissen.
Skeletal Radiology 2014; 2014;43(2):149-55 ___________________________________________ 119
Chapter 8

Chapter 9

Chapter 10
Measuring change in pseudotumor dimensions after metal-on-metal hip resurfacing. A pilot study. Walter van der Weegen, Peter Pilot, Renate Raidou, Thea Sybesma, Bart Kaptein and Rob G.H.H. Nelissen Submitted

Chapter 11
General discussion

Chapter 12
Summary

Chapter 13
Samenvatting (summary in Dutch)

Disclosure

Dankwoord

List of publications

About the author
SMITH-PETERTSON PROSTHESIS
ACETABULAR CUP

1937
Chapter 1

Introduction and research questions
1.1 **Hip replacement for younger and active patients.**

Hip replacement surgery is one of the most successful medical procedures performed in an elderly population suffering from disabling osteoarthritis (OA).\(^1\) The current hip replacement surgery has excellent long term results\(^1\), in the Netherlands about 21,000 THA are implanted annually\(^2\) and worldwide an estimated 750,000 Total Hip Arthroplasty (THA) procedures are carried out every year.\(^3\) Despite these large numbers, implant survival and optimal postoperative functioning of both the artificial joint as well as the patient are still challenges in THA. As for the artificial joint, wear-resistance of the bearing surfaces is still a key-issue and one of the challenges, especially in the physically demanding younger patient.\(^4,5\)

During the nineteen-nineties of the last century, Ultra High Molecular Weight Polyethylene (UHMWPE) was considered the benchmark for surface bearings in THA. At that time in our clinic, younger, physically active patients with severe hip osteoarthritis (OA) were treated with an uncemented THA with a standard UHMWPE acetabular liner (ArCom\(^\text{TM}\), Biomet, Warsaw, USA, Figure 1.1).

![Figure 1.1, Acetabular metal shell, UHMWPE acetabular liner and ceramic femoral head.](image)

This type of UHMWPE was compression molded and Argon packaged, to prevent ageing of the material before implantation. The in vivo wear rate of compression molded PE was shown to be 50% less than the more commonly used UHMWPE machined from extruded bars.\(^5\) In a further effort to reduce PE wear, cross-linking of UHMWPE by heat-treatment was developed during the 1990s and quite recently anti-oxidant treatment of this (highly) cross-linked UHMWPE was introduced, by infusion of vitamin E into the UHMWPE material (Figure 1.2).
Also around the millennium, Metal-on-Metal (MoM) was reintroduced as a bearing surface in THA and was promoted to be especially wear resistant in younger- and more active patients. MoM arthroplasty could either be applied as a resurfacing technique or as a ball and socket stemmed THA. Both were used in our clinic, replacing the uncemented THA with a standard, compression molded Argon packaged UHMWPE bearing, with a MoM protheses for indicated patients. The latter would in theory have not only lower wear rates but also, due to the larger femoral head, reduced dislocation rates, as well as preservation of femoral bone stock if a MoM resurfacing design was used (Figure 1.3).

1.2 Bearing surface issues
The orthopaedic literature from 2000 to 2010 is not conclusive on which bearing surface is superior in physically demanding, mostly younger, patients. The discussion on the limited longevity of standard UHMWPE bearings in younger patients\(^5\,^7\) and the demand for a better range of motion stimulated the reintroduction of MoM bearings, which eventually failed dramatically compared

![Figure 1.2](image1.png) Acetabular metal shell with vitamin E infused UHMWPE acetabular liner showing typical orange colouring.

![Figure 1.3](image2.png) Metal-on-Metal bearing surfaces in a hip resurfacing design.
to UHMWPE liner THA designs.\textsuperscript{8,9,10} The latter initiated the studies in this manuscript on MoM Total Hip Prostheses.

1.3 Aim of this thesis
This thesis addresses four main topics related to hip arthroplasty in young active patients with special emphasis on the use of MoM bearing surfaces: (1) A clinical and radiographic evaluation of THA survival in young active patients; (2) A systematic review of the different MoM hip resurfacing systems; (3) A study on the prevalence of Adverse Reaction to Metal Debris (ARMD) with MoM bearings, and; (4) Validation and quantification studies on presence of ARMD after MoM hip arthroplasty at MRI.

1.4 Outline of this thesis
First, to put current issues with MoM bearings in the proper context, a critical review on the development and market (re-)introduction of MoM bearings was done (chapter 2). Next, a retrospective study on radiological liner wear and implant survival of our first 200 uncemented THA procedures with standard UHMWPE in younger, more active patients (chapter 3) was done. At the introduction of MoM hip resurfacing in our clinic (2004), all treated patients were included in a prospective clinical follow up study on implant survival and functional outcomes. Since little was known on survival and outcome of most types of resurfacing hip arthroplasty, prior to analysing the short to mid-term results of this cohort (chapter 5), we systematically reviewed the peer-reviewed literature on implant survival of MoM hip resurfacing (chapter 4). By the end of the first decade of this century, an increasing number of papers were published on the adverse reactions related to in vivo release of metal wear particles. Clarke is recognized as first to start serious discussions on the possible downsides of the “modern” MoM total hip arthroplasty, expressing concern regarding the long term toxicological systematic effects such as immune modulation, chromosomal damage and carcinogenesis in 2003.\textsuperscript{11} The first occurrence of ARMD in response to in vivo released metal ion particles was described in 2008\textsuperscript{12}, with Clayton using the term “pseudotumor”\textsuperscript{13} in relation to current MoM bearings, a term previously introduced by Picard in 1997.\textsuperscript{14} These adverse reactions (pseudotumors) appeared to be related to a variety of factors including implant design characteristics, implant positioning, edge loading
and implant size. These pseudotumors, defined as a peri-articular mass caused by an immunological delayed hypersensitivity response to metal particles and characterised by a lymphocyte-dominated histological pattern\textsuperscript{15} lead to worse clinical outcomes after revision surgery compared to other reasons for MoM revision.\textsuperscript{16} Since pseudotumors are soft tissue masses, they are usually not detected with standard radiographs, although this was until recently the standard method to evaluate MoM case series. At first, most studies focused on metal ion concentrations as an indicator for the amount of wear to predict the occurrence of ARMD. To evaluate the occurrence of pseudotumors in our own cohort of well documented hip resurfacing arthroplasty (HRA) patients, a pilot study using an intensified screening protocol based on Metal-Artefact Reducing Sequence (MARS) MRI was performed (chapter 6). In this study we compared the prevalence of pseudotumors in a subgroup of MoM HRA patients with high risk for pseudotumor to a group with low risk for pseudotumor formation. The validity of pseudotumor classification systems was evaluated as well (chapter 7), and clinical pseudotumor dimension measurements were validated with a three-dimensional region-of-interest based method (chapter 10).

Screening our whole cohort of MoM hip resurfacing patients using metal ions analysis and MARS-MRI for every patient (chapter 8) provided detailed information on the prevalence of pseudotumors. A study on clinical symptoms and differences in MRI findings in unrevised MoM patients with repeated MARS-MRI scans at six to twelve months was done to elaborate on the clinical effect of presence of pseudotumors (chapter 9). A general discussion reflecting on the results of different implant designs and bearing surfaces of the last two decades, the results from studies of this thesis, and directions for future research is presented in chapter 11. Chapter 12 gives a complete summary of this thesis.
References


10. Deutman R. Experience with the McKee-Farrar total hip arthroplasty. 1974; Leiden University, thesis.


JUDET PROSTHESIS
ACRYLIC FEMORAL IMPLANT

1946
Chapter 2

The development and introduction of metal-on-metal bearings in total hip arthroplasty
Hip arthroplasty bearing choice and implant performance

Durability and performance of Total Hip Arthroplasty (THA) implants are still a challenge in younger and physically more demanding patients.\(^1\) Serious progress has been made since the “low-friction bearing” concept developed by Sir John Charnley in the early sixties of the last century\(^2\), resulting in a “forgotten” joint for many THA-patients nowadays.\(^3\) But especially in the younger and more active patients the longevity of the bearing surfaces is the limiting factor for long term implant survival.\(^4\)-\(^6\) Since younger patients tend to be physically more active than elderly patients, their implants have to withstand higher biomechanical stress and these stresses also need to be endured for a more prolonged period of time, leading to a higher risk of dislocation and accelerated wear of the bearing surface. To improve implant fixation and durability and to reduce the risk of other complications (i.e. dislocation, infection), surgeons, engineers and scientists have developed new materials and new surgical techniques, but also introduced new coatings and finishes (i.e. polished) of implants, and designed different implant forms such as collared stems, femoral resurfacing components and modular components. To reduce implant wear, different bearing surfaces were developed. For this purpose, hard-on-hard bearings such as Ceramic-on-Ceramic and especially Metal-on-Metal (MoM) surface bearings have a long tradition in THA. In this chapter we give a brief overview of the development of MoM surface bearings in the history of THA, and discuss how different generations of MoM surface bearings were introduced into clinical practice.

First generation Metal-on-Metal hip arthroplasty

After orthopaedic surgeons experimented with many different interposition materials, such as muscle, celluloid, silver plates, rubber struts and magnesium, Berliner Professor T. Glück (1853-1942) led the way in the development of hip implant fixation using an ivory ball and socket joint that he fixed to bone with nickel-plated screws.\(^7\) The first total hip implant with a MoM articulation is attributed to the British orthopaedic surgeon P. Wiles who, in 1938, implanted a bearing couple made of stainless steel, fixed to the bone with screws and bolts. His results with stainless steel were however disappointing.\(^7\) During this same period, Smith-Petersen introduced the concept of resurfacing the femoral head, using glass, celluloid and pyrex before settling on vitallium in 1938.\(^6\) Vitallium is a chrome-cobalt alloy which is remarkably inert. Wiles at that time however
preferred stainless steel and McKee, another pivotal orthopaedic surgeon in MoM arthroplasty, preferred brass and stainless steel at first. In the 1950s McKee and Watson-Farrar adopted a MoM articulation with a modified Thompson stem, which is considered the start of the first generation of MoM THA. McKee and Watson-Farrar initially treated 50 patients with this MoM prosthesis in which both the acetabular and the femoral component were made of Vitallium, later on they switched to a cobalt-chromium-molybdenum alloy in which both components were fixated to the bone using cement. Although later research showed unacceptably high revision rates, MoM arthroplasty was widely used in the 1960s, with besides the McKee-Farrar design the Ring design (also in the UK), the Mueller-Huggler implant in Switzerland, and the Sivash design in the Soviet Union.

Although many years later retrieval studies of first-generation MoM hips demonstrated low wear rates in individual cases, a tissue reaction to metal particles around MoM total hip prostheses was noticed with some retrieval cases. Large numbers of macrophages with metal particles in tissues around MoM prostheses were seen, with dark tissue staining and osteolysis after MoM arthroplasty being associated with impingement or with loose components. In studies using metal-on-polyethylene prostheses, a wide variety of marked tissue changes were also present around the hip implant, but these tissue response were associated with bone loss, rather than with soft tissue damage.

The disappointing results, poorly understood at that time, and the extraordinary mentoring of the "low friction" THA developed by Charnley, using metal-on-polyethylene bearings, 'red lined' the MoM bearing, and consequently the concept was abandoned before the reasons for its failure had been effectively analyzed. Later studies attributed the failure of MoM bearings to the factor of "high friction" resulting from inadequate manufacturing. By the mid-1970s, MoM had all been rejected in favour of Charnley’s technique for low-friction arthroplasty of the hip using polyethylene (PE). A full recounting of the historical events leading up to the “discovery” of PE for hip arthroplasty in 1962 by Sir John Charnley and his engineering associate, Harry Craven, can be found in Charnley’s monograph and biography, but it is instructive to briefly recall Ultra High Molecular Weight Polyethylene (UHMWPE) arrived in orthopaedics by chance rather than by design. Only after the catastrophic clinical failure of polytetrafluoroethylene (PTFE) and failure of his initial choices, glass-filled
variants of PTFE, that Charnley sought bearing material alternatives. With the introduction of UHMWPE particles in the body, resulting from inevitable wear on the implant bearing surface, the mechanism of osteolysis was described. In this process, UHMWPE wear products are thought to cause massive osteolysis by triggering foreign-body granuloma formation at the bone-cement interface, resulting in implant loosening and ultimately, implant failure. Long term implant survival results of standard UHMWPE are as a result often disappointing, especially for the acetabular component. In retrospect it was recognized that the results of "the McKee" could in fact differ only little from the results of "the Charnley". Some MoM implants by McKee-Farrar and Ring continued functioning extremely well and were "rediscovered" in the 1980s by Swiss and British surgeons. By the late 1980s, concerns over osteolysis attributed to PE wear debris led to the reintroduction of MoM bearings, the development of highly cross-linked PE and the more widespread use of Ceramic-on-Ceramic (CoC) bearings.

CoC bearing surfaces were developed in the early 1970s in France and Germany to reduce wear particles and subsequent osteolysis occurring with polyethylene THA bearings. CoC tribological properties are explained by its low surface roughness, high hardness for major scratch resistance, high wettability and fluid-film lubrication. The initial use of CoC bearings resulted in a high rate of aseptic loosening of the cemented socket and risk of component fracture, mainly related to bad design and material flaws. Incremental improvements in the manufacturing process, design, and quality control have since significantly decreased the risk of fracture to approximately 0.02% to 0.1%. However, there are still concerns regarding fracture of sandwich ceramic liners, squeaking, and impingement of the femoral neck on the rim of the ceramic liner leading to chipping, especially in younger and physically active patients, and according to a recent systematic review by Gallo et al, the use of CoC bearings leads to equivalent but not improved survivorship at 10 years follow-up compared to the best non-CoC THA. This is also shown in the 2012 Annual report of the Australian Joint Replacement Register, where the Yearly Cumulative Percent Revision rate for CoC bearing at 10 years is 4.8% for fixed femoral neck types, 9.8% for exchangeable femoral neck types and 4.6% for Metal-on-Polyethylene (MoP) bearing devices. Randomized clinical trials comparing CoC versus PE bearings also show similar clinical outcomes and dislocation rates between both groups.
Second generation Metal-on-Metal hip arthroplasty

After identifying the "polyethylene disease" in the beginning of the 1980s, MoM bearings were considered again as an alternative to PE. An elaborate analysis of the first generation MoM failures led Weber to initiate and then promote a second generation of MoM, cemented at first, then rapidly followed by non-cemented prostheses. At that time, the advantages of MoM were put into perspective due to survivorship analysis of the Charnley versus McKee-Farrar prostheses. Analysis of these results supported the reintroduction of MoM bearings in 1988. It was stated that significant numbers of MoM bearings were surviving at long term, due to polar bearing, a component orientation which avoided impingement and good cementation.

At the time of re-introduction, wear simulation tests showed that wear rates of second generation MoM bearings were 20 to 100 times lower compared to metal-on-conventional polyethylene, and MoM bearing couples started to experience widespread clinical use in both hip resurfacing and total hip arthroplasty. The material properties allowed the use of large heads in thin acetabular shells, promising of a reduced incidence of hip dislocation in younger and more active patients. From their arrival in the orthopaedic market in 1997, MoM bearings were strongly marketed as the latest advance in hip replacement and were targeted at young active patients who needed a hip that would last a whole lifetime.

In the case of MoM hip resurfacing, patients organised themselves on internet and started forums on this topic, as for example www.surfacehippy.info. In the same time, critical reports on the limitations of MoM hip resurfacing discussed poor medium term outcomes, with a two- to threefold difference in revision rate between different makes, based on the observation that prostheses were different in many details, such as shape, sizing, head coverage, clearance, metal alloy used, heat treatment, instrumentation, and so on. During the first years of the reintroduction however, resurfacings became very popular and the number of implantations rose to about 10-20% of all primary hip replacements in countries such as the UK, Australia, and the Netherlands. There remained a concern on the metal ion release over time and the potential detrimental effects of accumulated metal ions in the body. Its particular complication, Aseptic Lymphocytic Vasculitis-associated Lesions (ALVAL), was documented by Willert.
further development of this bearing, although at the time of introduction no complication could be attributed to this phenomenon.\(^\text{14}\) Throughout literature a variety of nomenclature describing implant failure mode as a reaction to metal wear particles is used, most notably the terms ALVAL\(^\text{34}\), pseudotumor\(^\text{35}\) and metallosis.\(^\text{36}\) Langton et al described a new umbrella term for these modes of failure: Adverse Reaction to Metal Debris (ARMD), to include MoM joint failures associated with pain, a large sterile effusion of the hip and/or macroscopic necrosis and metallosis.\(^\text{37}\) The end modes of failure requiring revision are ALVAL (a histological diagnosis made from tissue sampling at the time of surgery identifying an abundance of lymphocytes in the local pericapsular tissue)\(^\text{34}\) and pseudotumor (the development of a cystic mass in the periarticular region, which has a direct communication with the joint).\(^\text{35}\) These pseudotumors can be very large, extend into the pelvic region, can involve destruction of bone and muscle tissue, and compress vital surrounding structures such as nerves and blood vessels.\(^\text{38-40}\) Another concern with the toxicity of released chromium and cobalt is the increased risk for cancer but large comparative studies have demonstrated so far that patients with MoM hip prostheses were not at increased overall risk for cancer.\(^\text{41,42}\)

**Complexity of introducing new bearing devices into clinical practice**

Since there is an increasing necessity for innovative surgical techniques and designs for orthopedic surgeons to meet the demands of increasingly younger and more demanding patients, there is an inherent risk in the introduction of these innovations. Under current regulations, clinically important unknown modes of failure for newly introduced devices may not become known for several years after widespread adoption, affecting a large number of patients.\(^\text{43}\) There is a conflicting interest of making promising new hip implant materials and designs available so patients can benefit as soon as possible, and the fact that these same joint replacement devices have to perform well for over more than 10 years and preferably more than 20 years after implantation in the patient. These requirements make it difficult to design a model for market introduction that effectively and safely guards these requirements, without delaying the needed innovations. For example if more clinical trials are needed in one country before a device can be used in clinical practice, patients might prefer to have surgery in neighbouring countries where these specific requirements are absent. If a medical
device company spends more time on clinical research, including longer follow up studies before releasing new implant designs, other companies might actually introduce comparable devices with less clinical support in the mean time. Another concern is the absence of a clear definition of what is considered a new implant design. Typically, hip replacement devices undergo minor design alterations several times during their lifespan, for example CCD angle, conus, coating or just the manufacturing process. Although one has to bear in mind all this work is done with the benefit for patients in mind, it is not always clear how these minor design changes will effect implant performance.

When adopting more extensive regulations for the introduction of medical devices, it is therefore important this should be done in collaboration with all stakeholders involved. Of course the company which has developed the new implant, notified bodies, competent authorities (national authorities such as the Food and Drug Administration in the United States), the orthopaedic surgeons who have to start using this particular device, and last but not least the patient receiving the new implant. Increasingly, health insurance companies and hospital administrators are also influencing which devices are used by the professionals.

In general, we can conclude that the process of bringing medical devices into clinical practice is complex due to a discrepancy between the interest of introducing newer designs fast, the need for long term clinical data collection on implant performance, the involvement of many stakeholders, and lack of consensus on the definition of a new implant. There is both room to improve market introduction (or re-introduction) regulation and supervision by post market clinical research. A more gradual introduction of new implants, with the appropriate research modality should strike a balance in encouraging new technology which might improve clinical outcomes, while protecting patients from being exposed to new products which may produce unexpected complications. As witnessed with the re-introduction of MoM bearings in THA, serious complications which were unforeseen at the time of introduction became only known after a large number of patients worldwide (an estimated 1 million patients) had become at risk. In this particular case, RadioStereometric Analysis (RSA) studies which are nowadays considered an integral part of gradual introduction into practice did not detect these unforeseen complications of adverse soft-tissue reactions. However, better analysis of-, and anticipation on-, previous failure modes probably would have detected possible down sites of
MoM earlier. Consensus on who has to collect, manage and report this data is however not easily reached, with different interests of involved stakeholders on this topic. It is however clear that increasing post marked clinical research efforts in orthopaedic surgery might protect patients from unnecessary harm, reduce costs by preventing expensive revision surgery and by preventing loss of mobility and productivity in patients.

In conclusion, MoM surface bearings have a long history of use in total hip arthroplasty, with two distinct generations of these bearings. Reduced wear volume, the major advantage of the second generation MoM as seen with in vitro testing, was seriously challenged with in vivo use where less than optimal implant positioning resulted in edge loading and unexpected high wear. The released wear particles induced an, initially less known, local tissue response which is now generally known as ‘Adverse Reaction to Metal Debris’, of which the clinical importance is not yet fully understood. This unexpected failure mechanism has raised concern on how medical devices, including hip implants, are introduced into the market and has intensified the discussion on how to regulate this complicated process.
References


SAN-BAW PROSTHESIS
IVORY IMPLANT

1960
Chapter 3

Polyethylene wear in metal-backed cups. A retrospective analysis of 200 uncemented prostheses.

Walter van der Weegen, Shennah Austen, Thea Sijbesma and Henk J. Hoekstra

Introduction
Uncemented total hip prostheses were introduced some 40 years ago, after disappointing results with cemented hip prostheses in young and active patients. In orthopedic literature, research on uncemented hip prostheses has focused on the survival of the uncemented femoral stem, and in general, excellent results were reported. Although the femoral component showed excellent performance, recent in vivo studies have reported increased wear of the polyethylene (PE) liner of the uncemented acetabular cup. This PE wear results in PE particles being distributed in the tissue surrounding the prosthesis, with macrophages being activated by these particles. These activated macrophages induce osteolysis (Figure 3.1) which in the end results in aseptic loosening of the prosthesis.

![Figure 3.1, Osteolysis.](image)

Although uncemented hip prostheses vary greatly in design, they all have a metal-backed acetabular cup (Figure 3.2). This metal-backing is needed since direct contact between bone and PE results in osteolysis. Metal-backed cups are made more biocompatible by applying coatings which stimulate bone ingrowth. These coatings are either porous or hydroxyapatite (HA) coatings. When metal-backed cups were developed, a better force distribution with less peak forces was expected along the bone-prostheses interface.
Recent studies however stated there was less stress shielding with cemented cups than with uncemented cups.\textsuperscript{14,44} Another possible disadvantage from metal-backed cups is the dislocation or rotation of the PE liner from the metal-backing, resulting in additional wear and an increased number of released PE particles. This type of wear is known as “backside” wear.\textsuperscript{3,31}

![Figure 3.1](image)

\textbf{Figure 3.1}, Metal-backed acetabular cup.

Increased PE wear is most likely a multifactorial process influenced by, for example, the manner in which PE is produced and sterilized, the time between production and implantation (known as “shelf life”), the inclination angle of the cup, and the activity levels of the patient. Since we had concerns on the frequency of observed wear in our patient population, we retrospectively reviewed our first 200 uncemented hip prostheses using the Mallory-Head design (Biomet Inc., Warsaw, USA). The long-term survival of the femoral component of this particular prosthesis is well documented and has excellent results. Only a few studies report on acetabular wear and survival using this design. Yamamoto et al found a mean liner wear of 0.3 mm after 3 years, 0.55 mm after 5 years, and increasing to 0.7 mm after almost 7 years of follow-up.\textsuperscript{61} Kurtz concluded that the threshold for osteolysis is a head penetration rate of >0.1 mm per year. He also reported that osteolysis could not be detected with a head penetration rate of <0.05 mm per year.\textsuperscript{32} Other studies reported an osteolysis threshold at a head penetration rate of 0.1–0.2 mm per year.\textsuperscript{15,16,33,53,55} We therefore used a head penetration rate of >0.2 mm per year to classify any case as excessive wear. The primary objective of our study was to evaluate how many of the 200 implanted prostheses showed a
liner wear of more than 0.2 mm per year. The frequency of any osteolysis and implant survival was also evaluated.

Patients and Methods

Our first consecutive 200 uncemented total hip prostheses (Mallory-Head), implanted between November 1997 and September 2002, were retrospectively analysed (Table 3.1).

<table>
<thead>
<tr>
<th>Table 3.1, Patient demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (n)</td>
</tr>
<tr>
<td>Female (n)</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Bilateral (n)</td>
</tr>
<tr>
<td>Diagnosis: OA</td>
</tr>
<tr>
<td>AVN</td>
</tr>
<tr>
<td>FC</td>
</tr>
</tbody>
</table>

*OA: osteoarthritis; AVN: Avascular Necrosis; FC: fractured collum

In all cases, an uncemented porous-coated femoral stem was used with a 28-mm ceramic head and a porous-coated ringloc acetabular cup. The liner was made of conventional ultra-high molecular weight polyethylene (UHMWPE) (ArCom®, Biomet Inc., Warsaw), manufactured with compression molding and sterilized with gamma radiation in argon gas. Liner thickness ranged from 4.8 (cup size 48) to 11.8 mm (cup size 62). Mean shelf life was four months (range: 0 to 41). All prostheses were implanted through the posterolateral approach. All patients were asked to return for clinical follow-up including a standard anteroposterior (AP) radiograph. Medical file data were collected on primary diagnosis, BMI, complications and details of the used components. Of all patients, 89% completed a Duke Activity Index to measure current activity levels. There were 36 patients lost to follow-up (37 prostheses): 9 were deceased, 16 were revised, and we were unable to contact 10 patients. This left us with 163 prostheses (81.5%) available for analysis of PE wear. Liner wear was evaluated by measuring the two-dimensional displacement of the femoral head relatively to the cup position using software (Pro 3D software, Draftware Inc. Vevay, USA). We used the most recent AP radiograph (Figure 3.3). To check for interobserver reliability, a sample of ten
radiographs was measured by an experienced evaluator of Draftware Inc., and all 
PE wear measurements were 100% identical.

This is possible by using edge-detection features in the software, limiting the 
observer input on the obtained measurements. Besides the use of software, we 
retrospectively checked medical files if PE wear was noted by the orthopedic 
surgeon. We set the threshold for acceptable wear at <0.2 mm per year. A 
sensitivity analysis with a threshold of 0.1 mm per year was also calculated. We 
calculated the correlation between wear and the following subgroups: age, BMI, 
activity level, cup inclination angle, acetabular component size, liner thickness, 
and shelf life. Differences in wear between male and female patients were tested 
using an unpaired Student’s t-test. Implant survival was calculated using the 
Kaplan-Meier (KM) method. All statistics were performed using SPSS software 
(SPSS Statistics, version 17.0, IBM Corporation, Somers, USA). The most recent AP 
radiograph was screened for any radiolucency or osteolysis according to the zones 
described by DeLee and Charnley for the acetabular component and the zones 
described by Gruen for the femoral component.\textsuperscript{13,24}
Results

Wear and Osteolysis

The mean-measured PE wear was 0.2 mm per year (range: 0.07 to 0.5), after a mean follow-up of 8.3 years. In 53.4% of all cases, the PE wear was 0.2 mm per year (Figure 3.4), and if the threshold for acceptable wear was set at 0.1 mm per year, 96.3% of all liners showed excessive PE wear.

![Boxplot of wear rate per year.](image)

There was a significant correlation between PE wear and cup inclination angle and between PE wear and component size (Table 3.2). Mean PE wear was significantly higher in male patients than in female patients (respectively, 0.22 mm per year versus 0.19 mm per year, p = 0.02). On average, 24.3% of the original liner thickness was lost to PE wear (range: 10.7 to 42.7%). In 41 cases, PE wear was observed during routine clinical follow-up and noted in the medical file (24.8%), with a mean of 93 months after index surgery (range: 40 to 120). Osteolysis was observed in five cases (Table 3.3). The measured PE wear in these five patients had a mean of 0.22 mm per year (range: 0.19 to 0.26).
Table 3.2, Sub analyses PE wear

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.4</td>
<td>0.61</td>
</tr>
<tr>
<td>BMI</td>
<td>0.056</td>
<td>0.48</td>
</tr>
<tr>
<td>Activity level</td>
<td>0.166</td>
<td>0.053</td>
</tr>
<tr>
<td>Acetabular inclination</td>
<td>0.236</td>
<td>0.002*</td>
</tr>
<tr>
<td>Shell size</td>
<td>0.156</td>
<td>0.046*</td>
</tr>
<tr>
<td>Shelf life</td>
<td>0.065</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 3.3, Osteolysis

<table>
<thead>
<tr>
<th>Femoral component</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>160</td>
<td>98.2</td>
</tr>
<tr>
<td>- Gruen zone 1 or 7</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>- Gruen zone 2 – 6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acetabular component</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>- None</td>
<td>158</td>
<td>96.9</td>
</tr>
<tr>
<td>- DeLee &amp; Charnley zone 1</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>- DeLee &amp; Charnley zone 2</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>- DeLee &amp; Charnley zone 3</td>
<td>1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Implant Failure

Of the 200 prostheses, 16 were revised, and one was scheduled for revision. Most frequent reason for revision was PE liner wear (N=10), see tables 3.4 and 3.5. Of the ten patients revised for liner wear, a straightforward cup exchange was done in nine cases. In two cases, the liner was detached from the metal-backing, and in one of these two cases, metallosis was observed. In the other case, a fibrous tissue layer was observed between the PE liner and the metal-backing. Four cases needed bone impaction grafting for an acetabular cyst. Mean time to revision was 108 months (range: 77 to 144), and the mean observed wear in the revised patients was 0.28 mm per year (range: 0.21 to 0.45). The KM probability estimate of survival, with revision for any reason as end point, was 90.7% after 12 years of follow-up (95%–CI: 85.6–94.2). With only revision cases due to wear as end point, the KM survival estimate was 93.1% after 12 years follow-up (95%–CI: 79.9–100), see figure 3.5.
Table 3.4, Overview of revision cases

<table>
<thead>
<tr>
<th>Reason for revision</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-septic loosening</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Liner exchange</td>
<td>9 (4.5)</td>
</tr>
<tr>
<td>Dislocation</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Wound infection</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Breakage ceramic head</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Total</td>
<td>16 (8)</td>
</tr>
</tbody>
</table>

Table 3.5, Wear related revision

<table>
<thead>
<tr>
<th>Casus</th>
<th>Months to revision</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>2</td>
<td>104</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>3</td>
<td>107</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>4</td>
<td>107</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>5</td>
<td>109</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>6</td>
<td>109</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>7</td>
<td>110</td>
<td>Liner exchange, components well fixed</td>
</tr>
<tr>
<td>8</td>
<td>144</td>
<td>A-septic cup loosening</td>
</tr>
<tr>
<td>9</td>
<td>Unknown</td>
<td>Revised in other hospital, patient deceased</td>
</tr>
<tr>
<td>10</td>
<td>Planned</td>
<td>Wear observed</td>
</tr>
</tbody>
</table>

Figure 3.5, Kaplan-Meier survival probability estimate.
Discussion

In our study, we report a high proportion (53.4%) of UHMWPE liners with a wear rate of 0.2 mm per year, after a mean follow-up of 8.3 years. In contrast, implant survival after 12 years is acceptable (KM 90.1%). However, it is disturbing that in literature the liner wear rate is reported to be nonlinear, with an increase in PE wear 7 to 8 years after index surgery. These findings suggest that we have to expect an increasing number of revisions within the next few years of follow-up.

Parvizi conducted a study with longer mean follow-up than our study and found a revision rate of 20% after 11 years of follow-up. And McLaughlin reported a revision rate of 65% after 16 years. A possible explanation for the measured amount of wear can be found in the type of PE used. Free radicals, formed during the sterilization process, negatively influence the characteristics of conventional UHMWPE. Before and after implantation, these free radicals react to oxygen. This oxidation leads to accelerated wear rates. Wear can be reduced by using highly cross-linked polyethylene (HXLPE). Compared to conventional PE, HXLPE shows a significant reduction of the head penetration rate in several clinical studies. Currently, we do not know if in the long term, free radicals are released from HXLPE and can still cause oxidation. A recent method to prevent this happening is the infusion of vitamin E into (highly cross-linked) PE to scavenge any free radicals. This method is too new for clinical studies to be available. Alternatively, other bearing materials may be used such as metal or ceramics. Although there are some benefits of Metal-on-Metal (MoM) bearings such as low dislocation rates (due to the large diameter) and very low wear rates reported in in vitro studies, these benefits are outweighed by the occurrence of serious complications due to an adverse reaction to metal debris (ARMD), as reported in recent clinical studies. In general, recent clinical studies using MoM bearings report higher revision rates than expected with the introduction of these bearings. Clinical studies with ceramic bearings have good long-term results, but the use of ceramics is limited by high cost, “squeaking,” and difficult revision after liner fractures. The choice of material for the femoral head does not influence the PE wear rate significantly; only small differences in liner wear were observed between different materials for the femoral head. Wear is not only dependent on the used materials but indeed multifactorial. In our study cohort, more wear was observed in cups with a steeper inclination angle and in male patients. This corresponds with earlier publications.
earlier studies, we observed more wear with larger sizes of acetabular cups. We could not identify any possible explanation for this observation. We explored the hypothesis that larger cups would be more difficult to place, resulting in steeper cup placement. However, there was no significant difference in cup inclination angle between smaller (54 mm) and larger (56 mm) cup sizes. From our analysis on different subgroups, we could not detect any relation between age, BMI, shelf life or activity level, and the measured PE wear in our study cohort. This was unlike findings from other studies. There is however a large heterogeneity in number and characteristics of the included patients, making it difficult to compare these results. In our study, shelf life was quite short with an average of four months. The measured wear in all patients revised because of liner wear was more than 0.2 mm per year. However, 82.5% of all our patients with a PE wear rate of >0.2 mm per year had no radiolucent zones, no cyst formation, or such clinical symptoms that revision surgery was indicated. This might be due to the genetic profile of these patients, which makes them resistant to osteolysis. The observed wear in our study is comparable to other studies using metal-backed cups. Considering this comparable high wear rate, the number of cases with aseptic loosening (0.5%) and the number of observed osteolysis (5.5%) in our series is low in comparison to other studies. Although, most of these other studies had longer follow-up the retrospective nature of our study which makes it more difficult to classify aseptic loosening. Another explanation might be that the osteointegration of the coating is so effective that the acetabular component appears to be well fixed in place during revision surgery. Even if only a small area is integrated into the bone tissue, the optimal treatment if wear is observed and the best timing to perform revision surgery are clinical issues described in a treatment algorithm by Goosen et al (Figure 3.6). Strong points of our study are the large number of included prostheses, the use of a validated method to measure wear, and the analyses of multiple variables which might influence than our series. For example, Emms et al found a 17.1% osteolysis rate and a wear-related revision percentage of 20% after 11.5 years of follow-up. The fact that we only used the most recent radiograph for PE wear evaluation, might explain we only observed osteolysis instead of any radiolucency. It is also striking that the number of cases with aseptic loosening in our study cohort is very low.
This might either be because we revised patients early or by wear. Our study is limited by the retrospective design, the lack of a control group, the loss to follow-up, and the limited duration of the follow-up. Based on our results and the current literature, we strongly question the use of conventional UHMWPE in uncemented total hip prostheses with metal-backed cups. Detailed follow-up, especially in the long term, can prevent serious complications due to the use of conventional PE. Studies with longer follow-up, preferably more than 10 years, are necessary to validate the safety of conventional UHMWPE in uncemented total hip prostheses.
References


MCKEE-FARRAR METAL-ON-METAL TOTAL HIP PROSTHESIS

1965
Chapter 4

Survival of metal-on-metal hip resurfacing arthroplasty. A systematic review of the literature.

Walter van der Weegen, Henk J. Hoekstra, Thea Sijbesma Ellis Bos, Emil H. Schemitsch and Rudolf W. Poolman

Abstract
We systematically reviewed the peer-reviewed literature to relate the survival of hybrid Metal-on-Metal hip resurfacing arthroplasty devices to a National Institute of Clinical Excellence (NICE) benchmark for choosing a primary total hip replacement, which is a survival rate of 90% at a follow-up of ten years. A total of 29 articles (10,621 resurfaced hips) met the inclusion criteria. The mean follow-up ranged from 0.6 to 10.5 years and the survival of the implant ranged from 84% to 100%. Of the 10,621 hips, 370 were revised (3.5%), with aseptic loosening as the most frequent mode of failure. None of the hip resurfacing arthroplasty implants used to date met the full ten-year NICE benchmark of survival. A total of 13 studies showed satisfactory survival compared with the three-year NICE benchmark.

Introduction
Hip resurfacing arthroplasty (HRA) has regained popularity since the introduction of the third generation of implants in the mid-1980s. Both the first- (Metal-on-Polyethylene) and the second-generation (cementless Metal-on-Metal) resurfacings failed because of high rates of wear and aseptic loosening. The current third generation hip resurfacing implants consist of a cemented femoral component and a press-fit acetabular component. Some surgeons are hesitant to use HRA because of the failure rates of the first- and second-generation implants and the complications, which include fracture of the femoral neck, metal hypersensitivity and increased serum levels of metal ions. Those in favour of the technique indicate the possible advantages of conserved femoral bone stock, minimal wear and a reduced risk of dislocation due to the large diameter of the components. These advantages would suit the lifestyle of younger patients. With HRA promoted for use in young active patients, its use may not be entirely comparable with total hip replacement (THR). There remains a continuing debate on the possible advantages of HRA. The National Institute for Health and Clinical Excellence (NICE), as part of the National Health Service (NHS) for England and Wales, has indicated that a revision rate of 10% or less at ten years should be regarded as the current benchmark of the satisfactory performance of a primary THR. This applies to all forms of replacement including both conventional and resurfacing implants. Prostheses unable to satisfy these requirements should be appropriately investigated. In its appraisal of THRs,
implants may also be recommended if their reported implant revision rate at a follow-up of at least three years is consistent with this ten year benchmark.\textsuperscript{17} Although several reviews on HRA have been published recently, no studies have compared the survival of the HRA implant with an objective benchmark.\textsuperscript{18,19} In our systematic review, we hypothesised that primary hybrid Metal-on-Metal HRA is compliant with the NICE benchmark of a revision rate of 10% or less at a follow-up of ten years.

**Materials and Methods**

The Cochrane Library, EMBASE and MEDLINE electronic bibliographic databases were searched by an independent librarian. The search was conducted using standard software (Pubmed 2009 database for searching MEDLINE, OVID software (OvidSP_U102. 03.00.130; Ovid Technologies, Sandy, Utah) for searching EMBASE). The electronic search included articles published until June 2010. In combination with the booleans ‘AND’, ‘NOT’ and ‘OR’ the following search terms were used, with asterisks indicating where truncated search terms were used to yield the widest ranges of results: hip, femur head, femoral head, femur neck, femoral neck, resurfac*outcome*, follow-up, FU, prosthesis failure, treatment failure, re-operation, longevity, success, recovery of function, range of motion, joint instability, osteonecrosis, osteoarthritis, pseudotumor, pseudotumour, mechanical stress, gait, patient satisfaction, activity, activities, surviv* and risk factors. Reference lists in the included studies were hand searched for other relevant studies. Although only peer-reviewed publications were considered for inclusion, we tried to include all available studies by asking all implant manufacturers if they were aware of any (un-)published data. Also, experts in this field were contacted to determine if there were unpublished data. All the titles and abstracts were examined to assess their relevance. Only studies meeting the following eligibility criteria were included: any systematic review, clinical trial or case series using a Metal-on-Metal resurfacing prosthesis with a cemented femoral component and an uncemented acetabular component implanted after 1988; reports of the survival of the implant defined as time to revision; a minimum requirement of 75 HRA procedures to ensure that the learning curve was completed;\textsuperscript{20-26} basic clinical details including age, gender and aetiology; validated patient reported outcomes of pain, stiffness, functional impairment and quality of life such as the Harris hip score (HHS)\textsuperscript{27}, the Oxford hip score\textsuperscript{28}, the
University of California, Los Angeles (UCLA) hip rating system\textsuperscript{29}, the score of Merle d’Aubigné and Postel\textsuperscript{30}; and the mechanisms of failure such as fracture of the femoral neck and aseptic loosening. No language restrictions were applied. A native speaker was consulted for articles published in languages other than English. Case reports and articles published before 1988 were excluded, since current implants on the market had been introduced after 1988.\textsuperscript{1} The inclusion and exclusion criteria were checked in all identified abstracts by two independent reviewers (WvdW,TS). In case of disagreement, a third reviewer was consulted. The full texts were retrieved and further checked for inclusion and exclusion criteria. If articles described the same series of patients, only the most recent with the largest number of patients was included. Extraction of data focused on the baseline clinical details and aetiology, the types of implant used, details of follow up, standardised clinical scores, radiological findings, implant survival rates, complications not requiring revision and the modes of failure. Data extraction was undertaken by one author (WvdW) and validated by a second (TS). A third was consulted if there was disagreement. All the extracted data were summarized and pooled whenever possible. The survival rate of the implants was plotted against the follow-up mean for comparison with the NICE benchmark. The quality of the evidence was judged using the Grading of Recommendations Assessment, Development and Evaluations (GRADE) recommendations, resulting in the grading of quality as high, moderate, low or very low.\textsuperscript{31} In order to compare different clinical scores, we normalized the scores to a new score whenever possible, ranging from 0 to 100, with 100 being the best possible score.

Results

We identified 539 abstracts. Data were extracted from 29 papers. A flow chart, compliant with the Quality of Reporting of Meta-Analysis statement, detailing the study selection, is presented in figure 4.1. Four studies investigated the ASR hip resurfacing device (DePuy Orthopaedics Inc., Warsaw, Indiana), 13 the BHR (prior to October 2008: Finsbury Orthopaedics, Leatherhead, United Kingdom; thereafter Smith & Nephew Inc., Memphis, Tennessee), four the Conserve Plus (Wright Medical Technology Inc., Arlington, Tennessee), two the Cormet 2000 (Corin Group PLC, Cirencester, United Kingdom) and four the Durom (Zimmer, Warsaw, Indiana).\textsuperscript{9,10,22,24,25,33-54}
One study described the use of both the McMinn (Corin Group PLC) and the BHR device\textsuperscript{11} and another a second-generation McMinn prosthesis (McMinn Hybrid Resurfacing; Corin Group PLC), which was only in use around 1996.\textsuperscript{55} Data were thus presented on five of 11 resurfacing devices on the market. We could not identify studies which met our inclusion criteria describing the use of the Accis (Implantcast GmbH, Buxtehude, Germany), Adept (Finsbury Orthopaedics), Eska-Bionik (Eska Implants, Lubeck, Germany), Icon (International Orthopaedics, Geisingen, Germany), Mitch (Stryker, Kalamazoo, Michigan) or ReCap (Biomet Inc., Warsaw, Indiana) resurfacing devices. The studies included one randomised,
clinical trial, 27 prospective case series and one retrospective case series. The mean follow-up ranged from 0.6 to 10.5 years. The highest reported loss to follow-up was 8.3% and 11 studies reported no loss to follow-up. The survival of the implant ranged between 84% and 100% (Figure 4.2). In 13 of 17 studies with a follow-up of between three and 11 years, the survival rate was compliant with the NICE benchmark. These 13 studies used either the BHR implant (eight), the Conserve Plus (two), the Durom (one), the Cormet 2000 (one) or both the McMinn and the BHR implants (one). The four studies not compliant with this benchmark, but with follow-up of more than three years, used either the ASR device (two), the BHR (one) or the second-generation McMinn device (Figure 4.2). The only randomised, controlled trial reported a lower survival rate for the HRA group compared with the THR group, 96.3% versus 98% at 5.6 years. The mean follow-up of more than three years, used either the ASR device (two), the BHR (one) or the second-generation McMinn device (Figure 4.2). The only randomised, controlled trial reported a lower survival rate for the HRA group compared with the THR group, 96.3% versus 98% at 5.6 years.

Figure 4.2, Implant survival versus time.
up was less than three years in 12 studies. Details on the number of hips per study, the clinical details, the duration of follow-up, the loss to follow-up and implant survival are presented in table 4.1. Careful patient selection based on gender, age and the preoperative diagnosis is important for HRA. In two of the 29 studies which were included, most of the patients were female, reporting a survival rate of 88.7% at one year and of 94.2% at a mean of 2.8 years. With these results, neither study was compliant with the NICE benchmark. The mean age of the patients in the included studies ranged from 42 to 58 years. In six studies the mean age was more than 55 years, and three of these studies had an implant survival rate higher than that required by the three-year entry NICE benchmark. In three studies, all the included patients had a preoperative diagnosis of primary osteoarthritis. Two were compliant with the NICE benchmark. In two less than 50% of patients had primary osteoarthritis as the preoperative diagnosis. One was compliant with the NICE benchmark, the other was not. Four studies presented details on the surgeon’s learning curve. All reported that there were fewer cases of revision at the end of the series compared with the beginning. According to the GRADE recommendations, the quality of the evidence was very low. The clinical outcome was presented in a very heterogeneous manner, using six different scoring systems, some in modified form. Most frequently used (22 studies) was the HHS and in those studies, the mean score improved by 40.8 points (95% confidence interval (CI) 36.6 to 45.2), from 52.2 pre-operatively to 93.0 post-operatively. The clinical score improved significantly in all studies (Table 4.2). The radiological findings were also reported very heterogeneously. Seven studies did not report any radiological details. A summary of the radiological findings is shown in table 4.3. The postoperative levels of metal ions in the blood were not reported in any of the included studies. That by Ollivere et al investigated the association between early clinical failure with metallosis and soft-tissue necrosis. This response was possibly due to an acquired sensitivity to metal ions, leading to aseptic lymphocytic-vasculitis associated lesions (ALVAL). At follow-up at five years the rate of revision which was related to the metallosis was 3.1%. The reported risk factors for metallosis were female gender, a small femoral component, a high abduction angle and obesity. Three other studies reported a marked inflammatory response when performing a revision procedure for pain. The 29 studies represented a total of 10 621 HRA procedures. A total
of 370 were revised (3.5%). The reported reasons for failure were aseptic loosening (1.4%), fracture of the femoral neck (1.1%), infection (0.2%), avascular necrosis of the femoral head (0.2%), ALVAL (0.13%), persistent pain (0.1%), dislocation (0.08%), malpositioning of a component (0.08%) or other unspecified reasons (0.2%). The study by Daniel et al\textsuperscript{55} using a specific series of double-heat-treated resurfacing devices, which are no longer in use, can be regarded as an outlier. In this series, 16% of hips were revised for aseptic loosening. In all other studies, this percentage was less than 6% (Table 4.3). Clinical complications and adverse events without the need for revision were reported in 25 of the 29 studies. In these 25 studies (9446 patients), 529 complications were reported (5.6%). Steffen et al\textsuperscript{10} reported no major complications without providing further details. In their series of 337 HRAs, Stulberg et al\textsuperscript{50} reported hip-related complications in 83 (24.6%) and implant-related complications in 32 (9.5%), without further details. Beaulé et al\textsuperscript{47} reported re-operation on 28 patients (24.1%) because of loosening of the internal fixation in seven and complete removal of the fixation because of bursitis in 21. Both reasons for reoperation were specifically associated with the Ganz approach used in this series. Based on the remaining studies involving HRAs (8338) in which the complications were reported in detail, the most frequent was painless clicking of the hip (1.2%), followed by a nerve palsy (0.8%), deep-vein thrombosis (0.6%), dislocation (0.3%), squeaking (0.2%), wound infection (0.1%) and pulmonary embolism (0.1%). Heilpern et al\textsuperscript{37} and Witzleb et al\textsuperscript{45} each reported one patient with an undisplaced fracture of the femoral neck which healed with conservative treatment.
<table>
<thead>
<tr>
<th>Study design</th>
<th>Implant</th>
<th>Hips</th>
<th>% Male</th>
<th>Mean age</th>
<th>% prim OA/Sec OA/Other</th>
<th>Mean FU (yrs)</th>
<th>% LTFU</th>
<th>Revisions (n)</th>
<th>Survival% (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amstutz 2010</td>
<td>Conserve plus</td>
<td>1107</td>
<td>74</td>
<td>50</td>
<td>NS</td>
<td>6.8</td>
<td>0.2</td>
<td>44</td>
<td>96</td>
</tr>
<tr>
<td>Aulakh 2009</td>
<td>BHR</td>
<td>202</td>
<td>75</td>
<td>42</td>
<td>50/0/50</td>
<td>7.5</td>
<td>0</td>
<td>6</td>
<td>97.7AVN group; 95 OA group</td>
</tr>
<tr>
<td>Beaulé 2009</td>
<td>Conserve plus</td>
<td>116</td>
<td>81</td>
<td>47</td>
<td>81/9/10</td>
<td>3.2</td>
<td>1.9</td>
<td>2</td>
<td>98.1</td>
</tr>
<tr>
<td>Bergeron 2009</td>
<td>ASR</td>
<td>228</td>
<td>80</td>
<td>54</td>
<td>97/0/3</td>
<td>4.6</td>
<td>3</td>
<td>8</td>
<td>94.8</td>
</tr>
<tr>
<td>Daniel 2004</td>
<td>McMinn and BHR</td>
<td>446</td>
<td>79</td>
<td>48</td>
<td>100/0/0</td>
<td>3.3</td>
<td>0</td>
<td>1</td>
<td>99.8</td>
</tr>
<tr>
<td>Daniel 2010</td>
<td>McMinn secd gen</td>
<td>184</td>
<td>59</td>
<td>54</td>
<td>82/0/18</td>
<td>10.5</td>
<td>0.5</td>
<td>30</td>
<td>84</td>
</tr>
<tr>
<td>DeSmet 2005</td>
<td>BHR</td>
<td>252</td>
<td>69</td>
<td>50</td>
<td>81/6/13</td>
<td>5</td>
<td>1.5</td>
<td>3</td>
<td>98.9</td>
</tr>
<tr>
<td>Giannini 2007</td>
<td>BHR</td>
<td>350</td>
<td>52</td>
<td>51</td>
<td>52/22/26</td>
<td>2</td>
<td>NS</td>
<td>4</td>
<td>98.8</td>
</tr>
<tr>
<td>Gravius 2009</td>
<td>Durom</td>
<td>82</td>
<td>56</td>
<td>53</td>
<td>93/7/0</td>
<td>2.4</td>
<td>2.4</td>
<td>2</td>
<td>97.6</td>
</tr>
<tr>
<td>Hellpern 2008</td>
<td>BHR</td>
<td>113</td>
<td>58</td>
<td>54</td>
<td>88/4/8</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>96.3 (92.8-99.8)</td>
</tr>
<tr>
<td>Hing 2007</td>
<td>BHR</td>
<td>230</td>
<td>66</td>
<td>52</td>
<td>NS</td>
<td>5</td>
<td>0.5</td>
<td>2</td>
<td>97.8 (97.1-100)</td>
</tr>
<tr>
<td>Jameson 2010</td>
<td>ASR</td>
<td>214</td>
<td>60</td>
<td>56</td>
<td>68/5/27</td>
<td>3.6</td>
<td>0</td>
<td>12</td>
<td>93 (80-98)</td>
</tr>
<tr>
<td>Study design</td>
<td>Implant</td>
<td>Hips</td>
<td>% Male</td>
<td>Mean age</td>
<td>% prim OA / Sec OA / Other</td>
<td>Mean FU (yrs)</td>
<td>% LTFU</td>
<td>Revisions (n)</td>
<td>Survival% (95% CI)</td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>------</td>
<td>--------</td>
<td>----------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Khan 2009</td>
<td>case series prosp</td>
<td>BHR 679</td>
<td>60</td>
<td>51</td>
<td>NS</td>
<td>6</td>
<td>2</td>
<td>29</td>
<td>95.7 (94.4-97.4)</td>
</tr>
<tr>
<td>Killampalli 2009</td>
<td>case series prosp</td>
<td>Cormet 100</td>
<td>61</td>
<td>56</td>
<td>97/1/2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Kim 2008</td>
<td>case series prosp</td>
<td>Conserve plus 200</td>
<td>78</td>
<td>49</td>
<td>86/4/10</td>
<td>2.6</td>
<td>0</td>
<td>14</td>
<td>93</td>
</tr>
<tr>
<td>Klein 2008</td>
<td>case series prosp</td>
<td>ASR 115</td>
<td>47</td>
<td>58</td>
<td>100/0/0</td>
<td>1</td>
<td>0</td>
<td>13</td>
<td>88.7</td>
</tr>
<tr>
<td>Le 2010</td>
<td>case series prosp</td>
<td>Durom 90</td>
<td>52</td>
<td>47</td>
<td>50/0/50</td>
<td>2.3</td>
<td>0</td>
<td>1</td>
<td>98.9</td>
</tr>
<tr>
<td>Madhu 2010</td>
<td>case series prosp</td>
<td>BHR 117</td>
<td>58</td>
<td>54</td>
<td>56/44/0</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>91.5 (85.4-97.6)</td>
</tr>
<tr>
<td>Marulanda 2008</td>
<td>case series prosp</td>
<td>BHR 230</td>
<td>73</td>
<td>55</td>
<td>87/0/13</td>
<td>1.3</td>
<td>NS</td>
<td>3</td>
<td>98.7</td>
</tr>
<tr>
<td>Mc Andrew 2007</td>
<td>case series prosp</td>
<td>BHR 180</td>
<td>NS</td>
<td>56</td>
<td>94/0/6</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>98.3</td>
</tr>
<tr>
<td>McBryde 2010</td>
<td>case series prosp</td>
<td>BHR 2123</td>
<td>62</td>
<td>55</td>
<td>100/0/0</td>
<td>3.5</td>
<td>10</td>
<td>48 97.5 (96.3-98.3)</td>
<td></td>
</tr>
<tr>
<td>Mont 2007</td>
<td>case series prosp</td>
<td>Conserve plus 1016</td>
<td>28</td>
<td>50</td>
<td>77/11/12</td>
<td>2.8</td>
<td>6.3</td>
<td>54</td>
<td>94.2 (90.0-96.7)</td>
</tr>
<tr>
<td>Ollivere 2009</td>
<td>case series prosp</td>
<td>BHR 463</td>
<td>66</td>
<td>56</td>
<td>NS</td>
<td>5</td>
<td>0.6</td>
<td>13</td>
<td>95.8 (94.1-96.8)</td>
</tr>
<tr>
<td>Siebel 2006</td>
<td>case series prosp</td>
<td>ASR 300</td>
<td>64</td>
<td>57</td>
<td>NS</td>
<td>0.6</td>
<td>0</td>
<td>8</td>
<td>97.2</td>
</tr>
<tr>
<td>Study design</td>
<td>Implant</td>
<td>Hips</td>
<td>% Male</td>
<td>Mean age</td>
<td>% prim OA/Sec OA/Other</td>
<td>Mean FU (yrs)</td>
<td>% LTFU</td>
<td>Revisions (n)</td>
<td>Survival% (95% CI)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>------</td>
<td>--------</td>
<td>----------</td>
<td>------------------------</td>
<td>---------------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Steffen 2008</td>
<td>BHR</td>
<td>610</td>
<td>59</td>
<td>52</td>
<td>8/9/6</td>
<td>7</td>
<td>0.33</td>
<td>23</td>
<td>95 (85.3-99.2)</td>
</tr>
<tr>
<td>Stulberg 2008</td>
<td>Cormet</td>
<td>337</td>
<td>68</td>
<td>50</td>
<td>86/0/14</td>
<td>2</td>
<td>8.3</td>
<td>24</td>
<td>92.9</td>
</tr>
<tr>
<td>Swank 2009</td>
<td>Durom</td>
<td>128</td>
<td>62</td>
<td>51</td>
<td>71/5/24</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>98.1</td>
</tr>
<tr>
<td>Vendittoli 2010</td>
<td>RCT</td>
<td>109</td>
<td>63</td>
<td>49</td>
<td>31/9/60</td>
<td>4.7</td>
<td>0</td>
<td>4</td>
<td>96.3</td>
</tr>
<tr>
<td>Witzleb 2008</td>
<td>BHR</td>
<td>300</td>
<td>57</td>
<td>49</td>
<td>19/63/18</td>
<td>2</td>
<td>0.7</td>
<td>6</td>
<td>98</td>
</tr>
<tr>
<td>Study</td>
<td>Hips</td>
<td>Implant</td>
<td>Clinical scores</td>
<td>Pre op (indexed)</td>
<td>Post op (indexed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amstutz 2010</td>
<td>1107</td>
<td>Conserve plus</td>
<td>UCLA pain</td>
<td>35</td>
<td>92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA walking</td>
<td>60</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA function</td>
<td>54</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA activity</td>
<td>43</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aulakh 2009</td>
<td>202</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>62 (ON group)</td>
<td>96 (ON group)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Harris Hip Score</td>
<td>58 (OA group)</td>
<td>95.8 (OA group)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beaulé 2009</td>
<td>116</td>
<td>Conserve plus</td>
<td>Harris Hip Score</td>
<td>53.1</td>
<td>90.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bergeron 2009</td>
<td>228</td>
<td>ASR</td>
<td>Harris Hip Score</td>
<td>46</td>
<td>91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>44</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daniel 2004</td>
<td>446</td>
<td>McMinn and BHR</td>
<td>Modified UCLA</td>
<td>not presented</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oxford hip score</td>
<td>not presented</td>
<td>77.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daniel 2010</td>
<td>184</td>
<td>McMinn second gen</td>
<td>Oxford hip score</td>
<td>not presented</td>
<td>68.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DeSmet 2005</td>
<td>252</td>
<td>BHR, Conserve plus, Durom</td>
<td>Harris hip score</td>
<td>not presented</td>
<td>97.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giannini 2007</td>
<td>350</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>57</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gravius 2009</td>
<td>82</td>
<td>Durom</td>
<td>Harris Hip Score</td>
<td>40.1</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>4.6</td>
<td>8.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Merle d'Aubigne-Postel</td>
<td>1.9</td>
<td>17.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Hips</td>
<td>Implant</td>
<td>Clinical scores</td>
<td>Pre op (indexed)</td>
<td>Post op (indexed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>---------</td>
<td>--------------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hellpern 2008</td>
<td>113</td>
<td>BHR</td>
<td>Oxford hip score</td>
<td>30.2</td>
<td>74.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Harris hip score</td>
<td>not presented</td>
<td>96.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>39.1</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hing 2007</td>
<td>230</td>
<td>BHR</td>
<td>Oxford hip score</td>
<td>not presented</td>
<td>76.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Harris hip score</td>
<td>62.2</td>
<td>95.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jameson 2008</td>
<td>254</td>
<td>ASR</td>
<td>Harris Hip Score</td>
<td>51.3</td>
<td>94.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>39</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khan 2009</td>
<td>679</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>47</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Killampalli 2009</td>
<td>100</td>
<td>Cormet</td>
<td>Oxford hip score*</td>
<td>37.5</td>
<td>89.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>30</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim 2008</td>
<td>200</td>
<td>Conserve plus</td>
<td>Harris hip score</td>
<td>55.8</td>
<td>92.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WOMAC pain</td>
<td>49</td>
<td>92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WOMAC stiffness</td>
<td>41.2</td>
<td>80.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WOMAC Function</td>
<td>46.2</td>
<td>89.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klein 2008</td>
<td>115</td>
<td>ASR</td>
<td>Harris hip score</td>
<td>59</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lei 2010</td>
<td>90</td>
<td>Durom</td>
<td>Harris hip score</td>
<td>57</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Madhu 2010</td>
<td>2010</td>
<td>BHR</td>
<td>Harris hip score</td>
<td>not presented</td>
<td>84.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oxford Hip Score</td>
<td>no: presented</td>
<td>64.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Hips</td>
<td>Implant</td>
<td>Clinical scores</td>
<td>Pre op (indexed)</td>
<td>Post op (indexed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>-------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marulanda 2008</td>
<td>230</td>
<td>BHR</td>
<td>Oxford Hip score**</td>
<td>27</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McAndrew 2007</td>
<td>180</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>44</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McBryde 2008</td>
<td>909</td>
<td>BHR</td>
<td>Oxford Hip score</td>
<td>34.5</td>
<td>95.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mont 2007</td>
<td>1015</td>
<td>BHR</td>
<td>Harris Hip score</td>
<td>not presented</td>
<td>93.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ollivere 2009</td>
<td>463</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>not presented</td>
<td>not presented</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siebel 2006</td>
<td>300</td>
<td>ASR</td>
<td>Harris Hip score</td>
<td>44</td>
<td>89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steffen 2008</td>
<td>610</td>
<td>BHR</td>
<td>Harris Hip score</td>
<td>not presented</td>
<td>93.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Oxford Hip score</td>
<td>not presented</td>
<td>73.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>UCLA</td>
<td>not presented</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stulberg 2008</td>
<td>337</td>
<td>Comet</td>
<td>Harris Hip score</td>
<td>50.1</td>
<td>96.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swank 2009</td>
<td>128</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>49</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vendittoli 2010</td>
<td>109</td>
<td>Durom</td>
<td>UCLA</td>
<td>not presented</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WOMAC</td>
<td>45.1</td>
<td>90.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Merle d’Aubigne-Postel</td>
<td>60</td>
<td>97.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Witzleb 2008</td>
<td>300</td>
<td>BHR</td>
<td>Harris Hip Score</td>
<td>51</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>Hips (n)</td>
<td>Revised (n)</td>
<td>FFN* (n)</td>
<td>Asept L** (n)</td>
<td>Infections (n)</td>
<td>ALVAL (n)</td>
<td>Disl# (n)</td>
<td>Malp## (n)</td>
<td>Pers Pain### (n)</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>---------------</td>
<td>----------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Amstutz</td>
<td>1107</td>
<td>44</td>
<td>9</td>
<td>26</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aulakh</td>
<td>202</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Beaulé</td>
<td>116</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bergeron</td>
<td>228</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Daniel</td>
<td>446</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Daniel</td>
<td>184</td>
<td>30</td>
<td>0</td>
<td>29</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DeSmet</td>
<td>252</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Giannini</td>
<td>350</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gravius</td>
<td>82</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Heilpern</td>
<td>113</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hing</td>
<td>230</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*FFN indicates Femoral Fracture of the Neck; **Asept L indicates Aseptic Loosening; #Disl indicates Dislocation; ##Malp indicates malpositioning; *# Pers Pain indicates Persistent Pain; **# AVN indicates Avascular Necrosis of the femoral head.
Table 4.3, Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Hips (n)</th>
<th>Revised (n)</th>
<th>FFN* (n)</th>
<th>Asept L** (n)</th>
<th>Infections (n)</th>
<th>ALVAL (n)</th>
<th>Disl# (n)</th>
<th>Malp## (n)</th>
<th>Pers Pain*# (n)</th>
<th>AVN### (n)</th>
<th>Other (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jameson 2010</td>
<td>214</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Khan 2009</td>
<td>679</td>
<td>29</td>
<td>11</td>
<td>14</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Killampalli 2009</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kim 2008</td>
<td>200</td>
<td>14</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Klein 2008</td>
<td>115</td>
<td>13</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lei 2010</td>
<td>90</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Madhu 2010</td>
<td>117</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Marulanda 2008</td>
<td>230</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>McAndrew 2007</td>
<td>180</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mont 2007</td>
<td>1016</td>
<td>54</td>
<td>27</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Olivere 2009</td>
<td>463</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.3, Revision details. *FFN indicates Femoral Fracture of the Neck; **Asept L indicates Aseptic Loosening; #Disl indicates Dislocation; ##Malp indicates malpositioning; *# Pers Pain indicates Persistent Pain; ###AVN indicates Avascular Necrosis of the femoral head
### Table 4.3, Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Hips (n)</th>
<th>Revised (n)</th>
<th>FFN* (n)</th>
<th>Asept L** (n)</th>
<th>Infections (n)</th>
<th>ALVAL (n)</th>
<th>Disl# (n)</th>
<th>Malp## (n)</th>
<th>Pers Pain*# (n)</th>
<th>AVN**## (n)</th>
<th>Other (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siebel 2006</td>
<td>300</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Steffen 2008</td>
<td>610</td>
<td>23</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Stulberg 2008</td>
<td>337</td>
<td>24</td>
<td>8</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Swank 2009</td>
<td>128</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Vendittoli 2010</td>
<td>109</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Witzleb 2008</td>
<td>300</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*FFN indicates Femoral Fracture of the Neck; **Asept L indicates Aseptic Loosening; #Disl indicates Dislocation; ##Malp indicates malpositioning; *# Pers Pain indicates Persistent Pain; **# AVN indicates Avascular Necrosis of the femoral head
Discussion

All but one of the implants studied had insufficient follow-up to be compliant with the NICE benchmark, of a revision rate of less than 10% at ten years, for choosing a prosthesis for primary THR. The study reporting a follow-up of longer than ten years had a revision rate of 16%, mainly for aseptic loosening of the implant. This high failure rate was attributed to the double-heat-treatment manufacturing process which is no longer in use. The prosthesis was superseded by the Cormet 2000 implant in 1996. Compared with the three-year NICE entry-benchmark of implant survival ≥ 97%, 13 studies (44.8%) showed satisfactory survival. Eight used the BHR implant, two the Conserve plus, one the Durom, one the Cormet 2000 and one both the McMinn and BHR implants. There appeared to be a difference in the performance of the implants, with only the ASR appearing below the benchmark in four studies, in two of which the follow-up was very short. The BHR appeared above the line in 12 of 13 studies. Both the Conserve plus (four studies) and the Cormet implant (two studies) had an equal number above and below the line. The Durom implant had three studies above and one below the line. Since we excluded studies with incomplete learning curves, these results are more likely to be attributed to the characteristics of the implant such as the design and manufacturing process (Figure 4.2). No survival data were analysed for six of 11 HRA devices on the market (Accis, Adept, Eska, Icon, Mitch and Recap). However, the implants in the studies which were included represented most of the HRA implants worldwide. Aseptic loosening was the most frequent cause of failure (1.4%), followed by fracture of the femoral neck (1.1%). The variation in frequency of fracture of the femoral neck among studies was large, between 0% and 16.3% compared with the frequency of aseptic loosening (0.0% to 5.5%). Clinical outcome scores were reported very heterogeneously, but all studies showed a significant improvement from the pre-operative score. In all of the included studies on HRA the patients were relatively young with the mean age ranging between 42 and 58 years. This is important when comparing the failure rates of HRA with those of conventional THR, since most patients within this age range will be considerably more active than those aged more than 60 years. National Joint Registries are useful for this comparison since they combine different types of conventional THR and publish their data stratified by age. Our pooled revision rate of 3.5% is higher than the revision
rates at three years for conventional THRs in patients aged under 55 years reported in the sixth annual report from the National Joint Registry for England and Wales\textsuperscript{57}, but lower than that for HRA of 4.5% (95% CI 3.9 to 5.3).\textsuperscript{57} Both the cumulative revision rate at eight years for THR and for HRA in patients aged less than 55 years in The Australian National Joint Replacement Registry Report (4% and 4.7%, respectively) are higher than our pooled revision rate.\textsuperscript{61} As with the results for HRA presented in our review, there is much variation in reported survival rates for THR, ranging from less than 80% to 99% at ten years for patients aged less than 55 years.\textsuperscript{24,58} However, the pattern of the modes of failure is different. For HRA, fracture of the femoral neck is a unique mode of failure, but according to our results aseptic loosening occurs slightly more frequently, 1.1% versus 1.4% respectively. According to the Finnish Arthroplasty Registry, in patients under 55 years of age with primary osteoarthritis, the survival rate at ten years of less than 80% for THR is mainly due to wear of the liner. This required only its exchange at revision surgery.\textsuperscript{62} Excessive wear in HRA leads to increased levels of metal ions in the blood with ALVAL as a possible serious consequence. This is not unique to HRA. In a randomised trial, Garbuz et al\textsuperscript{63} reported a 46-fold increase in metal ion concentrations in the blood in patients with large-diameter Metal-on-Metal THRs, compared with a tenfold increase in patients who had undergone HRA.\textsuperscript{63} Failure due to dislocation, a common cause of revision of THR, is rare in HRA. In our review, only 27 of 10 621 HRAs were revised for dislocation (0.3%). Arguably, the same low rate of dislocation could be achieved using a large-headed THR, but without the perceived benefit of retained femoral bone stock. However, the use of bearings of large diameter appears to be effective against the risk of dislocation. In 2002 two systematic reviews of the literature up to 2001 were published, both presenting the results of one literature search.\textsuperscript{64,65} Based on the publication date, our systematic review included 29 new studies. Compared with several recent reviews on HRA, our study was designed as a systematic review focusing on survival of the implant compared with the NICE benchmark.\textsuperscript{18,19} The strengths of our review included a comprehensive and reproducible search strategy, exclusion of duplicate case series and studies with an incomplete learning curve, contact with authors for clarification, a comparison with an objective revision rate benchmark and the use of the GRADE system for assessment of quality. Finally, survival of the implant defined as years to revision surgery was an objective and patient-relevant endpoint. Our study is limited since
28 of the 29 were case series, possibly introducing bias. A requirement of a minimum of 75 treated patients resulted in the exclusion of studies describing metal ion release. These are expensive to perform and accordingly restrict the number of patients studied. The few which included this information and had more than 75 patients were excluded since they did not present data on survival of the implant. The quality of the included studies was very low according to the GRADE system, which is mainly based on the design of the study and heterogeneous reporting of clinical scores and radiological findings. Based on our findings there remain concerns on the long term effectiveness and safety and longer follow-up is needed. The large variation in the incidence of fracture of the femoral neck as a mode of failure in studies is poorly understood. However, the most frequent mode of failure after HRA remains aseptic loosening.
References


CHARNLEY HIP PROSTHESIS
LOW FRICTION
POLYETHYLENE

1960
Chapter 5

Hip resurfacing in a district general hospital: 6-year clinical results using the ReCap hip resurfacing system

Walter van der Weegen, Henk J. Hoekstra, Thea Sijbesma, Shennah Austen and Rudolf W. Poolman
BMC Musculoskelet Disord. 2012 Dec 13;13:247
Abstract

Background
The purpose of our study was to prospectively report the clinical results of 280 consecutive hips (240 patients) who received a ReCap Hip Resurfacing System implant (Biomet Inc., Warsaw, USA) in a single district general hospital. Literature reports a large variation in clinical results between different resurfacing designs and published results using this particular design are scarce.

Methods
Mean follow up was 3.3 years (1.0 to 6.3) and four patients were lost to follow-up. All patients were diagnosed with end-stage hip osteoarthritis, their mean age was 54 years and 76.4% of all patients were male.

Results
There were 16 revisions and four patients reported a Harris Hip Score <70 points at their latest follow up. There were no pending revisions. Kaplan-Meier implant survival probability, with revision for any reason as endpoint, was 93.5% at six years follow-up (95%-CI: 88.8-95.3). There were no revisions for Adverse Reactions to Metal Debris (ARMD) and no indications of ARMD in symptomatic non-revised patients, although diagnostics were limited to ultrasound scans.

Conclusions
This independent series confirms that hip resurfacing is a demanding procedure, and that implant survival of the ReCap hip resurfacing system is on a critical level in our series. In non-revised patients, reported outcomes are generally excellent.
Background

Hip resurfacing arthroplasty (HRA) has been widely used in recent years. Possible advantages of conserved femoral bone stock, low wear rates and low dislocation rates were the main reasons for surgeons to use HRA. Recent concerns on the use of Metal-on-Metal (MoM) bearings have intensified the discussion on HRA. The reported increase of metal ion levels after HRA with subsequent local Adverse Reactions to Metal Debris (ARMD) and poor results with revision for this complication have diminished the support for HRA. In the published literature there is a wide range of clinical results between different HRA designs. Although numeral clinical studies report short- and mid-term survival of different HRA systems, these studies focus on a limited number of HRA designs. To our knowledge, there are four studies published using the ReCap Hip Resurfacing System (Biomet Inc., Warsaw, USA). Gagala reported there were no significant complications after a maximum follow up (FU) of 20 months, using this implant design (n=23). Baad-Hansen reported no significant translation or rotation using this implant design (n=25), after two year FU using RadioStereometry Analysis (RSA). A larger number of ReCap procedures (n=137) with a three year FU are described in the Australian National Joint Replacement Registry. In this report a cumulative percent revision rate of 7.6% is presented for this specific HRA design. Recently, Gross and Liu presented the mid-term results of 740 hip resurfacings with a 3.4% revision rate. In this prospective study, we report the clinical results of 280 consecutive HRA’s using the ReCap Hip Resurfacing system, with a maximum FU of six years (range: 1 to 6). We hypothesised that implant survival would be compliant with the National Institute for Health and Clinical Excellence (NICE) benchmark (a revision rate of 10% or less at ten years, or consistent survival if only shorter FU is available). We further hypothesised that the risk for revision in subgroups based on gender, age and component size is comparable to findings in published literature.

Methods

Patients

Between September 2004 and September 2010 our first 280 consecutive, non-selected HRA procedures (240 patients) in a general district hospital were included in a prospective cohort study (Table 5.1). Patients diagnosed with end stage osteoarthritis (OA) were indicated for HRA. The entire group involved 240
patients (280 resurfacings) with a mean follow-up of 3.3 years (1 to 6.3) of whom 45 were followed-up for five years and 30 for six years. Prior to surgery, a dual energy X-ray absorptiometry (DEXA) scan was made of all female patients and in all male patients suspected of osteoporosis. When T and Z values were below normal, patients were excluded from HRA. After informing the patient on the

<table>
<thead>
<tr>
<th>Table 5.1, Demographics of the Study Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td>Age at surgery (yr)</td>
</tr>
<tr>
<td>BMI</td>
</tr>
<tr>
<td>Hospital stay (days)</td>
</tr>
<tr>
<td>Follow up (months)</td>
</tr>
<tr>
<td>Sex (n=240 patients)</td>
</tr>
<tr>
<td>Males</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>Diagnosis (n=280 hips)</td>
</tr>
<tr>
<td>Primary OA*</td>
</tr>
<tr>
<td>DDH**</td>
</tr>
<tr>
<td>Posttraumatic OA</td>
</tr>
</tbody>
</table>

*OA indicates osteoarthritis; **DDH, developmental dysplasia of the hip

expected benefits and risks associated with HRA, informed consent on the surgery procedure and on study participations was obtained from all patients. Our study was approved by the Institutional Review Board. Patients with renal failure, femoral cysts, osteoporosis or a-vascular necrosis (AVN) of the femoral head were excluded. Female patients with a possible child wish were also excluded.

Surgical technique and rehabilitation
Two experienced joint arthroplasty surgeons (HJH,TS) used the ReCap Hip Resurfacing System (Biomet Inc., Warsaw, USA) in all patients in a standard manner. Prophylactic antibiotics were administered on induction. Both the press-fit acetabular component and the cemented femoral component are manufactured from “as-cast” cobalt chrome (Co-Cr-Mo) with a high carbon content (>0.2%). The acetabular outside is a full-hemisphere design and has four
pairs of fins for initial rotational stability. It has a titanium porous plasma spray surface coating (Figure 5.1). The outer geometry of the cemented femoral component extends approximately 23 degrees beyond a full-hemisphere. The critical inner bearing surface has a coverage arc ranging from 155–164 degrees from smallest to largest component. The posterolateral approach was used in all procedures. After dislocating the hip joint, acetabular osteophytes were removed, the acetabulum was reamed and the acetabular component was impacted into the anatomical position.

![Figure 5.1, ReCap hip resurfacing system.](image)

Next, a femoral guide wire was inserted into the femoral head, directed with a jig. The femoral head was then circumferentially reamed and the bone-bed was prepared with drill holes and pulse lavage for cementing. After applying high viscosity cement (Refobacin® Bone Cement R, Biomet Europe, Dordrecht, Netherlands) to the inner surface, the femoral component was carefully put in place. Patients were mobilised the first post-operative day using two crutches and weight bearing as tolerated. All patients received antibiotic prophylaxis with a cephalosporin preoperatively and 24 hours post-operatively, fourteen days of indometacin for periarticular ossification prophylaxis, diclophenac for pain management and thrombosis prophylaxis with dalteparine 5000 units for six weeks postoperatively. Patients were discharged if the patient was fully mobile and the wound was without problems. Physiotherapy was prescribed to all patients. Patients were instructed to avoid all high impact activities in the first six months and discouraged to participate in high impact sports. All bilateral procedures were staged interventions with at least a three months interval.
Study protocol

Patients were recruited at the time of surgery and prospectively followed six weeks after surgery and yearly thereafter. Bilateral cases were followed up as separate cases. Standard antero-posterior (AP) and lateral radiographs, and the Harris Hip Score\textsuperscript{12} were collected at each visit, except for the six week FU. Only radiographs were collected at this visit. Any patient who was symptomatic post-operatively was analysed with a diagnostic ultrasound scan to check for ARMD. On the AP radiograph, the acetabular angle of inclination and femoral stem shaft angle were measured as described by Beaulé et al.\textsuperscript{13} Radiolucencies were measured in millimeters and acetabular radiolucency was classified in three zones according to DeLee and Charnley (Figure 5.2).\textsuperscript{14} Any femoral radiolucencies were classified in the three zones as described by Beaulé et al. (Figure 5.2).\textsuperscript{13} Heterotopic bone formation was classified as described by Brooker et al.\textsuperscript{15} Neck narrowing was measured as described by Grammatopoulos et al., using the first post-operative radiograph and the most recent radiograph for comparison.\textsuperscript{16} Clinical and radiological FU and statistical analyses were done by an independent observer, with a sample set of radiographic measurements audited by an experienced radiologist.

\textbf{Figure 5.2,} Acetabular zones according to DeLee & Charnley and femoral resurfacing zones according to Beaulé.
Statistical analysis
Revision for any reason was the primary endpoint of this study. Kaplan-Meier survivorship curves were calculated. Since we support the recent notion in literature that implant survivorship is a limited endpoint to define a successful outcome for joint arthroplasty, a HHS score of <70 points on the latest FU (two years or more) was also used as an endpoint for implant failure. The NICE benchmark (a revision rate of 10% or less at ten years, or consistent survival if only shorter FU is available) was used to evaluate survivorship. Relative risks (RR) were calculated to evaluate sub-group results based on gender, age, component size and acetabular inclination angle. A femoral head size <50 mm and an acetabular inclination angle of ≥ 550 were considered to be a risk factor for ARMD and therefore revision. SPSS software (SPSS Statistics, version 17.0, IBM Corporation, Somers USA) was used for all statistical analyses. The occurrence of femoral neck narrowing as a consequence to head downsizing can also be indicative for ARMD, as described by Grammatopoulos et al. Neck narrowing values were calculated as a percentage and ranges were presented for the whole cohort and for the patients who were revised > six months after index surgery.

Results
Four patients were deceased for reasons not related to the HRA procedure (four prostheses, 1.4%) and no other patient was lost to FU. Three patients were contacted by phone since they were unable to return for FU. Therefore, radiological FU was complete for 277 patients. There were 16 revisions at the time of final FU. Seven were for fracture of the femoral neck, five for aseptic loosening of the acetabular component, two for component malpositioning (one femoral and one acetabular) and two for persistent pain (Table 5.2). The Kaplan-Meier implant survival probability with revision for any reason as endpoint was 93.5% at six years FU (95%-CI: 88.8-95.3) (Figure 5.3). The mean time to revision was 14 months (range: 0 to 56) with eight out of 16 revisions within two months from index surgery. Female patients had a RR for revision of 1.1 compared to male patients (95%-CI: 0.92-1.06). The RR for revision in the group of patients with a femoral head <50 mm, was 1.1 compared to the group of patients with larger components (95%-CI: 0.98-1.09). In the patients younger 55 years the RR for revision was 0.9 compared to patients 55 years or older (95%-CI: 0.95-1.07).
### Table 5.2, Revision details

<table>
<thead>
<tr>
<th>Failure mode</th>
<th>Gender</th>
<th>Age</th>
<th>Fem.comp.</th>
<th>Months to revision</th>
<th>Revision details</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN#*</td>
<td>Male</td>
<td>61</td>
<td>48mm</td>
<td>0.5</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Female</td>
<td>55</td>
<td>46mm</td>
<td>0.5</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Male</td>
<td>57</td>
<td>52mm</td>
<td>1</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Male</td>
<td>57</td>
<td>50mm</td>
<td>1</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Male</td>
<td>60</td>
<td>52mm</td>
<td>1</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Male</td>
<td>54</td>
<td>50mm</td>
<td>2</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>FN#</td>
<td>Male</td>
<td>48</td>
<td>48mm</td>
<td>18</td>
<td>Femoral revision</td>
</tr>
<tr>
<td>Mal Fem Comp**</td>
<td>Male</td>
<td>60</td>
<td>52mm</td>
<td>0</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Mal Acet Comp#</td>
<td>Male</td>
<td>67</td>
<td>50mm</td>
<td>12</td>
<td>THP other hospital</td>
</tr>
<tr>
<td>Asep Loosening##</td>
<td>Male</td>
<td>58</td>
<td>54mm</td>
<td>1</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Asep Loosening</td>
<td>Male</td>
<td>64</td>
<td>50mm</td>
<td>23</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Asep Loosening</td>
<td>Female</td>
<td>49</td>
<td>44mm</td>
<td>32</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Asep Loosening</td>
<td>Male</td>
<td>28</td>
<td>50mm</td>
<td>43</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Asep Loosening</td>
<td>Female</td>
<td>49</td>
<td>42mm</td>
<td>56</td>
<td>Both comp. revised</td>
</tr>
<tr>
<td>Persistent pain</td>
<td>Male</td>
<td>43</td>
<td>50mm</td>
<td>7</td>
<td>THP other hospital</td>
</tr>
<tr>
<td>Persistent pain</td>
<td>Female</td>
<td>52</td>
<td>50mm</td>
<td>27</td>
<td>Both comp. revised</td>
</tr>
</tbody>
</table>

*FN# indicates fracture of the femoral neck; **Mal Fem Comp: malpositioned femoral component; #Mal Acet Com: malpositioned acetabular component; ##Asep Loosening: Aseptic loosening; Fem comp: femoral component size;

---

**Figure 5.3,** Kaplan-Meier survival probability.
Revision details

In all seven femoral neck fracture cases, the acetabular shell was left in situ and a stemmed, uncemented femoral prosthesis was inserted. Six out of seven neck fractures occurred within two months of the index surgery, one case was a late neck fracture 18 months post-operatively. During revision surgery of this one case it was observed that the femoral component was loose, which was thought to be caused by avascular necrosis of the femoral head. In all other cases both components were replaced. All cases of aseptic loosening only involved the uncemented acetabular component. Of the none-revised patients, there were four patients with a HHS score <70 points at their latest FU (two at two years and two at three years FU). Revision and clinical score combined as endpoint for implant failure, resulted in 20 failed prostheses at the time of final FU. During revision surgery no metallosis, soft tissue cysts or solid masses were observed, although postoperative histopathological analyses showed chronic inflammatory signs including synovial hyperplasia en some metallosis in both patients revised for persistent pain, indicating adverse local tissue reaction to metal debris. A diagnostic ultrasound was made in 27 patients (9.6%) with unexplained hip or groin pains, all were normal. In our series there were 81 patients with an acetabular inclination angle of 55°-65° (of which 23 had a femoral head size <50 mm) and 10 patients with an acetabular inclination angle of >65° (of which four had a femoral head size <50 mm). In none of these patients any signs of ARMD were observed during any revision surgery or additional diagnostic ultrasound scans.

Complications without need for revision

There were 30 (10.7%) complications without need for revision (Table 5.3). The majority of these complications were transient such as post-operative bleeding (n=18). There was one deep wound infection which was eradicated after surgical debridement and antibiotic treatment. Seven other patients with signs of a post-operative wound infection were treated successfully with antibiotics. There was one patient with persistent paraesthesia and pareses of the foot due to a sciatic nerve lesion. One other patient had a transient nerve palsy of the sciatic nerve. Another patient was treated conservatively for a non-displaced fracture of the femoral neck, which he sustained due to a fall three months after surgery. He recovered without any persistent symptoms. A healed non-displaced femoral neck
stress fracture was discovered with routine FU two years post-operatively (Figure 5.4A and 5.4B). This patient had experienced some groin pain after running, which completely resolved when he did not run for a couple of weeks. There were no dislocations or thromboembolic events in our series.

<table>
<thead>
<tr>
<th>Table 5.3, Complications without need for revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complication</td>
</tr>
<tr>
<td>Nerve damage</td>
</tr>
<tr>
<td>Non-displaced femoral neck fracture</td>
</tr>
<tr>
<td>Deep wound infection</td>
</tr>
<tr>
<td>Superficial wound infection</td>
</tr>
<tr>
<td>Post-operative bleeding</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
Outcomes

At one year FU, mean HHS had improved significantly from pre-operative scores (from 49.3 to 92, \( p < 0.0001 \), Table 5). At six year FU, 36 patients had an “excellent” HHS (66.7%), 16 a “good” HHS (29.6%) and two a “fair” HHS (3.7%). For the revised patients, the mean HHS after revision was 77 (range: 41 to 91).

<table>
<thead>
<tr>
<th></th>
<th>HHS</th>
<th>Fem. Pos.</th>
<th>Cup abd. Angle</th>
<th>Brooker 1/2/3/4 (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre op (n=280)</td>
<td>49.3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>6 wks (n=280)</td>
<td>-</td>
<td>+2.2°</td>
<td>51.3°</td>
<td>13/2/0/0</td>
</tr>
<tr>
<td>1 yr (n=280)</td>
<td>92</td>
<td>-</td>
<td>-</td>
<td>29/7/2/0</td>
</tr>
<tr>
<td>2 yrs (n=221)</td>
<td>88.3</td>
<td>-</td>
<td>-</td>
<td>26/5/3/0</td>
</tr>
<tr>
<td>6 yrs (n=54)</td>
<td>89.3</td>
<td>-</td>
<td>-</td>
<td>11/3/1/0</td>
</tr>
</tbody>
</table>

Radiological findings

At one year FU, the mean implant femoral shaft angle was 135.10 (range: 1160 to 1560). Mean acetabular angle of inclination was 51.30 (range: 260 to 770). With further FU, no radiolucencies were observed. Ectopic bone formation was noted in 13.8% of all cases. Mean HHS for patients who had a Brooker grade two or three ectopic bone formation was 91 points (range: 74 to 91) (Table 5.4). Neck narrowing was observed in 136 patients with a mean of 2.3% (range: 0% to 18.5%). In the patients with revisions later than 6 months after index surgery, neck narrowing was present in 3 out of 9 patients. One patient had 2.5% neck narrowing and two patients had 6% neck narrowing.

Discussion

Our KM-survival probability of 93.5% at six years FU (95%-CI: 88.8-95.3) is not compliant with the three year entry NICE benchmark. Longer FU is needed to compare our results with the full 10-year benchmark. Of the non-revised patients, there were only four patients with implant failure based on their HHS score. The combined endpoints of revision (n=16) and HHS score <70 points (n=4), resulted in 20 failed prosthesis (7.1%). Since no other studies on MoM hip resurfacing have combined implant survival and Patient Reported Outcome scores to define
implant performance, we cannot compare this result to other studies. We were able to identify all failure modes, including those from patients revised in other hospitals. Most frequent reasons for revision were fracture of the femoral neck (n=7) and aseptic loosening (n=5). All cases of aseptic loosening occurred relatively early and involved only the uncemented acetabular component. We think that insufficient seating of the acetabular component, which might occur due to deformation of the relatively thin cup during the impaction procedure, may have caused these early revision cases. In our series we have not observed any signs of ARMD during revision surgery, although post revision surgery two patients revised for persistent pain had histopathological evidence of adverse local tissue reaction (ALTR) to metal debris. Neither have we observed any signs of ARMD with diagnostic ultrasound scans in patients who were post-operatively symptomatic. We cannot completely rule out the presence of ARMD in our series, but since we observed two cases of ALTR, future follow-up will include routine metal ion analysis. Our complete FU, our detailed information on revision cases and the excellent clinical scores at the time of final FU are in contrast to other designs of HRA, of which failure rates of 25% for ARMD after six years FU are reported. Risk factors for ARMD are the inclination angle of the acetabular cup, implant design, small component sizes and occurrence of neck narrowing. Steep inclination angles and an acetabular cup with less than hemispherical coverage result in a small contact patch area (CPA), which increases the wear rate. Another risk factor is component size, with small sizes resulting in more friction, releasing more metal debris. In our series there were 81 patients with such risk factors, but no ARMD was observed in any of these patients, neither with a diagnostic ultrasound scan nor during revision surgery. The critical inner bearing surface of the ReCap has a coverage arc ranging from 155–164 degrees from smallest to largest component which is similar to other designs with a larger CPA such as the Birmingham Hip Resurfacing design (Smith and Nephew PLC, London, UK), the Conserve plus (Wright Medical Technology, Inc., Arlington, USA) and the Cormet resurfacing design (Corin Group PLC, Cirencester, UK). Our findings on ARMD are in line with several other studies. Malviya found a 0.15% incidence of pseudotumors using the Birmingham Hip Resurfacing (BHR). Beaulé et al found a 0.1% prevalence of pseudotumors with MoM resurfacing after surveying nine Canadian Academic centers. Glyn-Jones et al extensively studied the risk factors for pseudotumor formation in a large series of hip resurfacings. Gender and age
had a significant independent influence on the revision rate for pseudotumor formation, and the incidence increased with time, with a mean time to pseudotumor revision of 3.5 years (1 to 8.3 years). In the series presented by Steffen et al, there were three revision cases possibly related to metal debris. Two of these cases were revised around two years post-operatively, the other one at 5.6 years after surgery. These mean times to pseudotumor revision are within the maximum follow-up time of our case series (6.3 years), but we will have to stay alert on ARMD occurrence with longer follow-up. Grammatopoulos reported a mean 10.1% neck narrowing in patients revised for pseudotumors. In our cohort the mean percentage of neck narrowing was considerably lower (2.3%), although individual cases had greater neck narrowing. We did observe neck narrowing in three out of the nine patients who were revised > six months after index surgery, but these three patients had less than 10% neck narrowing. Neck narrowing data from our cohort is supplementary to the observations by Gross and Liu. They also report <1% revisions for adverse wear and based on their report and on data from our cohort we believe that the risk for adverse wear using this resurfacing design is low. Gross did report a lower revision rate compared to our study (3.4% versus 7.3%) but in his series the learning curve was avoided since the surgeon had performed 400 hip resurfacings before the presented series was started. As noted in the study by Gross, we also now have begun recommending routine metal ion tests in all our patients. Strong points of our study are its prospective study design, a large consecutive study cohort, limited lost to FU and comparison to an objective benchmark. There is detailed FU on all revised patients including those revised in other hospitals, and both clinical outcome scores and radiological FU were analysed. Another advantage is that this study was conducted in a general district hospital rather than a design institution. Our study also has limitations: FU time is limited and there is no control group. We also have to bear in mind that the NICE-benchmark is applied to an OA population of all ages, and literature describes higher revision rates in younger patients. Metal ion levels were not obtained and there were no diagnostic ultrasounds made to check for ARMD in non-symptomatic patients. Compared to published literature, our study reports the clinical results on more patients with longer FU using the ReCap Hip Resurfacing system than any other study. Gagala et al studied 25 patients (mean FU 11 months, range: 10 to 20) and found good short-term clinical results without significant complications. Baad-Hansen et al conducted a radiostereometric
analysis (n=23). There was no statistically significant translation or rotation of the femoral component observed after two years FU.\textsuperscript{8} The absence of any revisions in these series might be due to the small number of patients and the short FU. In the evaluation of risk factors for early failure with HRA, the Australian Arthroplasty Register reported on 137 procedures between 1999 and 2008 using the ReCap hip resurfacing system.\textsuperscript{9} Their cumulative percent revision rate of 7.6\% at three year FU using this system was worse than our implant survival at three years. A possible explanation might be that those 137 procedures were done by a large number of orthopedic surgeons in an extended period of time, limiting the individual expertise using this system. However, despite further enquiry, no more details could be provided by the Australian Arthroplasty Register. Regarding patient selection, in our series the RR for revision was slightly higher for female and for older patients, although statistically the difference was not significant. Patients with smaller component sizes had a higher risk for revision, but this was also not statistically significant. This is in line with several other publications which show a significantly higher risk for revision in female patients, older patients, and in patients with small components.\textsuperscript{5,29-32} The possible absence of ARMD in our series might explain the equal risk for revision in patients with small or large component sizes. Looking at diagnosis, literature reports that the best HRA results are obtained with OA.\textsuperscript{5,33} In our series, only patients with this diagnosis were included.

**Conclusion**

Although implant survival rate in our series is below the NICE benchmark, patient reported outcomes are excellent in the non-revised patients. Also, we were not able to detect signs of ARMD with standard radiographs and clinical outcome scores. As with other resurfacing designs, this resurfacing system should be regarded as a difficult but effective surgical procedure for a small and specific patient population.
References


RING HIP PROSTHESIS
METAL-ON-METAL
1960
Chapter 6

High incidence of pseudotumors after hip resurfacing even in low risk patients; results from intensified MRI screening protocol

Walter van der Weegen, Jose M.H. Smolders, Thea Sijbesma, Henk J. Hoekstra, Koen Brakel and Job L.C. van Susante

Hip Int. 2013; 23(3): 243-9
Abstract
We intensified our screening protocol for the presence of pseudotumors in a consecutive series of patients with a hip resurfacing arthroplasty (HRA), to establish whether we should be alert to the presence of ‘silent’ pseudotumors. Patients categorised with high risk (11 hips) and low risk (10 hips) for pseudotumor development and a control group (23 hips) were screened with Metal-Artefact Reduction Sequence (MARS) magnetic resonance imaging (MRI). The Anderson classification to grade any Metal-on-Metal (MoM) disease present on MARS-MRI images was used. In 15 out of 44 MRI scans pseudotumors were observed (34.1%), of which six were graded with mild (13.6%), eight with moderate (18.2%) and one with severe MoM disease (2.3%). Twelve pseudotumors were present in asymptomatic patients (27.3%). Metal ion levels were normal in 80% of the MARS-MRI screened patients. As a consequence of our intensified screening protocol, one patient was revised for pseudotumor formation and another patient was scheduled for revision. Silent pseudotumors were observed in all three groups. Before our intensified screening protocol was initiated, no pseudotumors were encountered in our cohort of 289 HRAs. We concluded that clinical outcomes and plain radiographs for screening MoM patients underestimates the presence of pseudotumors in MoM patients. The true clinical relevance of these pseudotumors is still unclear.
Introduction

Metal-on-Metal (MoM) bearings have been widely used in hip arthroplasty. Although wear rates are low, these bearings still release cobalt and chromium particles which may result in a periprosthetic soft tissue reaction, requiring revision surgery.\(^1,2\) This periprosthetic soft tissue damage, known as adverse reaction to metal debris (ARMD) compromises aseptic lymphocytic vasculitis-associated lesions (ALVAL), metallosis and pseudotumor formation.\(^3\) Revision surgery for pseudotumors is sometimes difficult and post-revision surgery clinical outcomes are less satisfying.\(^4\) The reported incidence of pseudotumors varies, depending on patient characteristics, type of follow-up and implant design features.\(^5,6\) Earlier MoM hip arthroplasty studies relied on clinical outcome scores and radiographs of large case series to report on good implant performance and excellent functional outcomes.\(^7-9\) Recently published data, however, report on a much higher incidence of pseudotumors in patients with MoM implants after all patients have been screened for the presence of these adverse peri-prosthetic reactions with Metal-Artefact Reduction Sequence (MARS) magnetic resonance imaging (MRI) or ultrasound.\(^10,11\) Suspicion arises that there may be a relatively large number of ‘silent’ pseudotumors present in otherwise well-functioning implants. There is reason to believe that the occurrence of pseudotumors is not solely observed with malpositioned implants with relatively high metal ion levels and poor clinical outcome.\(^11\) From this growing unease we decided to intensify our screening protocol for the presence of pseudotumors in a consecutive series of patients with HRA. The aim of this study was to clarify whether we should be alert to the presence of ‘silent’ pseudotumors in our cohort of hip resurfacing patients. According to previously defined patient and implant characteristics\(^6,11\), we categorised high and low risk patients for pseudotumor development, together with a non-stratified control group. Subsequently, in all three groups MARS-MRI screening for pseudotumors was performed.

Patients and methods

Patients

Between September 2004 and September 2010 we included 298 consecutive HRA procedures (240 patients) in a prospective cohort study. Females <60 years of age and males <65 years of age were the primary candidates for HRA if diagnosed with end stage osteoarthritis (OA) and had an active lifestyle. Older patients with
sufficient bone quality and an active lifestyle were considered for HRA on an individual basis. Dual energy X-ray absorptiometry was used to exclude patients with osteoporosis. Patients with renal failure, femoral cysts, avascular necrosis (AVN) of the femoral head and female patients trying to conceive were also excluded. Procedures were followed in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2000. After informing the patient on the expected benefits and risks associated with HRA, informed consent on the surgery procedure and on study participation was obtained. Our study was approved by the Institutional Review Board (IRB nr. 08.013, 18th December 2008).

**Implant system**

All procedures were performed by one of two experienced hip arthroplasty surgeons (TS, HH). The ReCap hip resurfacing system (Biomet Inc, Warsaw, USA) was implanted by a posterolateral approach. The press-fit acetabular component and the cemented femoral component are manufactured from “as-cast” cobalt-chrome (Co-Cr-Mo) with a high carbon content (>0.2%) without any heat treatment. The acetabular outside is a full hemisphere design and has four pairs of small fins for initial rotational stability. It has a titanium porous plasma spray surface coating facilitating bone ingrowth. The system offers 2 mm increment sizing. The surgical technique has been described earlier by Gross and Liu. All patients received antibiotic prophylaxis with a cephalosporin preoperatively and 24 hours post-operatively, fourteen days of indometacin for periarticular ossification prophylaxis, diclophenac for pain management and thrombosis prophylaxis with dalteparine 5000 units for six weeks postoperatively. Patients were rehabilitated with immediate unrestricted weight bearing according to the patient’s tolerance. All bilateral procedures were staged interventions with at least a three month interval.

**Study design**

To evaluate the occurrence and incidence of pseudotumor formation we defined three different groups of patients. The first group had a perceived high risk for pseudotumor formation based on gender, component size and cup inclination angle. Cup inclination angle was measured on the latest available standard anteroposterior radiograph using earlier described methods. Eventually we
allocated 11 female patients with a cup inclination angle >45° and a femoral component size <50 mm to this ‘high risk’ group. Five patients in this group had bilateral HRA; one patient fulfilled all high risk criteria bilaterally, four patients only unilaterally, and therefore 12 hips were included in the high risk group for MARS-MRI screening. The ‘low risk’ group consisted of 10 asymptomatic male patients with a unilateral HRA, cup inclination angle <45° and femoral component size >50 mm. The third group consisted of 19 patients (22 hips) who, regardless of risk factors, were scheduled for routine follow-up between November 2011 and May 2012 and acted as a ‘control’ group without risk stratification (Table 6.1). In all three groups, blood serum samples were collected and assessed on cobalt and chromium concentrations. Samples were collected in metal-free vacutainers; the first 5 mL blood was discarded to eliminate metal contamination from the needle. Tubes were stored at 2-8°C and sent to an external laboratory (Ziekenhuis Groep Twente, Hengelo, Netherlands) for analysis. The metal ion levels in whole blood were determined using Atomic Absorption Spectrophotometry (AAS) analysis. Cobalt levels were classified according to guidelines by the Dutch Orthopaedic Society15 with normal Cobalt <40 nmol/L, slightly elevated 40-85 nmol/L, elevated 85-170 nmol/L and extremely elevated >170 nmol/L. All MARS-MRI examinations were performed on a 1,5T MRI (Philips Medical Systems, Best, Netherlands). Scan parameters are listed in table 6.2. All MARS-MRI images were judged by an experienced musculoskeletal radiologist (KB) and validated by a second musculoskeletal radiologist (RH), who were both unaware of the clinical status of the patients. We used the description by Matthies et al of a pseudotumor being a sterile inflammatory lesion found in the soft tissues surrounding a MoM hip arthroplasty.16 Grading of MARS-MRI findings was based on the method described by Anderson et al17 (Table 6.3). Since Harris hip scores (HHS), Oxford hip scores (OHS)18,19 and anteroposterior and lateral radiographs were collected yearly as part of routine follow-up, these were available for all patients. The OHS results were calculated using the original scoring system (12 points being best possible score, 60 points being the worst possible score).
Statistical analysis
Descriptive statistics were used to compare the three study groups. Metal ion data distributions were asymmetric and are expressed as a group median with range. Symmetrical data are represented by a mean and standard deviation (SD). The significant level $\alpha$ is defined as .05 in this study. A post hoc analysis was used to measure the statistical power of the observed difference in pseudotumor occurrence between groups. SPSS software (SPSS Statistics, version 17.0, IBM Corporation, Somers USA) was used for all statistical analyses.

Results
Patient characteristics are shown in Table 6.1. Before the intensified screening protocol was implemented, no pseudotumors had been detected in our cohort of 298 HRAs. With the MARS-MRI screening completed, pseudotumors were observed in all three groups (Table 6.4). The risk for pseudotumor development in the high risk group was 0.45, 0.33 in the low risk group and 0.3 in the control group. However, the statistical power to detect a true significant difference in risk ratios between groups was low (0.11). Overall, in 15 cases of the 44 MARS-MRIs available for analysis, pseudotumor formation had occurred. In total 29 MARS-MRI images were classified as grade A, none as grade B, six as grade C1, eight as grade C2 and one grade as C3. In contrast to the MARS-MRI images, the cobalt levels were normal in 80% of the patients. Two patients had slightly elevated metal ion levels, four patients had elevated levels and two patients had extremely elevated levels. Median Cobalt level for all patients was 24 nmol/L (min-max: 11-
1897 nmol/L). Out of the 15 pseudotumors which were observed on MARS-MRI, there were 12 silent pseudotumors. These patients did not complain of any pain or other symptoms and had excellent clinical outcome scores (HHS >90, Oxford Hip Score <16) with normal radiographs. One female patient from the high risk group with severe MoM disease underwent revision surgery, and one male patient from the control group with moderate MoM disease is scheduled for revision. The revised patient had bilateral HRA: seven years after implantation on her right, six years on her left side.
### Table 6.2, MARS-MRI details

<table>
<thead>
<tr>
<th></th>
<th>TE (ms)</th>
<th>TR (ms)</th>
<th>TI (ms)</th>
<th>Slice thickness</th>
<th>FOV (mm)</th>
<th>Matrix</th>
<th>BW (HZ/pixel)</th>
<th>Coil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronal PDW</td>
<td>30</td>
<td>3000</td>
<td></td>
<td>2.5</td>
<td>230 x 197</td>
<td>328 x 220</td>
<td>435</td>
<td>sense body 16 ch</td>
</tr>
<tr>
<td>Coronal STIR</td>
<td>40</td>
<td>8645</td>
<td>130</td>
<td>2.5</td>
<td>230 x 198</td>
<td>256 x 168</td>
<td>437</td>
<td>sense body 16 ch</td>
</tr>
<tr>
<td>Transverse PDW</td>
<td>30</td>
<td>3576</td>
<td></td>
<td>3</td>
<td>240 x 199</td>
<td>344 x 198</td>
<td>437</td>
<td>sense body 16 ch</td>
</tr>
<tr>
<td>Transverse</td>
<td>40</td>
<td>105000</td>
<td>130</td>
<td>3</td>
<td>280 x 198</td>
<td>280 x 152</td>
<td>435</td>
<td>sense body 16 ch</td>
</tr>
<tr>
<td>Sagittal STIR</td>
<td>40</td>
<td>9570</td>
<td>130</td>
<td>3</td>
<td>230 x 230</td>
<td>256 x 189</td>
<td>438</td>
<td>sense body 16 ch</td>
</tr>
</tbody>
</table>

### Table 6.3, Anderson classification for MoM disease on MARS MRI\(^{13}\)

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Normal or acceptable</td>
<td>Normal post-op appearances including seromas and small haematomas</td>
</tr>
<tr>
<td>B</td>
<td>Infection</td>
<td>Fluid-filled cavity with high signal T2 wall; inflammatory changes in soft tissues, ± bone marrow oedema</td>
</tr>
<tr>
<td>C1</td>
<td>Mild MoM disease</td>
<td>Periprosthetic soft tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either less than 5 cm maximum diameter</td>
</tr>
<tr>
<td>C2</td>
<td>Moderate MoM disease</td>
<td>Periprosthetic soft tissue mass/fluid-filled cavity greater than 5 cm diameter or C1 lesion with either of following: (1) muscle atrophy or edema in any muscle other than short external rotator or (2) bone marrow edema: hyperintense on STIR</td>
</tr>
<tr>
<td>C3</td>
<td>Severe MoM disease</td>
<td>Any of the following: (1) fluid-filled cavity extending through deep fasci, (2) a tendon avulsion, (3) intermediate T1W soft tissue cortical or marrow signal, (4) fracture</td>
</tr>
</tbody>
</table>
There was no pseudotumor observed on her right side but on her left side she had a pseudotumor measuring 105 mm craniocaudally, 71 mm anteroposteriorly and 80 mm mediolaterally (Figure 6.1). Her Cobalt level was extremely elevated (1897 nmol/L). Her HHS score was 91 points and she never complained of pain after HRA. She did however regularly noticed squeaking on the left side, something we had not observed in any other patient from our series. Both cups had a steep inclination angle (left 70°, right 59°). During revision surgery a large fluid filled cyst was excised, extending from the lateral side to the anterior part of the hip joint.

**Figure 6.1,** Large fluid filled cyst left hip, indicating Anderson grade 2 MoM disease. Patient was revised.
Discussion

In our study group of patients with a Recap HRA the prevalence of pseudotumors appeared to be high, with pseudotumor occurrence even in the group defined as having a low risk for ARMD. With an established pseudotumor incidence of 34.1 percent in this concise exploratory study group, we can expect another 87 pseudotumors using an intensified MARS-MRI screening protocol on our entire group of 298 resurfacing hip arthroplasties. Of these 87 pseudotumors, an expected 17 would classify as a grade C2 or C3 pseudotumor with an increased revision risk. As confirmed by other authors, pain was not a very useful indicator for pseudotumor occurrence.\textsuperscript{20,21} Compared to the extent of damage noticed on MARS-MRI and at revision surgery, one has to wonder by which mechanism pseudotumors develop relatively pain free. Mild symptoms and relatively low metal ion levels can contribute to the difficulty of convincing patients to have their HRA revised. However, recent media attention about the negative effects of MoM bearings has scared many MoM patients, who even ask for revision surgery in absence of any symptoms. Although several authors report on pseudotumor rates, the number of studies using other imaging modalities than plain radiographs to detect pseudotumor occurrence is very limited. High rates of pseudotumor occurrence have been found in other studies which used MARS-MRI or computer tomography (CT) scanning. Wynn-Jones reported a similar pseudotumor rate of 36\% using the ASR resurfacing device.\textsuperscript{21} Compared to MoM hip resurfacing, higher pseudotumor rates are reported for MoM total hip arthroplasty. Mistry reported a 58.3\% pseudotumor rate using the Ultima TPS design\textsuperscript{20} and Bosker found a 39\% pseudotumor rate in MoM THA patients who received the M2a-Magnum femoral head and ReCap acetabular component.\textsuperscript{10} Langton described a 13.6\% revision rate for ARMD with the ASR design, but use of MRI or CT scanning was not reported in this paper.\textsuperscript{6} Malviya found a pseudotumor incidence of just 0.15\% using the BHR resurfacing device, although it is not clear from his paper if all patients routinely were scanned using MARS-MRI\textsuperscript{22} To our knowledge, there are no other studies which have investigated the prevalence of pseudotumors with this particular HRA design using imaging modalities other than plain radiographs. The studies by Baad-Hansen and Gagala were limited to 23 and 25 HRA patients respectively with a maximum follow-up of 24 months.\textsuperscript{23,24} Gross and Liu recently published a case series of 740 consecutive procedures with the ReCap HRA design with a follow-up of seven years maximum.\textsuperscript{25} The reported
Kaplan-Meier survivorship with any revision as an end point was 96.4% at 7 years, with only two revisions (0.3%) for adverse wear. Follow-up was limited to clinical outcomes and plain radiographs, but as the possibility of more adverse wear failures was acknowledged by the authors, they started taking metal ion samples routinely. There remains uncertainty on the risk factors for pseudotumor formation with current MoM hips. Studies have suggested that edge-loading resulting from adverse cup orientation and implant design leads to a higher wear of the components and subsequently increases blood metal ion levels. \(^{26,27}\) Clinical studies and reports from arthroplasty registers also implicate smaller components in connection with increased metal ion levels. \(^{13,28}\) Based on these finding, the use of MoM prostheses is supported for appropriately trained surgeons who select appropriate patients. \(^{29}\) Recently, studies have debated risk factors for pseudotumor formation. Kwon et al and Mistry et al showed that pseudotumors can be observed in asymptomatic patients with well positioned and well functioning prostheses. \(^{20,30}\) Recently, Matthies et al reported that pseudotumors are common in well positioned MoM prosthesis. \(^{16}\) These results are confirmed by our study in which pseudotumors were commonly found in asymptomatic patients with well positioned, large components. This suggests that development of pseudotumors is more likely to be dependent on patient susceptibility than on factors such as component size, component positioning or implant design. The risk for pseudotumor formation is higher for any patient with any MoM prosthesis than previously thought. Until now, clinical signs, radiographic evaluation and metal ion levels have been used to identify patients at risk for pseudotumor formation. The best protocol for detecting pseudotumors is not yet defined, but ultrasound scans, CT or MARS-MRI scans are commonly used. Our study indicates that follow-up methods of clinical outcomes and radiographs underestimate the prevalence of pseudotumors after MoM HRA. Moreover, metal ion levels alone are also not sufficient to detect all cases of ARMD. Our findings, especially those from the low risk ARMD group, have prompted us to start using MARS-MRI scans for our whole MoM cohort. Our findings suggest that radiographic screening with MARS-MRI, CT or ultrasound on all patients with a hip resurfacing might be the only option to discover the real magnitude of pseudotumor formation after MoM arthroplasty. There are several limitations of our study. Most importantly, the number of patients is small since we report on an exploratory study at this stage. In spite of this limited number of patients we still feel the need to report on our
preliminary findings of the high number of pseudotumors found on MARS-MRI even in low risk patients with few or no symptoms. In our study group, there were quite a few patients with a steep cup inclination angle, which is considered the only risk factor for ARMD by some authors.\textsuperscript{31} However, despite the fact that we differentiated amongst other factors between high and normal cup inclination, we still found pseudotumors with normally inclined cups. We believe that conventional radiological and clinical follow-up together with metal ion analyses will underestimate the true prevalence of MoM-disease. An intensified screening protocol for pseudotumors with MRI, CT scan or ultrasound is likely to become unavoidable. There is no consensus yet on the clinical relevance of pseudotumors and it may be possible that only some become problematic. There is increasing evidence that the incidence of pseudotumor formation with large diameter (>36 mm) MoM may be higher than assumed so far and the use of these implants has been suspended in the Netherlands.
References


LARGE DIAMETER METAL-ON-METAL HIP PROSTHESIS

1964
Chapter 7

Comparison of different pseudotumor grading systems in a single cohort of metal-on-metal hip arthroplasty patients

Walter van der Weegen, Koen Brakel, Roelof J. Horn, Jorgen A. Wullems, Dirk Das, Peter Pilot and Rob G. Nelissen

Skeletal Radiology 2014 Feb;43(2):149-55
Abstract

Objective. Follow up of pseudotumors observed with Metal-Artefact Reducing Sequence (MARS)-Magnetic Resonance Imaging (MRI) following Metal-on-Metal Total Hip Arthroplasty (MoMTHA) depends on how severe these pseudotumors are graded. Several pseudotumor grading systems for MARS-MRI have emerged but little is known of their validity. We studied the intra- and interobserver reliability of three different pseudotumor grading systems in a single cohort of MoMTHA.

Patients and Methods. Two experienced musculoskeletal radiologists independently used three different pseudotumor grading systems for classifying MARS-MRI results of the same cohort of 42 MoMTHA patients (49 hips, mean follow-up 5.2 years). Intraobserver and interobserver reliability for each grading system was measured using Cohen’s Kappa (κ). Variance in pseudotumor severity grading between systems was analysed.

Results. Intraobserver reliability on grading pseudotumor severity with the Anderson, Matthies and Hauptfleisch grading system scored 0.47, 0.10 and 0.35 (observer 1), and 0.75, 0.38 and 0.42 (observer 2) respectively. Interobserver reliability scores for pseudotumor severity were 0.58, 0.23 and 0.34 respectively.

Conclusion. Intraobserver reliability for grading pseudotumor severity on MARS-MRI ranged from poor to good, dependent on observer and grading system used. Interobserver reliability scored best with the Anderson system. A more succinct pseudotumor severity grading system is needed for clinical use.
Introduction
Although Metal-on-Metal (MoM) hip arthroplasty gained huge popularity in the beginning of this century, critical reports about Adverse Reactions to Metal Debris (ARMD) were published, eventually leading to a recall of some MoM designs\(^1\), and a stop of its use in some countries due to too many questions about its value and safety.\(^2\,3\) Manifestations of ARMD include the occurrence of pseudotumors (Figure 7.1 and Figure 7.2), which may cause severe symptoms, can be locally destructive and might require revision surgery in a proportion of patients.\(^4\,6\) Pseudotumors, defined as a peri-articular mass caused by an immunological delayed hypersensitivity response to metal particles and characterised by a lymphocyte-dominated histological pattern\(^7\), lead to worse clinical outcomes after revision surgery compared to other reasons for MoM revision.\(^8\) Besides the debate about risk factors, incidence and optimal management of pseudotumors, there is no consensus on how to grade the severity of pseudotumors observed on Computer Tomography (CT) or Magnetic Resonance Imaging (MRI) scans. Identified pseudotumors are graded to standardise and summarize results to allow concise management of treatment options for each individual patient. Grading is also of importance to determine changes in the severity of the pseudotumors more accurately when managed conservatively. Few studies were done on the validity of scoring systems for these pseudotumors and controversy exists.\(^9\,10\) The purpose of this paper is to validate three currently used pseudotumor grading systems by measuring their intraobserver and interobserver reliability in a single cohort of Metal-on-Metal Total Hip Arthroplasty (MoMTHA) patients.

Patients and Methods
We retrospectively reviewed a cohort of 42 consecutive MoMTHA patients (49 hips) with a Mallory Head femoral component, a Magnum M2A femoral head and a ReCap resurfacing acetabular component, who had Metal-Artefact Reducing
Sequence (MARS)-Magnetic Resonance Imaging (MRI) scanning, using a scanning protocol described in table 7.1. Since 2011, MARS-MRI scanning and metal ion analysis (determined with Atomic Absorption Spectrophotometry), is part of routine follow-up of MoM patients in our institution, regardless of symptoms. This

**Figure 7.1A** Transverse PDW MARS-MRI of a 60-year-old female showing a large, thick-walled pseudotumor 6 years after Metal-on-Metal total hip arthroplasty. This pseudotumor was graded C3 (Anderson classification) and grade 3 (Matthies and Hauptfleisch classification) by both observers.

**Figure 7.1B** PDW MARS-MRI of the same patient in the coronal plane.
approach is based on recent publications describing a high prevalence of asymptomatic pseudotumors after MoM hip arthroplasty.\textsuperscript{15,16} Clinical examinations (history taking and standard anteroposterior and lateral radiographs) were prospectively collected before surgery, 6 weeks and one year post-surgery and yearly thereafter. Study approval was obtained from the Hospital Ethical Committee. Demographic characteristics of patients are summarized in table 7.2. Two musculoskeletal radiologists (KB, RH), experienced in using pseudotumor grading systems\textsuperscript{11}, independently reviewed all MARS-MRI images, blinded to the clinical status of the patient.

**Figure 7.2A**, Transverse PDW MARS-MRI of a 40-year-old man 7 years after Metal-on-Metal total hip arthroplasty showing a small peri-articular pseudotumor located medial of the hip joint. This pseudotumor was graded C2 (Anderson classification), grade 2A (Matthies classification) and grade 2 (Hauptfleisch classification) by both observers.

**Figure 7.2B**, PDW MARS-MRI of the same patient in the coronal plane.
Both radiologists scored each MARS-MRI on three separate occasions, using a different pseudotumor grading system on each occasion. For intraobserver reliability testing, this was repeated two months later with observers blinded to their first reading and cases placed in random order. The used grading systems were described by Anderson et al\textsuperscript{9}, Matthies et al\textsuperscript{12}, and Hauptfleisch et al.\textsuperscript{13} Pseudotumor grading system details are compared in table 7.3, thereby grouping each severity grade into mild, moderate or severe. This was done according to the original publication\textsuperscript{9}, or by consensus if not described in the original publication.\textsuperscript{12,13} Descriptive statistics were used to report metal ion levels, symptoms and the number of identified pseudotumors per grading system. Differences in median metal ion levels were analysed between groups using the Kruskal-Wallis test. Intraobserver and interobserver reliability on grading pseudotumor severity was calculated for each grading system using Cohen’s Kappa ($\kappa$), excluding cases with no pseudotumor observed in this analysis. We also calculated $\kappa$ per observer on pseudotumor severity grading (mild, moderate or severe) between grading systems. Arbitrary, $\kappa <0.40$ was considered poor, $0.40$ to
0.75 as fair to good and >0.75 as excellent. Descriptive statistics were also used to describe complete agreement between observers on pseudotumor severity per grading system. Complete agreement was defined as both observers classifying one patient exactly the same (i.e. both observer 1 and 2 rate the same patient as having a grade 2a pseudotumor). A 95% Confidence Interval (C.I.) was provided where appropriate. A p<0.05 level was considered significant. All statistics were carried out using SPSS 19.0 software (SPSS Inc., Chicago, Illinois).

<table>
<thead>
<tr>
<th></th>
<th>Anderson</th>
<th>Matthies</th>
<th>Hauptfleisch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>(A) Normal post-op appearances including seromas and small haematomas</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Mild</td>
<td>(C1) Periprosthetic soft tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either &lt; 5 cm maximum diameter</td>
<td>(1) Flat shaped and thin-walled with walls mainly in apposition and fluid-like contents (hypointense on T1, hyperintense on T2)</td>
<td>(I) Thin-walled cystic mass (cyst wall &lt; 3mm)</td>
</tr>
<tr>
<td>Moderate</td>
<td>(C2) Peri-prosthetic soft tissue mass/fluid-filled cavity &gt; 5 cm diameter or C1 lesion with either of following: (1) muscle atrophy or edema in any muscle other than short external rotator or (2) bone marrow edema: hyperintense on STIR</td>
<td>(2a) Thick-walled or irregular with fluid-like contents (hypointense on T1, hyperintense on T2), not flat shaped with &gt; 50% of the walls not in apposition</td>
<td>(II) Thick-walled cystic mass (cyst wall &gt; 3mm, but less than the diameter of the cystic component)</td>
</tr>
<tr>
<td>Severe</td>
<td>(C3) Any of the following: (1) fluid-filled cavity extending through deep fasci, (2) a tendon avulsion, (3) intermediate T1W soft tissue cortical or marrow signal, (4) fracture</td>
<td>3: Solid throughout with mixed signal of any shape.</td>
<td>(III) A predominantly solid mass</td>
</tr>
</tbody>
</table>

Results
In this single cohort of 49 MoMTHA hips, observer 1 identified 23 pseudotumors (46.9%), regardless of the grading system used. Observer 2 identified 21 pseudotumors using the Anderson grading system (42.9%), 22 pseudotumors using the Matthies grading system (44.9%) and 20 using the Hauptfleisch grading (40.8%). Interobserver reliability on whether a pseudotumor was present or not was 0.92 (p <0.001) with the Anderson system, 0.84 (p <0.001) with the Matthies
system and 0.79 (p <0.001) with the Hauptfleisch system. Intraobserver reliability for grading pseudotumor severity with the Anderson, Matthies and Hauptfleisch grading system was 0.47 (p=0.001), 0.10 (p=0.257) and 0.35 (p=0.08) for observer 1, and respectively 0.75 (p<0.001), 0.38 (p<0.001) and 0.42 (p=0.001) for observer 2. Interobserver reliability for pseudotumor severity with the Anderson, Matthies and Hauptfleisch grading system was 0.58, (p =0.001), 0.23 (p =0.001) and 0.34 (p=0.015) respectively. A 60% complete agreement between observer 1 and observer 2 was reached for Anderson C1, 64% for Anderson C2 and 0% for Anderson C3. (Table 7.4).

**Table 7.4, Complete agreement (N) between observer 1 and 2 using the Anderson classification.**

<table>
<thead>
<tr>
<th>Observer 2</th>
<th>A</th>
<th>B</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>24</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C1</td>
<td>3</td>
<td>-</td>
<td>6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C2</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>C3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 7.5, Complete agreement (N) between observer 1 and 2 using the Matthies classification.**

<table>
<thead>
<tr>
<th>Observer 2</th>
<th>No pseudotumor</th>
<th>1</th>
<th>2a</th>
<th>2b</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No pseudotumor</td>
<td>25</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>-</td>
<td>5</td>
<td>3</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>2b</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 7.6, Complete agreement (N) between observer 1 and 2 using the Hauptfleisch classification.

<table>
<thead>
<tr>
<th>Observer 2</th>
<th>No pseudotumor</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No pseudotumor</td>
<td>25</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>

For the Matthies system, 23% complete agreement between observer 1 and observer 2 was reached for grade 1, 40% for grade 2a, 25% for grade 2b and 50% for grade 3 (Table 7.5). For the Hauptfleisch system, 38% complete agreement between observer 1 and observer 2 was reached for grade 1, 36% for grade 2 and 50% for grade 3 (Table 7.6). For observer 1, $\kappa$ on grading pseudotumor severity between the Anderson and Matthies system was 0.32 ($p=0.56$), 0.14 ($p=0.12$) between the Anderson and Hauptfleisch system, and -0.24 ($p=0.796$) between the Matthies and Hauptfleisch system. For observer 2 these scores were 0.11 ($p=0.274$), 0.03 ($p=0.77$) and 0.7 ($p=<0.001$) respectively.

Of the 49 hips, 4 were symptomatic. One patient had moderate symptoms but no evidence of pseudotumor on MARS-MRI, 3 patients had mild symptoms with small to moderately sized pseudotumor visible on MARS-MRI. Median Chromium and Cobalt levels were 54 (range: 10 to 344) and 37.5 (range: 10 to 526) nmol/L respectively. For the patients without a pseudotumor present, these values were 46 (range: 10 to 236) and 32.5 (range: 10 to 174) nmol/L respectively and for the patients with a pseudotumor present these values were 59 (range: 17 to 344) and 51.5 (range: 10 to 526) nmol/L (Table 7.7). Pseudotumors were treated based upon Anderson classification, combined with metal ion levels and symptoms. All C1 and C2 pseudotumors were scheduled for repeated MARS-MRI, one patient with a C3 pseudotumor had extremely elevated metal ion levels but no symptoms. After second opinion this patient was revised. All patients without pseudotumor were scheduled for clinical follow up including metal ion analysis.
Table 7.7, Metal-ion details per pseudotumor grading system

<table>
<thead>
<tr>
<th></th>
<th>Anderson</th>
<th>A</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrome (nmol/L)</td>
<td>46</td>
<td>45</td>
<td>96</td>
<td>344</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Cobalt (nmol/L)</td>
<td>37</td>
<td>43</td>
<td>72</td>
<td>526</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Matthies</th>
<th>No pseudotumor</th>
<th>1</th>
<th>2A</th>
<th>2B</th>
<th>3</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrome (nmol/L)</td>
<td>52</td>
<td>59</td>
<td>89.5</td>
<td>194.5</td>
<td>148</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Cobalt (nmol/L)</td>
<td>37</td>
<td>49.5</td>
<td>44.5</td>
<td>288</td>
<td>123.5</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Hauptfleisch</th>
<th>No pseudotumor</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrome (nmol/L)</td>
<td>52</td>
<td>60</td>
<td>108</td>
<td>148.5</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Cobalt (nmol/L)</td>
<td>38</td>
<td>53</td>
<td>50</td>
<td>123.5</td>
<td>0.83</td>
<td></td>
</tr>
</tbody>
</table>

* Kruskal-Wallis test

Discussion

Pseudotumors can be detected after MoM hip arthroplasty with MARS-MRI, but major clinical questions on severity grading of these pseudotumors are still open for debate. Little consensus exists on follow up of MoM prostheses and their optimal treatment policy (i.e. wait and see versus revision surgery). Even the relevance of elevated metal ion levels in the absence of symptoms or a pseudotumor, the necessity to screen a-symptomatic MoM patients with cross-sectional imaging, or the required frequency of such screening protocols are on debate. This uncertainty on the optimal management of MoM disease in general and pseudotumors in particular, might be partially due to the term pseudotumor being used for a broad variety of a spectrum of lesions, ranging from fluid-filled cysts (Figure 7.3A and B) which might be normal in artificial hip joints to large, complex, and destructive lesions with solid components (Figure 7.4A and B). The use of unvalidated pseudotumor grading systems might contribute to the controversy in the clinical management of problematic MoM implants. In clinical practice, the decision to revise or not will not be a sole consequence of CT or MRI results.

Therefore it is important to validate MARS-MRI based pseudotumor grading systems. Three frequently used pseudotumor grading systems for CT or MRI exists, which had a poor (Matthies and Hauptfleisch grading system) to fair (Anderson grading system) interobserver reliability when grading severity of pseudotumors identified on MARS-MRI.
Intraobserver reliability was not only dependent on observer, but also on the system used, with the Anderson system scoring fair for both observers, while observer 2 scored fair for both the Matthies and Hauptfleisch system and observer 1 scored poor with both these systems. For the Anderson system Chang et al also found a moderate interobserver reliability while Anderson et al found good interobserver reliability. These differences might be explained by the used methodology (we excluded the MARS-MRIs on which no pseudotumor was seen.

**Figure 7.3A** Transverse PDW MARS-MRI of a 59-year-old man 3 years after Metal-on-Metal total hip arthroplasty showing a thin-walled pseudotumor located dorsal of the collum femoris with a high T2 signal, indicating fluid content. Observer 1 graded this pseudotumor as Anderson C2, Matthies 2A and Hauptfleisch 2. Observer 2 rated this as grade C2, 1 and 1 respectively.

**Figure 7.3B** STIR MARS-MRI in the coronal plane of the same patient.
from analysis) but might also occur since the differences between the pseudotumor grades are rather subjective. On observer reliability of the Matthies or Hauptfleisch grading systems, no results could be found in literature. Anderson et al described their system based on a retrospective review of 59 patients (73 MoM hips) and reported that the strongest reliability appeared to be for the grade A, C2 and C3 categories, while the most disagreement appeared to be for categories B and C1. In our study, agreement was slightly higher for C2 than for C1 (64% vs. 60%), while the number of C3 cases was too small to draw any conclusions on observer reliability. Matthies et al retrospectively reviewed 105 revisions of a current-generation MoM hip prosthesis with MARS-MRI and classified pseudotumor contents according to the signal intensity on T1-weighted and T2-weighted images into four different categories. This grading system was later used in a study by Hart et al, who found comparable pseudotumors rates and discussed the high prevalence of fluid-filled cysts. It was hypothesized that these cysts might reflect the required capsulotomy during hip implantation resulted in a pathway of low resistance, allowing the formation of encapsuled fluid collections. As a result they placed less clinical importance of these types of pseudotumors and concluded that a fluid-filled periprosthetic lesion (pseudotumor) may not necessarily indicate the need for revision arthroplasty. No guidelines on clinical follow up based on type of pseudotumor could be deduced from the study by Matthies et al or Hart et al. Hauptfleisch et al retrospectively observed 33 hips with a pseudotumor which they divided into type I, II or III. They considered any solid or cystic mass, in continuity with the hip joint, as a pseudotumor. Isolated distension or thickening of a non-communicating trochanteric bursa was not included. A common characteristic of these grading systems was the analysis of pseudotumor content (i.e. fluid or solid), but other than this each system analysed different pseudotumors details such as size (Anderson system), apposition of walls and shape (Matthies system), or wall thickness (Matthies and Hauptfleisch system). In our experience, strong points of the Anderson grading system are the detailed description of each pseudotumor grade and the incorporation of grade A, allowing a grade for normal MRI scans. Its disadvantages are the absence of a clear description of normal appearance (including seromas and small haematomas), not taking pseudotumor wall thickness in account (which might be an important factor for predicting clinical outcome) , and the 5.0 cm cut-off is rather arbitrary. In our study, the Matthies
grading system had the advantage of a higher interobserver reliability on severe pseudotumors.

**Figure 7.4A**, Transverse PDW MARS-MRI of a 65-year-old female 6 years after Metal-on-Metal total hip arthroplasty showing a pseudotumor with mixed signal intensity 6 years after Metal-on-Metal hip arthroplasty. Both observer rated this as a Anderson C2, Matthies grade 3 and Hauptfleisch grade 3 pseudotumor.

**Figure 7.4B**, PDW MARS-MRI of the same patient in the coronal plane.

The grading system by Hauptfleisch had the advantage of having the least number of grades, making it a straightforward system to use. In our study we observed a
high incidence (41% to 47%, depending on observer and grading system) of pseudotumors after reviewing 49 MoM large head hip arthroplasty cases. Most were asymptomatic (19/23). This is higher than the 36% prevalence rate reported by Wynn-Jones et al\textsuperscript{16}, but lower than the 65% found by Anderson et al.\textsuperscript{9} This might be explained by a twice as long mean follow up in our study compared to the cohort described by Wynn-Jones et al (62 versus 31 months), while the cohort described by Anderson et al retrospectively selected MARS-MRI’s for review, possibly resulting in a higher pseudotumor incidence. Our study was limited since only a very small number severe pseudotumors was included. However this closely reflects daily clinical practice where the difficulty in grading mild to moderate pseudotumors is more of an issue than grading very large, extensive pseudotumors. Strong points of our study are the analysis of both intra and interobserver reliability of all current pseudotumor grading systems. In conclusion, our study is the first which validates different pseudotumor grading systems by applying these different systems to a single cohort of MoM total hip arthroplasties. Both intraobserver reliability and interobserver reliability for grading severity of pseudotumors is limited with all three pseudotumor grading systems. Further validation of all three classification systems on their prognostic value for pseudotumor management is needed.
References


STANDARD HIP PROSTHESIS
METAL-ON-
POLYETHYLENE
Chapter 8

Treatment of pseudotumors after metal-on-metal hip resurfacing based on Magnetic Resonance Imaging, metal ion levels and symptoms

Walter van der Weegen, Thea Sijbesma, Henk J. Hoekstra, Koen Brakel, Peter Pilot and Rob G.H.H. Nelissen
J Arthroplasty 2013  Feb;29(2):416-21
Abstract
Peri-prosthetic pseudotumor formation can be a severe complication following Metal-on-Metal hip resurfacing arthroplasty (MoMHRA), with limited data on the optimal management of this complication. The aims of this study were (1) to evaluate the prevalence and severity of pseudotumors in a consecutive cohort of 248 MoMHRA (214 patients, mean follow-up 4.6 years, range: 1 to 8.2), and (2) to present a clinical guideline for their treatment based on severity grading with Metal-Artefact Reduction Sequence Magnetic Resonance Imaging, metal ion levels and symptoms. Pseudotumor prevalence was 36.3%: 61 mild, 25 moderate and four were graded severe. Five revisions followed, all in symptomatic patients with elevated metal ion levels. Pseudotumor severity grading allowed us to be conservative with revision surgery for mild and moderate MoM disease.
Introduction
Recently controversies occur on the benefit of metal on metal arthroplasty (MoM), due to an increasing number of studies on pseudotumors occurrence next to these types of hip replacements.\textsuperscript{1-3} Adverse peri-prosthetic soft tissue reactions following MoM hip arthroplasty can include metallosis, Asymptomatic Lymphocyte Vasculitis-Associated Lesions (ALVAL) or pseudotumor formation.\textsuperscript{4} Pseudotumors, defined as a solid or fluid mass which has developed in the peri-prosthetic soft tissue\textsuperscript{5}, are considered a severe complication of these MoM implants, which may cause pain, swelling, deep vein thrombosis and extensive soft tissue damage.\textsuperscript{6-8} Interestingly, not all MoM prostheses seem to develop these pseudo tumor sequelae, a debate exists on the prevalence of these pseudotumors, which ranges from less than 1% to 39%.\textsuperscript{9,10} Currently the only treatment option in case of pseudotumors is revision surgery, during which the MoM articulation is replaced by a non-MoM articulation. However, outcome of revision surgery for pseudotumor is poor compared to MoM revision surgery for other reasons.\textsuperscript{11} Incomplete pseudotumor resection and recurrence of pseudotumor, both a reason for re-operation, is reported by Liddle et al\textsuperscript{12} while de Steiger et al found infection to be a major cause for re-revision surgery in MoM hip arthroplasty.\textsuperscript{13} In clinical practice, symptoms (both general health as well as local at the hip region) and metal ion levels are also used next to MARS-MRI pathology about the hip, to guide not only surgical treatment, but also follow up of these patients, despite that controversy exists on the validity of these variables.\textsuperscript{2,14-16} Furthermore, only poor consensus exists on detection of these MoM pseudotumors.\textsuperscript{2,17,18} The aim of this study was to evaluate the prevalence and severity of pseudotumors in a consecutive cohort of MoM hip resurfacings using MARS-MRI. Secondly, a clinical guideline for the treatment of these MoM pseudotumors will be presented based on pseudotumor severity as graded with MARS-MRI, combined with metal ion levels and symptoms.

Patients and Methods
A consecutive cohort of 258 patients (296 MoM hip resurfacing procedures) who had surgery between September 2004 and November 2011. The MoM prosthesis in all patients was the ReCap resurfacing hip (Biomet, Bridgend, South Wales, UK). Data was prospectively collected as part of an Investigational Device Exemption study for this specific MoM hip resurfacing design (Registration: NCT00603395),
before surgery, 6 weeks and one year post-surgery and yearly thereafter. Clinical outcomes and radiographs were collected per protocol from 2004 onwards. The study protocol was extended in 2011 to include baseline cross-sectional imaging (MARS-MRI or ultrasound) and metal ion blood analysis for each patient scheduled for follow up, as a response to the concerns raised on adverse reactions to metal debris. Forty-one patients had a bilateral MoM hip implant, two of these had a different design contra lateral hip resurfacing from another hospital, one received a contra lateral MoM Total Hip Arthroplasty (THA) in our hospital. These three MoM hips were excluded from analyses, all other bilateral cases (n=38) were analysed as separate cases. At the last follow-up in 2012, 17 patients (18 hips) had been revised of which details were published before. After excluding 21 patients (24 hips) for reasons explained in figure 8.1, pseudotumor prevalence using MARS-MRI could be evaluated in 214 patients (248 hips). Mean age of the 235 invited patients was 53.7 years (range: 31 to 76), mean follow up was 4.6 years (range: 1.0 to 8.2). In seven patients (eight MoM hips) a contra-indication for MRI was present, these patients were examined using ultrasound examination of the hip area. Ultrasound examinations were performed in supine, prone and left or right side position with different planes (coronal, transversal and sagittal) to detect hydrops and/or peri-articular masses and fluid collections; if needed duplex ultrasound was used to differentiate between vascular and non vascular lesions. Clinical examination was done using the Oxford Hip Score (OHS) and physical examination (i.e. hip Range of Motion, groin swelling and palpation tenderness). Patients were also questioned about their general health. Since public awareness existed on possible general symptoms of the MoM, questions on symptoms which could be attributed to the MoM implant, were nevertheless posed: “Did general health changed since their hip surgery” in a dichotomous way. Special notice was given to symptoms derived from the NHS advise on follow-up for MoM patients: chest pain or shortness of breath, numbness or weakness, changes in vision or hearing, fatigue, feeling cold or weight gain. An anterior-posterior radiograph of the pelvis and a lateral hip were made annually. At the latest follow up, particular attention was given to radiolucency, evidence of peri-articular masses and peri-prosthetic bone resorption. Radiographs were scored for position of the prosthesis (i.e. inclination of the cup, neck thinning etc). Blood serum samples were collected and assessed on cobalt and chromium
concentrations. Samples were collected in metal-free vacutainers; the first 5mL blood was discarded to eliminate metal contamination from the needle. Tubes were stored at 2-8°C and sent to an external laboratory (Ziekenhuis Groep Twente, Hengelo, the Netherlands) for analysis.

![Study flow diagram](image)

**Figure 8.1, Study flow.**

The metal ion levels in serum blood were determined using Atomic Absorption Spectrophotometry (AAS) analysis. The Medicines and Healthcare products Regulatory Agency (MHRA) statutory body that regulates resurfacing devices in the UK advocates 7 parts per billion (ppb) for chromium and cobalt after MoM hip arthroplasty as a safe upper limit. All MARS-MRI examinations were performed on a 1.5T MRI (Philips Medical Systems, Best, The Netherlands). Scan parameters are listed in table 8.1. MARS-MRI images were judged by one experienced musculoskeletal radiologist and validated by a second radiologist. If patients had two cysts observed on MRI, the maximum diameters of both were added up. In case of disagreement consensus was reached by discussion. Pseudotumor findings were classified according to the grading system described by Anderson (Table 8.2), which has a
good interobserver reliability (κ=0.78, 95% confidence intervals: 0.68–0.88).\textsuperscript{18} We defined pseudotumors to be asymptomatic if patients scored no pain on the Oxford Hip Score (OHS) pain question and if the total OHS score was less than 19\textsuperscript{20}. Our study was approved by the Institutional Review Board (IRB nr. 08.013, 18\textsuperscript{th} December 2008).

**Statistical analysis**

Descriptive statistics were used to report patient characteristics, clinical outcomes and radiographic measurements the number of (asymptomatic) pseudotumors detected with MRI scanning. Serum metal ion data are non-normally distributed, therefore median with interquartile ranges (IQR) were used. Normally distributed data are represented as mean and range. A priori sub analyses were planned on the odds ratios for pseudotumor prevalence based on gender, unilateral or bilateral MoM implants, cup inclination angle (55\textdegree or higher was considered a cut-off point for too steep), component size (femoral component less than 50mm was considered small), neck thinning (neck thinning versus no neck thinning), and elevated blood metal ion levels. The Pearson correlation coefficient between cup inclination and both chromium and cobalt serum levels was determined. The significant level α is defined as .05. All statistics were carried out using SPSS 19.0 software (SPSS Inc., Chicago, Illinois).
### Table 8.1, MARS-MRI details

<table>
<thead>
<tr>
<th></th>
<th>TE (ms)</th>
<th>TR (ms)</th>
<th>TI (ms)</th>
<th>Slice thickness</th>
<th>FOV (mm)</th>
<th>Matrix</th>
<th>BW (Hz/pixel)</th>
<th>Coil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronal PDW</td>
<td>30</td>
<td>3000</td>
<td>2.5</td>
<td>230 x 197</td>
<td>328 x 220</td>
<td>435</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Coronal STIR</td>
<td>40</td>
<td>8645</td>
<td>130</td>
<td>230 x 198</td>
<td>256 x 168</td>
<td>437</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Transverse PDW</td>
<td>30</td>
<td>3576</td>
<td>3</td>
<td>240 x 199</td>
<td>344 x 198</td>
<td>437</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Transverse</td>
<td>40</td>
<td>105000</td>
<td>130</td>
<td>280 x 198</td>
<td>280 x 152</td>
<td>435</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Sagittal STIR</td>
<td>40</td>
<td>9570</td>
<td>130</td>
<td>230 x 230</td>
<td>256 x 189</td>
<td>438</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8.2, Anderson classification for MoM disease on MARS-MRI

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Normal or acceptable</td>
<td>Normal post-op appearances including seromas and small haematomas</td>
</tr>
<tr>
<td>B</td>
<td>Infection</td>
<td>Fluid-filled cavity with high signal T2 wall; inflammatory changes in soft tissues, ± bone marrow oedema</td>
</tr>
<tr>
<td>C1</td>
<td>Mild MoM disease</td>
<td>Periprosthetic soft tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either less than 5 cm maximum diameter</td>
</tr>
<tr>
<td>C2</td>
<td>Moderate MoM disease</td>
<td>Peri-prosthetic soft tissue mass/fluid-filled cavity greater than 5 cm diameter or C1 lesion with either of following: (1) muscle atrophy or edema in any muscle other than short external rotator or (2) bone marrow edema: hyperintense on STIR</td>
</tr>
<tr>
<td>C3</td>
<td>Severe MoM disease</td>
<td>Any of the following: (1) fluid-filled cavity extending through deep fasci, (2) a tendon avulsion, (3) intermediate T1W soft tissue cortical or marrow signal, (4) fracture</td>
</tr>
</tbody>
</table>
Results

*Pseudotumors identified with MARS-MRI*

In 90 hips (85 patients) pseudotumors were detected at MARS-MRI (36.3%, table 8.3). The mean follow-up of these patients was 4.8 years (range: 1.0 to 8.2). No pseudotumors were detected in the seven patients scanned with ultrasound. There were no significant risk groups identifiable (Table 8.4) and there were 80 pseudotumors visible on MRI in patients with low chromium or cobalt levels (Tables 8.5 and 8.6).

<table>
<thead>
<tr>
<th>Table 8.3, Pseudotumor severity grading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Total (n)</td>
</tr>
<tr>
<td>Symptomatic (n)</td>
</tr>
<tr>
<td>Silent (n)</td>
</tr>
</tbody>
</table>

*Fluid collections not graded as MoM disease*

There were 41 cases of fluid-filled cysts observed on MR images which were graded normal (Anderson grade ‘A’). The mean size of these cysts was 26mm (range: 8 to 62).

*Metal ion levels*

Median chromium and cobalt values were 1.82 ppb (IQR: 1.1-3.2) and 1.47 ppb (IQR: 1.1-2.40), but increased per pseudotumor severity group (table 8.7). Eight patients had chromium and cobalt levels >7 ppb, another five patients had chromium values of >7 ppb but cobalt values of <7 ppb. No patients with cobalt values of >7 ppb had Chromium values <7 ppb. Bilateral patients had median chromium and cobalt levels of respectively 2.92 ppb (IQR: 1.82-4.46) and 2.35 ppb (IQR: 1.65-3.49) compared to 1.51 ppb (IQR: 0.98-2.19) and 1.29 ppb (IQR: 0.94-1.71) for unilateral patients. The Pearson correlation between acetabular cup inclination angle and chromium blood-levels was 0.22 (p <0.001). See figure 8.2A. The Pearson correlation between acetabular cup inclination angle and cobalt blood-levels was 0.19 (p = 0.002). See figure 8.2B.
Table 8.4, Odds ratio’s for pseudotumor prevalence

<table>
<thead>
<tr>
<th></th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.91 (0.5 – 1.64)</td>
<td>0.74</td>
</tr>
<tr>
<td>Unilateral MoM</td>
<td>1.25 (0.61-2.55)</td>
<td>0.06</td>
</tr>
<tr>
<td>Femoral head &lt;50mm</td>
<td>1.3 (0.78-2.2)</td>
<td>0.30</td>
</tr>
<tr>
<td>Cup inclination angle of &lt;55°</td>
<td>0.94 (0.53-1.66)</td>
<td>0.83</td>
</tr>
<tr>
<td>General symptoms present</td>
<td>0.71 (0.35-1.45)</td>
<td>0.35</td>
</tr>
<tr>
<td>Femoral neck thinning</td>
<td>1 (0.6-1.67)</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 8.5, 2 x 2 table for Chromium level and pseudotumor occurrence

<table>
<thead>
<tr>
<th></th>
<th>No pseudotumor</th>
<th>Pseudotumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromium &lt;7ppb</td>
<td>161</td>
<td>80</td>
</tr>
<tr>
<td>Chromium &gt;7ppb</td>
<td>4</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 8.6, 2 x 2 table for Cobalt level and pseudotumor occurrence

<table>
<thead>
<tr>
<th></th>
<th>No pseudotumor</th>
<th>Pseudotumor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cobalt &lt;7ppb</td>
<td>163</td>
<td>84</td>
</tr>
<tr>
<td>Cobalt &gt;7ppb</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Figure 8.2A, Chrome versus cup inclination.
Symptoms

Pain in or around the hip area (as a domain of the Oxford Hip score), was reported in 23.6% (n= 60) of all 256 cases, ranging from slight (n=32, 12.6%), mild (n=16, 6.3%), moderate (n=9, 3.5%) or marked (n=3 hips, 1.2%). A wide variety of general symptoms were reported by 44 of all 221 patients (19.9%) and ranged from poor vision, general fatigue, hypertension and other cardiovascular diseases to skin disease, strength loss, weight loss and stomach aches. General health symptoms as specified in the NHS advice on MoM implants are given in table 8.8. Eleven patients reported other cardiovascular symptoms than chest pain, such as hypertension or coronary bypass surgery. Another six patients reported tinnitus.

Plain radiographs

In none of the 221 patients, plain radiographs were indicative for MoM disease. The contrast between plain radiographs and MARS-MRI is seen in figure 8.3A and 8.3B.

Revision case description

Severe MoM disease

Of the four patients with a C3 pseudotumor, two were revised and one is scheduled for revision surgery. Besides a C3 pseudotumor these patients presented with either symptoms and/or metal ion levels >7 ppb. In both revision
cases a large, fluid filled cyst was excised, which was thick-walled in one patient. Post operative histopathology confirmed metallosis for each revised pseudotumor. In both cases an uncemented THA with a ceramic-on-polyethylene bearing was inserted. Metal ion levels dropped significantly six weeks after revision surgery (a 20-fold decrease in one patient and a 10-fold decrease in the other patient). One patient who was without general or hip symptoms and had metal ion levels <7ppb, was treated conservatively. This patient was hesitant to undergo revision surgery and pseudotumor evaluation, including MRI and metal ion levels, is scheduled after six months.

**Moderate MoM disease**
One patient was revised for a C2 pseudotumor with mild hip pain but no general symptoms, during which a fluid-filled cyst was excised. ALVAL was confirmed with post-operative histopathology. For one other patient with a C2 pseudotumor, mild hip pain but no general symptoms, revision surgery is scheduled. Repeated MR scanning and metal ion sampling with an interval of six months was scheduled for all non-revised patients with a C2 pseudotumor.

**Mild MoM disease**
We revised no patients for a C1 pseudotumor, and no revisions are pending for this reason. One patient with elevated metal ion levels (13.7 and 8.88ppb respectively) had no pain the first five post-operative years, but developed increasing pain around the hip during the last two years, which is now moderate. The observed pseudotumor had a maximum 49mm diameter (>50mm will classify as a C2 pseudotumor). Repeated MR scanning and metal ion sampling was scheduled for all patients with a C1 pseudotumor with a time interval of one year.

**Total MoM disease in our cohort**
Before this study, two patients were revised for persistent pain who post-operatively had histopathological evidence of ALVAL to metal debris and two patients were revised for pseudotumor diagnosed with MARS-MRI following our pilot study. Combined with the 90 pseudotumors detected with MR scanning, this results in 94 cases of MoM disease (36.7%) in our total cohort of 256 hips (excluding deceased, lost to follow up and unwilling patients). Until now, seven hips were revised and two revisions are pending MoM disease (3.5%).
Follow-up of patients without a pseudotumor on MR scanning

Patients without a pseudotumor seen with MR scanning are followed up yearly with a clinical examination and metal ion levels.

**Figure 8.3A**, Plain anteroposterior radiograph of patient with C3 pseudotumor.

**Figure 8.3B**, PDW MARS-MRI of same patient showing large pseudotumor.
### Table 8.7, Details per pseudotumor severity

<table>
<thead>
<tr>
<th>Anderson score</th>
<th>Male/Fem. (n)</th>
<th>Bilateral (%)</th>
<th>FU (yr) mean (range)</th>
<th>Acet. Incl. (%) mean (range)</th>
<th>Fem. neck thinning (%) mean (range)</th>
<th>Silent (%)</th>
<th>Chromium ppb median (IQR)</th>
<th>Cobalt ppb median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>113/43</td>
<td>22.7</td>
<td>5.1 (1.0-8.2)</td>
<td>50 (27-71)</td>
<td>2.1 (0-18.5)</td>
<td>n/a</td>
<td>1.82 (1.0-2.85)</td>
<td>1.41 (1.1-2.05)</td>
</tr>
<tr>
<td>C1</td>
<td>44/14</td>
<td>15.5</td>
<td>4.6 (1.1-8.2)</td>
<td>50 (38-66)</td>
<td>1.74 (0-8.7)</td>
<td>82</td>
<td>1.72 (1.25-3.15)</td>
<td>1.29 (1.03-2.09)</td>
</tr>
<tr>
<td>C2</td>
<td>17/6</td>
<td>21.7</td>
<td>4.6 (2.1-7.7)</td>
<td>57 (55-60)</td>
<td>2.1 (0-9.5)</td>
<td>68</td>
<td>3.18 (1.25-5.56)</td>
<td>2.47 (1.24-4.97)</td>
</tr>
<tr>
<td>C3</td>
<td>2/2</td>
<td>25</td>
<td>5.5 (3.3-7.7)</td>
<td>55 (46-70)</td>
<td>2.6 (0-9)</td>
<td>50</td>
<td>6.41 (2.42-42.35)</td>
<td>5.06 (1.86-60.56)</td>
</tr>
</tbody>
</table>

Fem. Indicates female, Acet. Incl. indicates acetabular inclination; Fem. indicates Femoral;

### Table 8.8, Reporting of general health symptoms according to NHS advice on MoM implants

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest pain/shortness of breath</td>
<td>4 (1.5%)</td>
</tr>
<tr>
<td>Numbness/weakness</td>
<td>5 (2.0%)</td>
</tr>
<tr>
<td>Change in vision/hearing</td>
<td>3 (1.2%)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>8 (3.1%)</td>
</tr>
<tr>
<td>Feeling cold</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Weight gain</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>
Discussion

MRI screening a complete cohort of MoM hip arthroplasty patients, we found a high prevalence of pseudotumors, the majority (70/90) asymptomatic. Other authors have confirmed the high prevalence (up to 30%) of asymptomatic pseudotumors in MoM patients, although screening for pseudotumors is generally advised if symptoms are present (FDA) or if the serum metal ions levels are above a certain threshold (UK).\textsuperscript{2,23} Based on our results and from previous reports, we believe that commonly used follow up methods (clinical examination and plain radiographs) will give a gross underestimation of (asymptomatic) pseudotumors in MoM hip arthroplasty. This conventional approach might result in late surgery for pseudotumor, increasing the risk of poor outcome of revision surgery. The early and low-threshold use of cross-sectional imaging might prevent this. In the discussion about the clinical value of both symptomatic and asymptomatic pseudotumors the true incidences are important facts to know. Furthermore, only 13 patients (5.9%) had metal ion levels >7 ppb, the latter also a threshold to initiate MRI screening. This confirms that ion levels do not correlate with visualized adverse local tissue reaction, either noted at the time of revision or on MRI. As for the usefulness of metal ions levels to detect pseudotumors, MacNair found a pseudotumor prevalence of 24% in patients with normal metal ion levels\textsuperscript{16} and Matthies found that patients revised with pseudotumors had similar whole-blood metal ion levels to those who were not revised.\textsuperscript{24} These findings, together with the findings of our study, underline the importance of cross-sectional imaging in MoM patients. The high prevalence of up to 30% or more of asymptomatic pseudotumors in MoM hip arthroplasty, does raise ethical concerns both for the patients as well as for society.\textsuperscript{25,26} However, there is little knowledge about the clinical relevance of these silent pseudotumors and the natural course of pseudotumors. Further, there is no validated follow-up for detected pseudotumors. We propose a conservative approach for mild to moderate pseudotumors (Anderson grade C1 and C2) which are asymptomatic and have normal metal ion levels. Since there is no clear consensus on the optimal treatment of pseudotumors, and revision surgery of these pseudotumors result in poor outcome\textsuperscript{11-13}, future studies with multiple follow-up time points including cross-sectional imaging are needed to validate the optimal management of pseudotumors. Until the optimal management of conservatively treated pseudotumors is established, we suggest that cross-sectional imaging is repeated.
every six months until lesion stability is confirmed. This will provide new insight in
the yet unknown natural history of conservatively treated pseudotumors, while at
the same time minimizing the burden for both patients and for society (economic
costs). The management of pseudotumors after MoM hip resurfacing is hindered
at this moment since only a few, unvalidated, qualitative grading systems exist.\textsuperscript{2,17,18} Although the interrater reliability of the Anderson grading system is
good ($\kappa=0.78$, 95% confidence intervals: 0.68–0.88)\textsuperscript{18} the clinical validation of this
grading system is still limited. This is also the case for other published
pseudotumor grading systems.\textsuperscript{2,17} The importance of a validated management of
pseudotumors is stressed even more since it is estimated that more than a million
large diameter MoM implants were inserted worldwide.\textsuperscript{27} Using a validated
quantitative pseudotumor grading system would also prevent an overly aggressive
surgical treatment of pseudotumors. We advocate an approach of conservative
policy with intensified follow up if a moderate to mild (Anderson class C1 or C2)
pseudotumor at MRI is present with low metal ion levels (<7 ppb) and no
symptoms. We based revision surgery of pseudotumors primarily on
pseudotumor appearance on MRI (Anderson grade C3), and secondly on metal ion
levels (>7ppb) and symptoms.

Limitations
Since our study is cross-sectional in design, no conclusions on the development of
pseudotumors throughout follow-up can be made. The natural course of adverse
reactions to metal debris is unclear, but based on two studies Fary et al suggested
the likelihood of progression.\textsuperscript{15} Sequential MR scanning will be needed to evaluate
any change in pseudotumor size, shape and location.
Despite the problems with these MoM implants some authors still claim they are
useful in the correct setting and if the implant is correct.\textsuperscript{10} But, this approach is
only possible when all risk factors for pseudotumor formation are well
understood. We found an increased risk (however not significant) in men, for
smaller components and for unilateral MoM hip resurfacing. In previous studies,
female patients and age <40 years were found as risk factors for pseudotumors.\textsuperscript{20}
A second limitation is, that a small number of patients did not have MR scanning
due to contra-indications or unwillingness to participate. However, the complete
follow up of MARS-MRI, metal ion levels and hip and general symptoms of the
remaining, consecutive series of this large cohort with a single hip resurfacing
design has not been presented before. One has to keep in mind that the amount of wear depends on details of each specific resurfacing design such as acetabular arc of cover and clearance, thereby limiting the ability to extrapolate our results to other resurfacing designs.\textsuperscript{28,29}

In conclusion, although prevalence of pseudotumors in a single design MoM hip resurfacing is high, the majority of these patients having subclinical appearance of the pseudotumors, and chromium and cobalt levels <7ppb. In contrast to guidelines from national orthopedic boards, we believe that clinical examination and plain radiographs only have a limited role in the detection of pseudotumors. On the other hand, only a small number of pseudotumors is graded severe on MRI. For now, this allows us to be conservative in the management of detected pseudotumors. Data on the future development of mild to moderate pseudotumors is however lacking and there is a clear need for studies presenting multiple follow up points with cross-sectional imaging of these type of pseudotumors.
References


STANMORE CEMENTED
HIP PROSTHESIS
METAL-ON-POLYETHYLENE
Chapter 9

Asymptomatic pseudotumours after metal-on-metal hip resurfacing show little change within one year

Walter van der Weegen, Koen Brakel, Roelof J. Horn, Henk J. Hoekstra , Thea Sijbesma, Peter Pilot and Rob G.H.H. Nelissen

Bone and joint journal 2013 Dec;95-B(12):1626-31
Abstract
We aimed to establish the natural course of unrevised asymptomatic pseudotumours after Metal-on-Metal (MoM) hip resurfacing during a six to twelve month follow-up period. We used repeated Metal-Artefact Reduction Sequence (MARS)-Magnetic Resonance Imaging (MRI) scanning, metal ion analysis and clinical examination to study 14 unrevised cases (mean age 52.7 years) with pseudotumour and a control group of 23 cases (mean age 52.8 years) without pseudotumour. Mean postoperative time to the first MARS-MRI was 4.3 years (range: 2.2 to 8.3), mean time between first and second MARS-MRI was 8 months (range: 6 to 12). With the second MRI, 35 out of the 37 hips (95%) had not changed in pseudotumour severity, one new pseudotumour (Anderson C2 score, moderate) was observed and one pseudotumour was downgraded from C2 (moderate) to C1 (mild). In general, pseudotumour details were hardly changed. Repeated MARS-MRI within one year follow-up in unrevised patients with asymptomatic pseudotumours after MoM hip resurfacing shows little to no variation. In 23 controls without pseudotumour, one new pseudotumour was detected (4%). Since this is the first longitudinal study on pseudotumours using MARS-MRI, our findings need to be interpreted with caution.
Introduction
Metal on Metal prostheses caused a tremendous change in thought on performance of hip arthroplasty. Although problems with this type of implant are now known to society, the policy on what to do with the aftermaths of this implant are still obscure. A simple revision has mediocre results, the effect of the existing pseudotumours caused by these MoM implants is unknown since no follow-up studies are available. There is ample debate on the prevalence of pseudotumours following Metal-on-Metal (MoM) hip replacement. Pseudotumours are believed to develop in reaction to the release of metal debris of the articulating metal surfaces. A retrieval study by Doorn et al report that about one trillion small nanoparticles are released per year in a MoM bearing (14,000 times more particles than with a polyethylene low friction articulation), but little is known of the biological effects of the metals—predominantly cobalt, chromium, and molybdenum—that are released into the body by these implants. Unlike most organic chemicals, metals cannot be eliminated from tissues by metabolic degradation, but only by renal or gastrointestinal excretion. The formation of pseudotumours is believed to be either an allergic response to a normal level of metal wear particles, or a toxic effect of a very high level of particles. Currently the only treatment of a pseudotumour is revision surgery in which the MoM articulation is replaced by a non-MoM articulation. Outcome studies on MoM revision surgery are scarce and have short follow up, but tend to report moderate results, with even a 25% re-revision rate being reported. The clinical relevance of smaller pseudotumours detected with MARS-MRI is unknown. Moreover, there is a lack of knowledge on when and how fast pseudotumours develop, since all cross-sectional imaging studies on pseudotumours except one, have been retrospective in design with only one follow up. Almousa et al recently published the natural history of 15 pseudotumours in a sample unrevised asymptomatic patients using ultrasound examination, and observed both increase (n=6) in pseudotumour size, decrease (n=1) and complete disappearance of pseudotumours (n=3). However, we do not know if and when new pseudotumours are detected with repeated cross-sectional imaging, and, in case of a pseudotumour without the need for immediate revision surgery, (i.e. smaller, less severely graded pseudotumours in asymptomatic patients with (near) normal metal ion levels), what the short term natural history of these pseudotumours is. Our primary aim was to study the
natural course of unrevised mild to moderate pseudotumours in unrevised patients during a six to twelve month follow-up period, using MARS-MRI; Our secondary aim was to study if new pseudotumours were observed in this follow-up period.

**Patients and Methods**

From a previously published cohort of 44 MoM hip replacements\(^{12}\), 37 cases were available for prospective follow-up, who all had a second MARS-MRI (Table 9.1). Two cases were revised after the first MARS-MRI, and four patients (5 hips) refused further MARS-MRI scanning. MARS-MRI scan parameters are given in Table 9.2, all MARS-MRI examinations were performed on a 1.5T MRI (Philips Medical Systems, Best, The Netherlands. Each patient had received a MoM hip resurfacing arthroplasty (ReCap, Biomet, Warsaw, USA) for primary hip osteoarthritis (OA). MARS-MRI was used to score severity of pseudotumours, which was graded by an experienced musculoskeletal radiologist (KB) and validated by a second musculoskeletal radiologist (RH), using the Anderson method (Table 9.3). This method has good interobserver reliability (κ=0.78, 95% confidence intervals: 0.68 to 0.88) as shown in the original publication by Anderson et al.\(^ {13}\) At follow-up, clinical examination, Oxford Hip Score,\(^ {14}\) and a MARS-MRI was made at mean 4.3 years (range: 2.2 to 8.3). Mean time between the first and second MARS-MRI was 8 months (range: 6 to 12). Pseudotumour details (classification, maximum diameter, localisation with respect to the hip joint -anterior, lateral or posterior-, wall thickness and solidity) are shown in Table 9.4. We defined a pseudotumour as a peri-prosthetic cavity, either fluid-filled or having a solid content, which in case of being fluid-filled communicates with the hip joint. Pseudotumour wall thickness was measured at the site where wall thickness appeared to be thickest: ≥3mm was considered to be thick, <3mm was considered thin.\(^ {15}\) High MRI signal intensity was associated with fluid, low signal intensity with solid pseudotumour content. Bone marrow edema and compromise of nerve or blood vessel structures was systematically analysed for each MRI scan by both radiologists. Serum ion samples (Chromium and Cobalt) were collected at both MRI time points and analyzed as previously described.\(^ {12}\) Since little is known on short term variability of chromium and cobalt levels, a difference of +/- 5% between metal ion levels was considered a true difference. This was based on the findings by Khan that a short exercise bout resulted in 11% to 13% increased...
metal ions concentration. The Oxford Hip Score (OHS) ranges between 48 (least problems) to 0 (most problems) and was also recorded at both time points.

<table>
<thead>
<tr>
<th>Table 9.1, Patient demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Age (years)*</td>
</tr>
<tr>
<td>FU (years)*</td>
</tr>
<tr>
<td>Femoral Component size (mm) **</td>
</tr>
<tr>
<td>Cup inclination*</td>
</tr>
<tr>
<td>Oxford Hip score *</td>
</tr>
<tr>
<td>Chrome (ppb)**</td>
</tr>
<tr>
<td>Cobalt (ppb)**</td>
</tr>
</tbody>
</table>

*mean is presented with range between brackets. ** median is presented with IQR between brackets
### Table 9.2, MARS-MRI details

<table>
<thead>
<tr>
<th></th>
<th>TE (ms)</th>
<th>TR (ms)</th>
<th>TI (ms)</th>
<th>Slice thickness</th>
<th>FOV (mm)</th>
<th>Matrix</th>
<th>BW (HZ/pixel)</th>
<th>Coil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coronal PDW</td>
<td>30</td>
<td>3000</td>
<td>2,5</td>
<td>230 x 197</td>
<td>328 x 220</td>
<td>435</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Coronal STIR</td>
<td>40</td>
<td>8645</td>
<td>130</td>
<td>230 x 198</td>
<td>256 x 158</td>
<td>437</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Transverse PDW</td>
<td>30</td>
<td>3576</td>
<td>3</td>
<td>240 x 199</td>
<td>344 x 198</td>
<td>437</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Transverse</td>
<td>40</td>
<td>105000</td>
<td>130</td>
<td>280 x 198</td>
<td>280 x 152</td>
<td>435</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
<tr>
<td>Sagittal STIR</td>
<td>40</td>
<td>9570</td>
<td>130</td>
<td>230 x 230</td>
<td>256 x 189</td>
<td>438</td>
<td>sense body 16 ch</td>
<td></td>
</tr>
</tbody>
</table>

### Table 9.3, Anderson classification for MoM disease on MARS-MRI

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Normal or acceptable</td>
<td>Normal post-op appearances including seromas and small haematomas</td>
</tr>
<tr>
<td>B</td>
<td>Infection</td>
<td>Fluid-filled cavity with high signal T2 wall; inflammatory changes in soft tissues, ± bone marrow oedema</td>
</tr>
<tr>
<td>C1</td>
<td>Mild MoM disease</td>
<td>Periprosthetic soft tissue mass with no hyperintense T2W fluid signal or fluid-filled peri-prosthetic cavity; either less than 5 cm maximum diameter</td>
</tr>
<tr>
<td>C2</td>
<td>Moderate MoM disease</td>
<td>Peri-prosthetic soft tissue mass/fluid-filled cavity greater than 5 cm diameter or C1 lesion with either of following: (1) muscle atrophy or edema in any muscle other than short external rotator or (2) bone marrow edema: hyperintense on STIR</td>
</tr>
<tr>
<td>C3</td>
<td>Severe MoM disease</td>
<td>Any of the following: (1) fluid-filled cavity extending through deep fascia, (2) a tendon avulsion, (3) intermediate T1W soft tissue cortical or marrow signal, (4) fracture</td>
</tr>
</tbody>
</table>
## Table 9.4. Pseudotumor details per case

<table>
<thead>
<tr>
<th>Casus</th>
<th>Anderson score</th>
<th>Location*</th>
<th>Wall</th>
<th>Max diam (mm)**</th>
<th>Connected to joint</th>
<th>Nerve/blood vessel compromised</th>
<th>MRI signal***</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C1</td>
<td>lateral</td>
<td>thick</td>
<td>40</td>
<td>yes</td>
<td>no</td>
<td>mixed</td>
<td>bone marrow edema</td>
</tr>
<tr>
<td>2</td>
<td>C1</td>
<td>dorsal</td>
<td>thin</td>
<td>18</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>C1</td>
<td>dorsal</td>
<td>thin</td>
<td>19</td>
<td>yes</td>
<td>no</td>
<td>mixed</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>C1</td>
<td>lateral</td>
<td>thin</td>
<td>32</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>C1</td>
<td>lateral</td>
<td>thin</td>
<td>50</td>
<td>no</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>C1</td>
<td>lateral</td>
<td>thin</td>
<td>50</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>C1</td>
<td>dorsal</td>
<td>thin</td>
<td>30</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>C2</td>
<td>lateral-dorsal</td>
<td>thin</td>
<td>53</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>C2</td>
<td>dorsal-medial</td>
<td>thick</td>
<td>45</td>
<td>yes</td>
<td>no</td>
<td>mixed</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>C2</td>
<td>lateral</td>
<td>thin</td>
<td>69</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>C2</td>
<td>lateral</td>
<td>thin</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>C2</td>
<td>dorsal-lateral</td>
<td>thin</td>
<td>78</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>C2</td>
<td>dorsal</td>
<td>thin</td>
<td>80</td>
<td>yes</td>
<td>no</td>
<td>high</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>C3</td>
<td>dorsal-ventral</td>
<td>irregular thick</td>
<td>70</td>
<td>yes</td>
<td>no</td>
<td>mixed</td>
<td>bone cyst</td>
</tr>
</tbody>
</table>

* relative to hip joint; **maximum pseudotumour in any direction; *** high MRI signal intensity is associated with fluid content, low signal intensity with solid content
**Statistical analysis**

Descriptive statistics were used to report patient characteristics and observations such as the number, size and appearances of pseudotumours detected with MRI scanning. Metal ion data and pseudotumour dimension distributions were asymmetric and are expressed as a group median with interquartile range (IQR). Normal distributed data are represented by a mean and range. A qualitative analysis was done for each change in pseudotumour details. Differences in mean values were tested with (two-sided) t-test, differences in median values with the Mann-Whitney test. Significance level was defined at 0.05, 95% Confidence Intervals (C.I.) are provided were appropriate.

**Results**

Details of the 14 pseudotumours observed with the first MRI are given in table 9.4. At first MARS-MRI, the majority of pseudotumours (10/14) were fluid-filled cysts, only four showed a mixed MRI signal intensity indicating a more solid content. Three out of four solid pseudotumours were thick-walled, whereas all 10 fluid-filled pseudotumours were thin-walled. Maximum diameter ranged from 18mm to 80mm. One pseudotumour was graded as Anderson score C3 (severe MoM disease), six as Anderson C2 (moderate MoM disease) and 7 as Anderson C1 (mild MoM disease). Median Chromium and Cobalt for the solid pseudotumours was 3.1ppb (IQR: 1.7-5.6) and 2.1ppb (IQR: 1.8-4.5) versus 3.0 (IQR: 1.6-5.1) and 2.3 (IQR: 1.1-5.1) for the fluid-filled pseudotumours. There were no changes observed in pseudotumour position, wall thickness or content (based on MRI signal intensity) for any of the pseudotumours between both time points. Median pseudotumour diameter decreased from 50 mm (IQR: 32-70) to 46mm (IQR: 37-69). There were five pseudotumours where the maximum diameter had not changed, five pseudotumours had become smaller (mean absolute change -13mm, range: -32 to -2mm), and four had grown (mean absolute change +26mm, range: 7 to 33mm) (Figure 9.1). Thirty-five out of the 37 hips had not changed according to the Anderson grading system. In two cases (5%), the Anderson pseudotumour grade had changed between the two MRI scans: one C2 pseudotumour was observed on the second MRI, which had not been there at the first MRI scan (Figure 9.2A and 9.2B). One pseudotumour was downgraded from C2 to C1 (Figure 9.3A and 9.3B). See table 9.5. Median chromium increased from 1.7 ppb (IQR: 1.0-3.8) to 2.1 ppb (IQR: 1.1-3.6) and median cobalt decreased from...
1.4 ppb (IQR: 0.9-2.5) to 1.3 ppb (IQR: 0.9-3.5) but no metal ion level had changed more than +/− 5%. In the pseudotumour group, mean OHS improved from 32.1 (range: 42 to19) points pre-operatively to 43.2 (range: 48 to 39) at first MRI follow-up time point (40.7, range: 48 to-31 at second MRI time point). In the control group OHS improved from 28.9 (range: 39 to 11) points pre-operatively to 42.1 (range: 48-27) at first MRI follow-up time point (42.2, range: 48 to 27 at second MRI time point).

*median value is presented with IQR between brackets. Pseud.: indicates Pseudotumour

**Table 9.5, Pseudotumour and metal ion levels comparison between first MRI (mean 4.3 years postoperative, range: 2.2 to 8.3) and second MRI (mean 5.0 years postoperative, range: 3.2 to 9.1)**

<table>
<thead>
<tr>
<th>MRI 1 → MRI 2</th>
<th>N</th>
<th>Pseud. size (mm) MRI 1*</th>
<th>Pseud. size (mm) MRI 2*</th>
<th>Chromium 1 (ppb)*</th>
<th>Chromium 2 (ppb)*</th>
<th>Cobalt 1 (ppb)*</th>
<th>Cobalt 2 (ppb)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A → A</td>
<td>22</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1 → C1</td>
<td>7</td>
<td>40 (19-50)</td>
<td>38 (30-41)</td>
<td>1.4 (1.1-4.0)</td>
<td>1.7 (1.4-4.8)</td>
<td>1.8 (1.1-3.5)</td>
<td>1.2 (0.9-3.5)</td>
</tr>
<tr>
<td>C2 → C2</td>
<td>5</td>
<td>74 (51-80)</td>
<td>70 (68-78)</td>
<td>3.4 (1.1-6.0)</td>
<td>4.2 (1.6-13.1)</td>
<td>3.2 (1.2-5.4)</td>
<td>4.2 (1.9-8.9)</td>
</tr>
<tr>
<td>C3 → C3</td>
<td>1</td>
<td>70</td>
<td>70</td>
<td>2.4</td>
<td>1.5</td>
<td>3.0</td>
<td>2.2</td>
</tr>
<tr>
<td>A → C2</td>
<td>1</td>
<td>n/a</td>
<td>55</td>
<td>0.5</td>
<td>0.8</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>C2 → C1</td>
<td>1</td>
<td>53</td>
<td>39</td>
<td>3.6</td>
<td>3.9</td>
<td>6.4</td>
<td>5.3</td>
</tr>
</tbody>
</table>

*median value is presented with IQR between brackets. Pseud.: indicates Pseudotumour

**Figure 9.1, Absolute change pseudotumor size.**
**Change in treatment**

After the first MARS-MRI, one patient was considered for revisions surgery and 13 for intense follow-up, without immediate need for revision surgery of their MoM implant. Based on the results of second MARS-MRI, metal ion levels and symptoms 6 to 12 month later, this clinical advice was not changed in any of the patients.

**Description of the 2 cases changed in Anderson grading**

**Case 1**
The first MRI images of a 67 year old male patient who did not have any evidence for pseudotumour formation at that time (Figure 9.2A) are compared with the images of the second MARS-MRI (Figure 9.2B) when a C2 pseudotumour was detected, 3.5 years after implantation. Time between scanning was 11 months. A thin-walled fluid-filled cyst developed lateral to the hip joint with a maximum diameter of 55 mm in cranio-caudal direction and a thin dorsal connection to the joint space. Based on size, signal intensity and connection to the joint space, this cyst was classified as a C2 pseudotumour. Between MRI scanning, OHS score deteriorated from 41 points to 33 points, although hip pain was unchanged (mild). Chromium and Cobalt levels remained stable at 0.9 and 0.8ppb respectively.

**Case 2**
In a 57 year old male patient, the pseudotumour was downgraded from C2 to C1, see figure 9.3A and 9.3B. This patient had bilateral MoM hip resurfacing, with bilaterally a pseudotumour observed. The pseudotumour on the right hip, reduced from 53 mm to 39 mm in the six months between MRI scanning. Consequently, the Anderson classification changed from C2 to C1. Between MRI scanning, OHS deteriorated from 44 to 36 points, with hip pain deteriorating from very mild to mild, while Chromium and Cobalt levels improved from 6.4 to 3.6ppb and from 5.3 to 3.9ppb respectively.
Figure 9.2A, First MARS-MRI.

Figure 9.2B, Second MARS-MRI.
Figure 9.3A, First MARS-MRI.

Figure 9.3B, Second MARS-MRI.
Discussion

We found that only 5% of the included, small to moderate sized, asymptomatic pseudotumours after MoM hip resurfacing, changed in severity using a six to twelve months interval to repeat MARS-MRI. In the control group without pseudotumour (23 hips), one new pseudotumour was detected (Anderson grade C2) but this patient had no change in metal ion levels or hip pain. In the pseudotumour group (n=14), pseudotumour severity was downgraded in one case (from Anderson grade C2 to C1). Accordingly, metal ion levels decreased in this patient but in contrast his hip pain deteriorated from very mild to mild.

Based on these results clinical treatment was left unchanged for all included patients, indicating that a >1 year interval between consecutive cross-sectional imaging appears to be safe. On this last topic no evidence is available. How much deterioration of symptoms and metal ion levels should trigger additional cross-sectional imaging cannot be concluded from our results, since we observed only a very small variation and sometimes contradictory development in metal ion levels and symptoms between both MRI time points. Longer follow up with an extensive screening protocol is needed. Analysing all included pseudotumours, maximum diameter both increased (n=6) and decreased (n=6), although the observed differences were small to very small. None of the pseudotumours changed in appearance or location.

Previous studies using cross-sectional imaging of pseudotumours after MoM hip arthroplasty were retrospective in design, used only one time point for imaging and had considerable variation in follow up duration. Recently, Almousa et al published the first report using repeated ultrasonography (US) in a cohort of 15 pseudotumours and five isolated fluid collections in a variety of hip replacement types (13 MoM THA, four MoM hip resurfacings and three metal-on-polyethylene bearings). In their series, three pseudotumours had such an increase in size (2.2-fold to 11.4-fold) that it was deemed clinically significant. In our series, we observed no clinically relevant change in pseudotumour size or severity. This might be explained by the shorter follow-up in our study (mean of eight months versus 25.8 months).

There is limited data available on when pseudotumours develop and on how fast pseudotumours change over time. There is also no consensus on the exact definition of a pseudotumor, with different lesions included such as solid pseudotumors or fluid-filled lesions (which might fluctuate more in time), making
it more difficult to guide clinical management of pseudotumours after MoM arthroplasty. Most orthopedic societies and national boards advise computer tomography (CT) or MARS-MRI only in symptomatic patients. However, high prevalence rates of asymptomatic pseudotumours after cross-sectional imaging were reported by Kwon et al (6.5%), Wynn-Jones et al (36%) and Mistry et al (58.3%). How pseudotumours can remain asymptomatic is not known. To our knowledge, no explanation for the absence of symptoms in case of pseudotumours has been presented in literature. Since we know that asymptomatic pseudotumours will be missed, the validity of the advices issued by the FDA and national boards can be questioned. Accepting the risk of missing pseudotumours might outweigh the potential risk of overtreatment based on positive MRI findings, since the clinical relevance of mild to moderate pseudotumours is not yet fully known. On the other hand, one can state that all MoM patients need to be investigated with cross-sectional imaging at least once, to establish a pseudotumour baseline status for each individual patient. Furthermore the FDA MoM safety communication does provide little detail on how to interpret more detailed cross-sectional imaging results and how observed pseudotumours should be treated. There is no study comparing the effectiveness of US, MRI or CT for detecting pseudotumors. US diagnostics is user dependent and provides less detailed imaging compared to MRI but the presence of a metal prosthesis does not compromise US imaging, it is relatively cheap to perform and is widely available, and is therefore considered the preferred initial investigative tool by several authors. According to Fary et al CT diagnostics is not suitable as a screening tool for pseudotumor detection but they consider MRI a suitable tool for making a definitive diagnosis of a mass resulting from an adverse reaction to metal debris. All three modalities have advantages and disadvantages regarding radiation, costs and accuracy and therefore it remains debatable which modality is best for (initial) screening for pseudotumor occurrence. The exact description and grading of pseudotumours has not fully matured. As pointed out by Anderson et al, validation of a grading system is likely to take several years, since mild degrees of disease in asymptomatic patients do not warrant intervention, thereby preventing surgical or histopathological outcome data for this group. Only stability in a longitudinal study will be a useful marker of the validity of mild disease grades. In our own studies experienced musculoskeletal radiologists reported a learning curve evaluating pseudotumours.
We therefore recommend that more than one radiologist is involved in analysing MARS-MRI’s. Also, the use of maximum diameter as an important part of grading pseudotumour severity has limitations, and since the changes in pseudotumour size are very small during a six to 12 months period, measurement error has to be taken into account. Possible factors influencing the MR images when a pseudotumour is present, such as time of day or any physical activity shortly before acquiring the images need to be established. Furthermore, a long thin pseudotumour might be considered a grade C2 (moderate) or C3 (severe) pseudotumour based on maximum diameter, without actually involving a large volume. Besides maximum diameter, other considerations such as MRI signal intensity, cyst wall thickness and position might also be important to evaluate pseudotumour changes in time. The observation within our series that mild to moderate pseudotumours remained fairly stable with MARS-MRI evaluation over a six to 12 month period, for now validates our conservative approach for these pseudotumours, which is in agreement with other authors.\textsuperscript{13,24} Using a >12 months interval to repeat cross-sectional imaging for smaller, non-revised asymptomatic pseudotumours, might help to control the enormous worldwide costs involved. Lloyd et al estimated that annual metal ion analysis and MRI scanning of MoM patients would increase UK nationwide costs with 72.6 million UK pounds for a 5 year period, compared to standard THA follow up costs.\textsuperscript{25} One even has to consider the possibility to treat larger asymptomatic pseudotumours conservatively if metal ion levels are normal and the pseudotumour is positioned in a relative safe position, although the current consensus is that larger pseudotumours need to be revised.\textsuperscript{17,27,28} The need for revision is unquestioned for more extensive pseudotumours which cause symptoms, extensive soft tissue damage and compromise other structures such as blood vessels and nerves.

In conclusion, we show little value to repeat MRI within one year for mild to moderate sized asymptomatic pseudotumours after MoM hip resurfacing, since the few observed changes were minimal and did not change clinical treatment. But there is a value for repeated examinations with longer term follow-up as was shown by Almousa et al.\textsuperscript{11} Since our study is the first longitudinal study on pseudotumours using MARS-MRI, our findings need to be interpreted with caution.
References


25. Lloyd et al. EFFORT 2012 proceedings.


SHORT STEM UNCEMENTED TOTAL HIP PROSTHESIS CERAMIC-ON-CERAMIC
Chapter 10

Measuring change in pseudotumor dimensions after metal-on-metal hip resurfacing. A pilot study.

Walter van der Weegen, Peter Pilot, Renate Raidou, Thea Sijbesma, Bart Kaptein and Rob G.H.H. Nelissen

Submitted
Abstract

Objectives: To validate clinical measurements of (change in) pseudotumor maximum diameter and estimated volume with three-dimensional region-of-interest (3-D ROI) volume measurements.

Methods: Repeated Metal-Artefact Reduction Sequence (MARS)-Magnetic Resonance Imaging (MRI) scans of 13 cases of non-revised pseudotumors after Metal-on-Metal hip resurfacing were reviewed. Mean time between first and second MARS-MRI was 7.5 months (range: 6 to 12 months). Pseudotumor dimensions were measured by: (1) Maximum diameter in one plane (MD); (2) Estimating pseudotumor volume based on maximum diameter in three different planes (EV); (3) Three-dimensional (3-D) Region-Of-Interest (ROI) based volume (V) method.

Results: Correlation was strongest between EV and V (0.86, p<0.000). EV overestimated V with a mean of 72.6%, more so in non-ellipsoid pseudotumors than in ellipsoid pseudotumors. Median change for MD between first and second MARS-MRI was 0.0cm (range: -1.5 to 3.4); -0.5ml for EV (range: -16.4 to 45.5); and 0.5ml for V (range: -7.7 to 5.2). The percent change in pseudotumor dimensions was not significantly different between MD, EV and V.

Conclusion: Estimating pseudotumor volume in clinical practice using maximum diameter in three different planes is easily attainable, has a strong correlation with a 3D-ROI method, and can be used for monitoring pseudotumor volume over time. Clinicians need to be aware of overestimating pseudotumor volume when using this method.
Introduction

Pseudotumors, defined as a peri-articular mass caused by an immunological delayed hypersensitivity response to metal particles and characterised by a lymphocyte-dominated histological pattern, are regularly observed after Metal-on-Metal (MoM) hip arthroplasty. They may cause severe symptoms, have been found to be locally destructive and might require revision surgery. The clinical relevance of pseudotumors is however not fully understood. For that purpose, clinicians need to know if pseudotumors containing fluid are less destructive than more solid pseudotumors, if thick-walled pseudotumors are more aggressive than thin-walled pseudotumors. Furthermore, the location of the pseudotumor (i.e. near nerves or blood vessels) influences the need for re-operation. The clinical relevance of these details are validated by recent publications.

Another important clinical issue in these patients with MoM bearing is pseudotumor size. The latter is taken into account for the assessment of pseudotumor severity and to determine if revision surgery is indicated. Most reports on pseudotumors are cross sectional studies, only two longitudinal reports (one with sonography and one with MRI) on pseudotumor development exist. These indicate that pseudotumors with time not only can increase in size but that some pseudotumors can diminish in size or even disappear. To study these changes in pseudotumor size over time, it is important that pseudotumor dimensions can be measured accurately. Although with cross-sectional imaging measuring maximum pseudotumor diameter in clinical practice is straightforward, it might be that this is not accurate enough for detecting changes in pseudotumor size over time. Measuring pseudotumor volume could potentially detect changes in pseudotumor size over time more accurately. Pseudotumor volume might also be more relevant than maximum diameter to grade pseudotumor severity and to quantify the potential for soft tissue damage. Although this issue is not yet discussed in MoM pseudotumor literature or incorporated in current pseudotumor grading systems, authors have used estimated pseudotumor volume to describe their cross-sectional imaging results. Lack of experience on measuring volume can partly be explained by the relative short period of research into pseudotumors after MoM hip arthroplasty. In other medical fields, like oncology, much more experience with tumor volume measurements is available. Imaging-based tumor volume
measurements before, during, and after therapy have become essential components of cancer management. Since volume measurements in oncology are that vital, the technique of volume measuring has evolved enormously over the last decade. With the advancement and availability of commercially available computer software for quantitative analysis in cross-sectional imaging, the entire tumor, regardless of its shape, can be identified and traced as a region of interest (ROI) on each imaging slice, allowing 3-dimensional (3D) ROI-based quantitative measurement of tumor volume.

The primary aim of our study was to validate clinical measurements of (change in) pseudotumor dimensions (maximum diameter and estimated volume) with a three-dimensional region-of-interest-based volume (V) method.

**Patients and methods**

**Patient population**

All patients treated with a MoM hip resurfacing in our institution were part of a registration study for which ethical approval was obtained from the Institutional Ethics Committee. Metal-Artefact Reducing Sequence (MARS)-Magnetic Resonance Imaging (MRI) was made for routine follow up for all patients with MoM hip resurfacing arthroplasty, after a pilot study showed a high pseudotumor prevalence. All MARS-MRI’s were evaluated by an experienced musculoskeletal radiologist who used the Anderson classification to grade pseudotumor severity. For this study, 11 patients (13 hips) with mild (n=6) to moderate (n=7) pseudotumors which were treated conservatively, were available for follow-up (Figure 10.1). None of these patients had elevated metal ion levels and all were asymptomatic, mean follow-up at the first MARS-MRI was 5.3 years (range: 2.4 to 7.5 years). To monitor the course of pseudotumor development, a sequential MRI was done at a mean of 7.5 months (range: 6 to 12 months). See table 10.1 for patient details.

<table>
<thead>
<tr>
<th>Patients/hips (n)</th>
<th>Age Mean (range)</th>
<th>Male/female (n)</th>
<th>FU (years) Mean (range)</th>
<th>Chrome (ppb) Median</th>
<th>Cobalt (ppb) Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/13</td>
<td>52.7 (41-61)</td>
<td>6/5</td>
<td>5.3 (2.4-7.5)</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

186
MRI analysis

All MARS-MRI studies were performed on a 1.5 T scanner (Philips Gyroscan, Best, The Netherlands) using the system’s body coil. All patients were scanned supine. On the acquired DICOM images, stored in the Picture Archiving and Communication System (PACS), maximum diameter was measured on the workstation. Maximum pseudotumor diameter was measured three times: once in the transversal, once in the coronal and once in the sagittal plane. For analysis, the largest diameter of these three measurements was used, resulting in a clinically measured maximum diameter (MD). All clinical distance measurements were done by one observer (WvdW) and independently repeated by a second observer (PP). Pseudotumor volume was measured using a ROI-based volume measurement (V). V was derived from the MRI data in the three different directions, which were combined to generate one single 3D image with high resolution and isotropic sampling (Figure 10.2). For this, the coronal, sagittal, and transverse MRI datasets were rigidly transformed to one single dataset using translation and rotation. This was possible because the images were acquired during one image session and at the location of the hip, the effect of movement artefacts caused by for example the breathing of the patient was minimal. The parameters for the transformations were obtained from the MRI image header information. The voxel values of the newly generated image were calculated by averaging the values of the aligned MRI images. A ROI-based method was used to

Cases MRI scanned in pilot study: n=44

Excluded cases: n=31;
- Cases without pseudotumor (n=29)
- Cases revised after pilot study (n=2)

Cases available for pseudotumor dimension measurements: n=13

Figure 10.1, study flow.
calculate pseudotumor volume from the newly generated images. For this, the entire pseudotumor region was identified and traced using Amira software (Visualization Sciences Group, Bordeaux, France). A 3D ROI-based volume was calculated by the summation of all pseudotumor areas in each slice and multiplication by the slice profile. Correction for partial volume effects was done by multiplying the volume of the surface voxels by a filling fraction. Filling Fraction = 0.5*(MAX pixel intensity + MIN pixel intensity) / (MAX pixel intensity).

A less complicated measurement of pseudotumor volume is to multiply the transversal maximum diameter with the coronal and sagittal maximum diameters, resulting in a clinically estimated volume (EV) representing a simple box model. To test how much this method of measuring EV is influenced by pseudotumor shape, we classified pseudotumors into two categories: ellipsoid (Figure 10.3) and non-ellipsoid shape (Figure 10.4). Outcomes of EV for ellipsoid and non-ellipsoid were compared with pseudotumor volume V.

Figure 10.2, 3-D reconstruction of pseudotumor.
For comparison, all three measurements (MD, EV and V) were done with the first and second MRI dataset of each included patient. Outcomes >2 standard deviations were considered as outlier and were excluded from analysis.

**Figure 10.3**, Ellipsoid pseudotumor.

**Figure 10.4**, Non-ellipsoid pseudotumor.

**Statistical analysis.**
Descriptive statistics were used to report MD, EV and V, with subgroup analysis for ellipsoid versus non-ellipsoid pseudotumors. Diameter and volume data were
tested for normal distribution using the Shapiro-Wilk test. The relationship between MD, EV and V was calculated using Spearman’s rank correlation coefficient. Change in MD, EV or V between the first and second MRI dataset were tested with a Wilcoxon signed rank test. The Friedman test was used to test if using a different measuring method (MD, EV or V) influenced the observed change significantly. For this purpose, the percent change per case was calculated. The significance level $\alpha$ was set at 0.05. All statistics were carried out using SPSS 19.0 software (SPSS Inc., Chicago, Illinois).

**Results**
MR Imaging Analysis per cases
By using 3D images with high resolution and isotropic sampling, identification of the entire pseudotumor region using Amira was easier than by using the coronal, sagittal and transverse MRI datasets separately. This is because the voxels belonging to the pseudotumor were better connected with each other in the adjacent slices and it was easier to review the pseudotumor shape in the different views (x, y and z) because of the isotropic sampling. Eight pseudotumors were ellipsoid, MD ranged from 1.7cm to 8.6cm, EV from 1.5ml to 87ml and V ranged from 1.1ml to 35.4ml (Table 10.2).

<table>
<thead>
<tr>
<th>Case</th>
<th>Ellipsoid</th>
<th>MD1 (cm)</th>
<th>EV1 (ml)</th>
<th>V1 (ml)</th>
<th>Anderson grade 1</th>
<th>MD2 (cm)</th>
<th>EV2 (ml)</th>
<th>V2 (ml)</th>
<th>Anderson Grade 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>1.80</td>
<td>2.48</td>
<td>1.33</td>
<td>C1</td>
<td>1.74</td>
<td>1.97</td>
<td>1.10</td>
<td>C1</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>4.46</td>
<td>67.51</td>
<td>35.37</td>
<td>C2</td>
<td>6.65</td>
<td>51.11</td>
<td>27.69</td>
<td>C2</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>3.57</td>
<td>7.87</td>
<td>11.13</td>
<td>C2</td>
<td>6.93</td>
<td>31.62</td>
<td>16.36</td>
<td>C3</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>8.63</td>
<td>41.52</td>
<td>15.09</td>
<td>C2</td>
<td>8.63</td>
<td>86.99</td>
<td>15.56</td>
<td>C2</td>
</tr>
<tr>
<td>5</td>
<td>Yes</td>
<td>7.40</td>
<td>18.15</td>
<td>10.07</td>
<td>C2</td>
<td>6.33</td>
<td>7.82</td>
<td>9.56</td>
<td>C2</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>5.43</td>
<td>24.97</td>
<td>9.42</td>
<td>C2</td>
<td>3.91</td>
<td>8.61</td>
<td>9.66</td>
<td>C1</td>
</tr>
<tr>
<td>7</td>
<td>Yes</td>
<td>4.96</td>
<td>8.70</td>
<td>2.79</td>
<td>C1</td>
<td>4.41</td>
<td>4.15</td>
<td>2.75</td>
<td>C1</td>
</tr>
<tr>
<td>8</td>
<td>No</td>
<td>4.15</td>
<td>9.80</td>
<td>5.74</td>
<td>C1</td>
<td>3.91</td>
<td>6.71</td>
<td>6.36</td>
<td>C1</td>
</tr>
<tr>
<td>9</td>
<td>Yes</td>
<td>5.04</td>
<td>9.69</td>
<td>3.63</td>
<td>C1</td>
<td>5</td>
<td>15.45</td>
<td>4.57</td>
<td>C1</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>3.24</td>
<td>2.34</td>
<td>1.58</td>
<td>C1</td>
<td>2.95</td>
<td>1.45</td>
<td>1.29</td>
<td>C1</td>
</tr>
<tr>
<td>11</td>
<td>Yes</td>
<td>5.95</td>
<td>10.40</td>
<td>18.12</td>
<td>C2</td>
<td>7.89</td>
<td>23.39</td>
<td>22.85</td>
<td>C2</td>
</tr>
<tr>
<td>12</td>
<td>Yes</td>
<td>6.43</td>
<td>32.90</td>
<td>23.14</td>
<td>C2</td>
<td>6.68</td>
<td>46.52</td>
<td>27.37</td>
<td>C1</td>
</tr>
<tr>
<td>13</td>
<td>Yes</td>
<td>3.02</td>
<td>2.78</td>
<td>2.18</td>
<td>C1</td>
<td>3.89</td>
<td>10.10</td>
<td>4.70</td>
<td>C1</td>
</tr>
</tbody>
</table>
Pseudotumor maximum diameter and volume data had a non-parametric distribution. Between both time points, median values for MD, EV and V were not significantly different (Table 10.3).

<table>
<thead>
<tr>
<th>Table 10.3, Median values at both MARS-MRI time points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MARS-MRI 1</strong> (median, min-max)</td>
</tr>
<tr>
<td><strong>MD (cm)</strong></td>
</tr>
<tr>
<td><strong>EV (ml)</strong></td>
</tr>
<tr>
<td><strong>V (ml)</strong></td>
</tr>
</tbody>
</table>

* Wilcoxon signed rank test

**Relationships between MD, EV and V**

MD, EV and V were all highly correlated, with the strongest correlation for EV and V (Table 10.4). EV was larger than V in 22 case, and smaller in 4 cases. On average, EV was 72.6% larger than V (range: -42.6% to 238.1%). Boxplots based on the data from the first MRI measurements show a distinct difference for MD, EV and V results per Anderson pseudotumor classification (Figure 10.5).

<table>
<thead>
<tr>
<th>Table 10.4, Correlations for MD, EV and V</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spearman correlation</strong></td>
</tr>
<tr>
<td>MD – EV</td>
</tr>
<tr>
<td>MD– V</td>
</tr>
<tr>
<td>EV – V</td>
</tr>
</tbody>
</table>
Figure 10.5, Boxplot of maximum diameter (A), estimated volume (B) and 3-D volume (C) versus Anderson classification.
**EV in ellipsoid versus non-ellipsoid shaped pseudotumors**

With 3-D ROI-based V results considered the reference value, the median percentage difference between EV and V for ellipsoid pseudotumors was 64.2% (range: -42.6 to 211.8), and 90.9% (range: -29.3 to 175.2) for non-ellipsoid pseudotumors.

**Comparison of Anderson score, MD, EV and V between first and second MARS-MRI.**

There were 10 cases with unchanged Anderson score, one case was upgraded from C2 to C3 and two cases downgraded from C2 to C1. For the upgraded case (Case 3), MD, EV and V had increased at the second MARS-MRI (MD: 3.57cm -> 6.93cm, EV: 7.87ml -> 31.62ml, V: 11.13ml -> 16.36ml). For one downgraded case (Case 12), MD, EV and V had also increased with the second MARS-MRI (MD: 6.43cm -> 6.68cm, EV: 32.90ml -> 46.52ml, V: 23.14ml -> 27.37ml). The other downgraded case (Case 6) had increased V but decreased MD and EV (MD: 5.43cm -> 3.91cm, EV: 24.97ml -> 8.61ml, V: 9.42ml -> 9.66ml). See table 10.2.

Median MD, EV and V for both time points are presented in table 10.3. There were no statistically significant differences in median values for either MD, EV or V. Median change in MD was 0.0cm (range: -1.5 to 3.4), median change in EV was -0.5ml (range: -16.4 to 45.5) and median change in V was 0.5ml (range: -7.7 to 5.2). Three pseudotumors had not changed in MD, in 5 cases MD was decreased and in 5 cases MD was increased (Figure 10.6). The median percentage difference in MD was -0.8% (range: -28% to +94. 1%). In 6 cases EV had increased, and 5 cases had decreased EV. The median percentage change in EV was of -20.6% (range: -65.5% to +301.8%). In 6 cases, V had increased and in 1 case V had decreased. The median percentage change in V was 3.1% (range: -21.7% to 115.6%). To detect change in pseudotumor dimension between two time points, there was no significantly different result if either of the three methods (MD, EV or V) was used (p = 0.74).
Figure 10.6, Scatterplot comparing MD1 with MD2

Figure 10.6A, Scatterplot comparing MD1 with MD2

Figure 10.6B, Scatterplot comparing EV1 with EV2

Figure 10.6C, Scatterplot comparing V1 with V2

Figure 10.6, Scatterplot comparing MD1 with MD2 (10.6A), EV1 with EV2 (10.6B) and V1 with V2 (10.6C).
Discussion

The clinical relevance of pseudotumors after MoM hip arthroplasty detected with cross-sectional imaging is still debated.\textsuperscript{3,19} However, several authors have published reports on extensive soft tissue damage due to pseudotumors after MoM hip arthroplasty,\textsuperscript{20,21} resulting in muscle necrosis, nerve compression and Deep Venous Thrombosis. Which aspects (i.e. volume or presence of solid masses rather than fluid-filled cysts) of pseudotumors are predictors for subsequent tissue lesions are not yet determined, although a common feature of these reports is that compression by the pseudotumor plays a role. To this end also a follow-up on how dimensions of conservatively managed pseudotumors develop through time, is of importance.

In a clinical setting, measuring pseudotumor dimensions on MRI slices is limited to either the maximum diameter (MD) in one plane, or an estimate of volume (EV) based on maximum diameter in three directions, representing a simple box model. This is an established method in oncology that is easy and fast to perform in a busy clinical practice.\textsuperscript{22} As expected, EV overestimated V more in non-ellipsoid pseudotumors (91%) than in ellipsoid pseudotumors (64%). From our results it is also clear that although there is a high correlation between EV and measuring volume with a three-dimensional region-of-interest (3D-ROI) per MRI slice method, the difference between these two methods can be substantial. In our study, EV overestimated pseudotumor volume by approximately 70%, which is something to consider in clinical practice when managing individual cases of pseudotumors.

Compared to MD, EV had a higher correlation with V and we therefore propose to use EV rather than MD to quantify pseudotumor dimensions in clinical practice. And although EV is partially dependent on MD, explaining the strong correlation between MD and EV, there are individual cases with a large MD who have only a relatively small volume (3D-ROI volume measurement). It is therefore evident that EV rather than MD is a more valid approach to assess the severity of pseudotumors, but clinical studies need to validate this approach. For assessing change in pseudotumor dimensions in sequential MARS-MRI scans, there was no significant difference between methods in the percentage change each method measured, although in clinical practice absolute changes might be more relevant than percentage change.
Although 3D-ROI volume measurements can be performed in a clinical setting using standard radiology software, they are very time consuming. This prohibits the use of this method for routine MARS-MRI evaluation of MoM patients. Estimated pseudotumor volumes were previously used in clinical studies.\textsuperscript{7,13} Hart et al used the maximal anterior-posterior, superior-inferior, and medial-lateral diameters to approximate a cuboid-base volume. Their median value was 25.1 ml with a range from 0.9 to 594.0 ml, which is considerably higher than our median value for EV (9.8 ml, range: 1.5 to 87 ml). This might be explained their study involving only a single MARS-MRI per case, and included painful MoM patients who were revised later on. In our study only small to moderate pseudotumors which were not considered for revision surgery after the first MARS-MRI were included, thereby excluding more extensive pseudotumors. Almousa et al used ultrasound imaging to report on estimated pseudotumor volume development in 15 unrevised and five revised patients. Similar to our findings, they reported both increased volume (n=6) and decreased volume (n=1) with repeated cross-sectional imaging, but they also observed pseudotumor disappearance (n=3) if left untreated.\textsuperscript{7} They considered three pseudotumors to have a clinical important increase in volume (46 ml, 56 ml and 134 ml respectively) but in contrast, we did not find such a change in pseudotumor dimensions between repeated MARS-MRI’s when using a more sophisticated 3-D ROI volume measurement. Using this method we observed a maximum increase of 5.2 ml, but with estimating volume using a similar approach as by Almousa et al, we also observed a 46 ml increase in one case. Another explanation might be by the difference in time between first and second cross-sectional imaging. In their study it was almost four times longer than in our study (25.8 months versus 7.5 months).

Our study was limited by the small number of included patients, and the inclusion of small to moderately sized pseudotumors only. Also, metal ion levels in all included patients were not elevated, and in theory pseudotumor development over time might be different in patients with elevated metal ion levels. This issue needs to be addressed in future studies. Our study is however the first to use 3D-ROI-based volume method for measuring pseudotumor dimensions after MoM hip arthroplasty. It also benefits from the use of MARS-MRI, allowing for more detailed imaging compared to ultrasound and less dependence on technician experience, although still manual input is required. This manual input might explain the four cases where V was smaller than EV. By definition, EV should be
equal or larger than $V$ since it uses a box model. MARS-MRI for MoM pseudotumor analysis also allows the use of a pseudotumor severity grading system, which incorporates more details of pseudotumors than only dimensions, such as pseudotumor wall thickness and pseudotumor content. A second issue would be to quantify the degree of solid mass of the pseudotumor.

In conclusion, estimating pseudotumor volume based on maximum diameter in three orthogonal planes is a feasible method in clinical practice and superior to measuring maximum diameter in one plane to evaluate (change in) pseudotumor dimensions. Clinicians however need to consider that this method overestimates pseudotumor volume in most cases, compared to more accurate 3D-ROI volume measurements.
References


MALLORY-HEAD UNCEMENTED TOTAL HIP PROSTHESIS
Chapter 11

General discussion
Introduction
Survival of hip arthroplasty in young patients is not as excellent as in elderly hip arthroplasty patients. This can be explained by the higher demands younger patients have. Larger and different biomechanical stresses are put to the prosthesis and a longer life of the prosthesis is needed due to patient’s longer life expectancy. These higher biomechanical stresses result in increased implant wear, and challenges the long term durability of artificial hip joints. Surgeons, biomedical engineers and scientists constantly strive to develop and improve designs and used materials that are better able to withstand these stresses and show reduced bearing surface wear characteristics. These materials are tested first in the laboratory using wear simulators and computerized models. When these materials are introduced in clinical practice, post market surveillance of their performance is limited and often left to individual surgeons. Increasingly however, national joint replacement registers report on the long term implant survival.

The constant drive to improve the clinical performance of hip prostheses has resulted in the re-introduction of Metal-on-Metal (MoM) implants during the late 1990s. Surgeons used this type of surface bearing previously between 1950 and 1970, but almost completely discontinued its use due to high failure rates. In the same period the successful concept of “low friction” arthroplasty was developed by sir John Charnley, using Ultra High Molecular Weight Polyethylene (UHMWPE) as a bearing surface. This became the most widely used bearing material. But wear rates for standard UHMWPE proved to be high in younger, more active patients, resulting in osteolysis, implant loosening and ultimately revision surgery. This high wear resulted in less favorable implant survival rates in these younger patients. This gave MoM bearings a second chance. With advanced production techniques allowing tighter material tolerances, these implants now were
believed to have reduced wear rates, as demonstrated in wear simulators. Wear rates were so much lower that the second generation of MoM bearings were introduced as “a life-time implant”, and were to be considered for the younger and more active patient. The issue of hip implant failure in younger patients is a concern of each surgeon. When a younger patient presents himself with severe clinical and radiological hip osteoarthritis and failed conservative treatment, hip surgery can be considered. In this case, the surgeon has to choose a bearing surface which will last as long as possible and leaves room for future revision surgery. When the research in this thesis was initiated, hip resurfacing became a popular option around the world. However, the clinical results achieved by this second generation MoM hip resurfacing were still under debate. We therefore started with prospectively collecting clinical outcomes and radiological data on our complete cohort of MoM hip resurfacing patients. Most chapters in this thesis present clinical results of implants using a MoM bearing surface, but we also report on implants using UHMWPE in age-matched patients with the MoM group, to have a baseline comparator. These results confirmed our assumption that survival in young patients is not as optimal as we expected. We found a high proportion (53.4%) of implants being above the accelerated wear threshold rate of >0.2 mm per year, after a mean follow-up of 8.3 years. Somewhat in contrast, implant survival at a maximum of 12 years was acceptable (Kaplan-Meier survival probability 90.1%), and just compliant to international guidelines such as the NICE criteria. To benchmark the results of our MoM hip resurfacing cohort, we systematically reviewed the peer-reviewed literature on the survival of these resurfacing implants. We found that aseptic loosening was the most frequent failure mode and that none of the contemporary hip resurfacing designs met the full 10 year NICE benchmark for survival. With the increasing attention in the international literature on the adverse reactions to metal debris (ARMD) in soft
tissue surrounding the MoM implant, we intensified our research on MoM resurfacing and focused the research more to the role of cross-sectional imaging in diagnosing these reactions. The results of several investigations on the role of cross-sectional imaging in detecting and grading of these adverse reactions are presented in the second part of this thesis and will be discussed in detail in this general discussion.

**Hip implant survival in younger patients using different bearing materials**

As described in chapter 2, long term hip joint replacement survival is often disappointing in younger patients and usually fails to meet the criteria of the National Institute for Clinical Excellence (NICE) for implant survival. Wear of the implant bearing surfaces is seen as one of the main failure reasons. This is the case for implant types that use standard UHMWPE as a bearing surface but, in retrospect, also for so called hard-on-hard bearings such as MoM and Ceramic-on-Ceramic (CoC). With the reintroduction of MoM bearings (both as THA and as hip resurfacing designs) during the 1990s, the main failure mode of the first generation MoM bearings was believed to be solved. The unacceptable high failure rate of the first generation MoM hip arthroplasty was mainly caused by short term aseptic implant loosening, due to high numbers of wear particles being released directly after implantation. The second generation MoM held the promise of low wear rates compared to standard UHMWPE and tighter production tolerances allowed the use of a thin acetabular shell with a large diameter femoral component, reducing the risk for dislocation. In case of MoM hip resurfacing, the preserved amount of bone stock compared to THA promised the benefit of easier future revision. These three promises of a longer lasting bearing surface combined with a reduced dislocation risk and easier future
revision surgery were tailored to the needs of younger and more active patient indicated for hip joint replacement surgery. 

During the introduction of MoM on the marked, the proposed benefits outweighed the concerns on metal ion debris released after implantation. Although these proposed advantages were tempting, the available evidence on MoM hip arthroplasty at that time was less than encouraging. In retrospect there is much debate about why large diameter MoM hip arthroplasty was (re)introduced around the millennium. With hindsight, marketing by orthopaedic device companies, media attention, internet and claiming patients can all be blamed for the introduction of MoM without the proper solid scientific evidence or a phased, controlled introduction in the market. Another big problem was the unavailability of MRI or CT that could deal with implanted metal implants. Scientific evidence was and still is conflicting regarding the benefits and complications associated with MoM arthroplasty. In 2000, Doorn, in his thesis on wear and biological aspects of MoM hip arthroplasty, concluded that wear volume was significantly less with MoM bearings compared to metal on polyethylene bearings, that less histocytic reactions occur with MoM bearings and that sensitivity and toxicity were not observed with MoM bearings. In 2011, Murray et al discussed possible risk factors for pseudotumor formation. Based on the argument that most of these risk factors could be avoided, they supported the continued use of resurfacing in appropriately selected patients by appropriately trained surgeons. However, several other authors were unable to confirm all these risk factors using data from their own case series. This conflicting evidence prompted us to study our MoM patients. In our prospective case series of 298 MoM hip resurfacings we found a six year survival rate of 92.7%. In comparison, we retrospectively found a 90.7% survival rate for implants using UHMWPE after 12 years. Held against the benchmark of a 90% survival rate after
10 years follow up as set by the NICE guideline, the first conclusion is that UHMWPE is compliant with this guideline, although only by a small margin, and the MoM implant is not meeting the 95% at five year landmark. This latter showed an insufficient follow up period of this particular MoM resurfacing device. In our systematic review of implant survival of MoM hip resurfacing devices, this finding was confirmed: none of the included MoM hip resurfacing designs met the NICE criteria. Moreover, at the time of review, there were no studies available on the particular hip resurfacing device used in our clinic. Later, a case series on this particular MoM hip resurfacing design was published by Gross. Although their survival rate, 96.4%, was better at 7 years, it was still not convincing. Another limiting factor of this study was that it was limited to clinical outcome scores and plain radiography only. This was comparable to our study first study on clinical follow-up of MoM resurfacing. In retrospect, clinical outcome scores and standard radiographs were insufficiently capable of detecting pseudotumors, as demonstrated in our pilot screening study using cross sectional imaging and confirmed after we screened our complete MoM hip resurfacing cohort using Metal-Artefact Reduction Settings (MARS) MRI. Applying MARS-MRI resulted in a 36.3% pseudotumor prevalence patients. These results were comparable with other cross-sectional imaging studies using different MoM designs, for example 28% for the Birmingham Hip Resurfacing (BHR), 33% for the Articular Surface Replacement (ASR) and 29% for the Durom design. The most severe cases were revised, adding a relatively new failure mechanism that negatively impacts implant survival of MoM hip implants.

Although survival rates with our implants which used UHMWPE were better than with our MoM hip resurfacing implants, the observed mean wear rate with standard UHMWPE was far from satisfactory. Further follow up of this particular case series should provide new data on whether this high wear rate will result in
an increased revision rate for osteolysis and implant loosening after the first decade. The few studies available on wear rates of UHMWPE with 10 to 20 years of follow up show that after so called “bedding in phase” during the first year, wear rates remain fairly stable up to around 8 to 10 years, but then increase again. The clinical relevance of this second decade of increased wear is not fully known, but a number of long term studies on the survival of the acetabular component report revision rates of 20% at 11 years\textsuperscript{13} up to 65% at 16 years.\textsuperscript{14} Future research should be directed towards constructing guidelines for implant survival in which the patients’ age at implantation is a consideration. Ideally, the implant survival in younger patients should not only be held against a 10 year benchmark but also against a 15 or 20 year benchmark, since the majority of younger patients will live more than 10 years after implantation. The Swedish hip register makes separation between different age categories, but NICE just uses 10 years as a benchmark.

During the last decade, more advanced UHMWPE materials have been developed to withstand wear and material fatigue. Clinical studies using cross-linked UHMWPE and second generation highly cross-linked UHMWPE are now published and compared to other bearings for wear performance and implant survival.\textsuperscript{15,16} Five to ten year clinical results of highly cross-linked UHMWPE reveal excellent clinical and wear results. Short term reports of vitamin infused highly cross-linked UHMWPE (developed to reduce material aging in highly cross-linked polyethylene in addition to wear resistance) are also encouraging. For now, we conclude that both standard UHMWPE and MoM bearings still not have succeeded in significantly improving implant survival in hip arthroplasty for younger patients. For MoM the unexpected occurrence of ARMD is the most important downside, for UHMWPE the high amount of wear with subsequent osteolysis and implant loosening.
**Surveillance for soft-tissue lesions after MoM hip arthroplasty**

The limited regulations for market introduction of hip implants have resulted in unforeseen problems. Currently there is attention to these deficits, but it needs to be seen if this is continued and applied to prevent future repetitions of this process, or if the orthopedic community, including surgeons, national boards and the medical device industry, turns its attention to a new design and forgets about the problems discussed in this thesis.

Fortunately, recent scientific publications have discussed how a more controlled introduction of joint replacement designs could be done, while balancing the protection of patients with the benefits of introducing new designs which might outperform current designs. With this reconsideration of how new joint replacement designs are introduced into the market, there is the possibility to define the role and responsibilities of all stakeholders involved.

The focus on improved implant introduction to the market should leave room for following up on the clinical results of currently used designs and on previously used but discontinued designs. Ongoing research on (discontinued) implant designs will benefit patients by making the optimal selection of revision implant designs and will learn us at what time point after the index surgery, revision surgery is best done if indicated in these cases. For example, if clinical results from early revisions for pseudotumor formation after MoM THA are worse than expected, surgeons should be more resistant to perform revision surgery.

To objectively study these issues, there are several needs. First there is a need for further development of pseudotumor classification systems and these systems should be more rigorously validated researching the consequent clinical actions based on the classification systems and other findings. These developments need to be incorporated into national guidelines to help clinicians treating their MoM patients. Secondly, more knowledge is needed on the development of
pseudotumors over time, their occurrence in non-MoM THA and which details of pseudotumors are predictive for the clinical outcome of conservative therapy and revision surgery. More research is also needed on the validation of imaging techniques like MRI. Can the circumstances under which conditions cross-sectional imaging was done, such as positioning of the patient, time of the day, etc., influence the results? In addition, not only in MoM hip arthroplasty, but also in non-MoM hip arthroplasty we need more insight in adverse soft tissue reactions incidences and consequences. Already, numerous case reports have been published on the occurrence of soft tissue masses near non-MoM total hip implants. So far, only very small observational studies have researched the occurrence of these adverse reactions in non-MoM hip arthroplasty, leaving the need for a larger study using cross-sectional imaging. It is also relevant for surgeons faced with a patient diagnosed with severe pseudotumor after MoM arthroplasty needing revision surgery, to have evidence on what bearing option to choose for the revision implant. Currently there is only a limited number of studies available presenting the clinical and radiological outcomes of MoM revision surgery for pseudotumor, all of them with only short follow up on a very limited number of cases. Different bearing options for MoM revision surgery are used such as large diameter ceramic-on-ceramic, dual mobility heads, or more standard THA using ceramic-on-polyethylene or metal-on-polyethylene.

**Introduction of hip implant designs into clinical practice**

The current questions around MoM implants, combined with issues like PIP (Poly Implant Prostheses) breast implants and failure of ICD implantable-cardioverter defibrillator (ICD) leads, have led to a global discussion on bringing medical devices to the market. Both the CE marking (Europe) and the IDE (US) process are
criticized.\textsuperscript{21,22} Currently a process in the European parliament is going on to change the CE marking legislation, however this is a quite complex process and it is questionable whether this will solve the current problems. Orthopaedic surgeons and biomedical engineers primarily question the use of specific implants from a performance perspective. Increasingly, national associations, medical insurance companies and hospital administrators also question the use of specific implants, often both from a performance and a costs perspective. All these stakeholders communicate with the medical device manufacturers that engineer and produce these implants, often in close collaboration with designer orthopaedic surgeons. In orthopaedic surgery, medical device companies have the infrastructure and knowledge for developing new orthopaedic devices, including laboratory testing. With the required testing standards (ASTM and ISO) and vigilance plan, the request for the CE mark is made. However, in contrast to pharmacy, the companies only need to present a vigilance plan since blinded, dose finding or placebo controlled studies are not possible. Dependency of post market surveillance is completely on orthopaedic surgeons who are the only ones that can apply these new techniques in the clinic. Many innovations or changes to the devices are made together and per request of the market (i.e. the surgeons).

Still, in comparison to pharmaceuticals which require multiple controlled clinical trials prior to approval, which take a mean of nine years and cost an average of 800 million U.S. dollars, medical devices such as a new hip implant design can be released onto the market after in vitro testing and limited supportive clinical data.\textsuperscript{23} To improve on this situation, several authors have advocated a stepwise clinical introduction of new implants. This involves pre-clinical testing, small prospective trials using high-precision methods such as Radiographic RadioStereometric Analysis (RSA) to assess initial fixation to predict long term
survival, larger multicentre trials and finally population-based register studies to keep devices on track.\textsuperscript{24,25} RSA studies limit the number of patients at risk while at the same time, with a short follow up period, provide sound predictions of long term implant performance.\textsuperscript{25} Uniform reporting of RSA and clear descriptions of the predicted migration pattern beforehand are essential to get high quality RSA results.

Surveillance of implant performance after introduction onto the market is done in a number of countries, but not all, by national joint replacement registries. For example in the United States, which is one of the largest markets, less than 200 of the 5724 registered hospitals participate in the American Joint Replacement Register (AJRR). Other countries such as Sweden however have a long history of nationally registering joint replacement procedures, with data entry compliance near 100%, enabling them to identify outliers in implant performance after market introduction. Still, a significant number of patients are put at risk before national joint registries can identify underperforming implant designs.

The re-introduction of second generations MoM bearings into clinical practice, which compromised the confidence of patients and professionals after reports on failing implants and even the recall of particular products, is now used to identify the shortcomings of the process governing the introduction of new THA implant designs in practice.\textsuperscript{25,26} There are however many considerations. For example RSA, a key element in the stepwise introduction of new implants, is a predictor for survival. The current issues raised with the MoM bearings could not have been prevented with RSA. For example, the RSA results of the Recap MoM resurfacing (the prosthesis described in this thesis) were excellent.\textsuperscript{27} It is therefore necessary that a balanced introduction will not only rely on clinical data of implant fixation, but that also both local tissue reactions and systemic reactions to released wear particles are monitored.
This phased introduction should strike a balance between optimizing patient safety while at the same time allowing maximum technology development. New testing protocols have been developed, in which THA surface bearings are tested in adverse conditions such as non-optimal mechanical placement, oxidative stress and more extreme temperatures. This should increase the validity of these test results for performance in clinical practice. In Europe, medical devices are allowed onto the marked after CE (Conformité Européenne) approval but since the number of medical devices regulated by CE marking is approximately 500,000, ranging from scoot mobiles and drapes to artificial joints or heart valves, it is extremely difficult to design specific guidelines for each medical device in Europe. Further from notified bodies cannot be expected to be experts on all devices and materials. They rely on the quality of the presented documents.

Benchmark criteria on THA implant survival are nowadays used to evaluate implant performance. The National Institute for Health and Care Excellence (NICE) criteria set a rate of revision for failure of 10% or less for a given prosthesis at 10 years.\textsuperscript{1} These guidelines however do not take into account factors such as indication or age. With the increasing number of patients who have received a THA, national benchmark guidelines should consider extending their criteria beyond the first decade.

A recent study by Anand demonstrated that the level of implant performance in modern hip arthroplasty is hard to beat: none of the new implant designs outperformed current hip implant designs, most even did worse.\textsuperscript{28} But one has to be careful to interpret these findings in such a way that development of new materials and designs is not halted. Moreover, there are examples of implant designs of which the use was discontinued after initial reports predicted poor long term performance. For example, based on RSA studies, the SHP stem was predicted to have poor long term performance but recent clinical studies showed
equal implant survival compared to well-established implants. An even more complicated discussion is raised on the issue of implant design changes. During the lifecycle of a certain implant, there might be one or more modifications to the original design. All these changes are made to further improve products to meet market demands. Although these minimal changes can have major consequences, it is difficult to say what research is needed to back up these changes. It might be useful to copy the automotive industry in this matter, where small changes to a certain model results in the addition of ‘Mark 2’ or ‘Mark 3’ suffix to the model name. This would allow surgeons to better judge the available evidence for certain implant designs and their alterations.

**Future research on hip joint replacement performance**

Finally, the most difficult consideration in the management of problematic MoM patients are the patients’ experiences and preferences. There are patients who want a revision but with no or mild symptoms, normal or slightly elevated metal ion levels and no pseudotumor visible with imaging techniques. Other patients are very hesitant to revision surgery but have large, pseudotumors visible on CT or MRI but are without symptoms or elevated metal ion levels. The phased introduction of new orthopedic implant designs should prevent future recurrence of these dilemma’s. Not only RSA studies on implant fixation should be a part of such a balanced introduction, but also local tissue reactions should also be monitored with either ultrasound, CT or MARS-MRI, while possible systemic reactions to released wear particles should be studied with blood analyses. Ultimately, these preliminary findings should be validated with strong clinical research results, thereby ensuring long term patient safety and guiding the orthopaedic community in which implant types to use for best results in the younger, more active patients.
References


METAL-ON-METAL BIRMINGHAM HIP RESURFACING
Chapter 12
Summary
Long term durability of hip replacement implants is mainly limited by wear of the bearing surfaces between the femoral and acetabular components. Different bearing materials have been used with the aim to reduce wear and prolong implant survival. Polyethylene (PE), commonly used as a bearing surface on the acetabular side, releases wear particles which induce osteolysis with subsequent component loosening and ultimately implant failure. In the constant strive to improve on implant design and materials, a second generation of Metal-on-Metal (MoM)surface bearings was introduced in the 1990s with the promise of reduced wear, thereby supposedly improving long term implant survival. The main support for this claim was achieved by in vitro testing, using wear simulators which were run under ideal conditions.

This re-introduction, after the use of the first generation MoM was discontinued due to unacceptable high failure rates, took place in the context of limited requirements on supportive clinical data to release new designs into the orthopedic market. After approximately a million MoM hip implants were inserted into patients (both total hip arthroplasty (THA) and resurfacing procedures), it became clear that unexpected complications occurred in soft tissue surrounding this MoM implant. These were due to metal debris released from the bearing surfaces. This eventually led in 2010 to a worldwide recall of one of the MoM hip implant designs, followed by a ban on the use of MoM large diameter hip implants in several countries. Since these reactions were unforeseen, no evidence based guidelines on how to diagnose and to treat these complications were available. Since there is no final data or consensus on the risk-benefit ratio for the use of MoM implants, the use of large diameter MoM hip implants is currently banned in some countries and still in use in many other countries.

Chapter 1
The first chapter introduces current issues raised with implant survival and bearing surfaces in THA and in Hip Resurfacing Arthroplasty (HRA). The four main aims of this thesis are presented, being (1) To report the implant survival of the current THA bearing options seen as gold standard for the young and more active patients. (2) To review all available literature on the different resurfacing systems (3) To investigate complications after MoM due to soft tissue reactions to metal
wear debris, and (4) To study the most used diagnostic tool, MRI, and the classification systems used to find and rate these complications.

Chapter 2
In chapter two we reviewed the long history of MoM bearing surfaces in hip arthroplasty. Since younger patients tend to be physically more active than elderly patients, their implants have to withstand higher biomechanical stress and these stresses also need to be endured for a more prolonged period of time, leading to accelerated wear of the bearing surfaces. To reduce bearing surface wear, surgeons, engineers and scientists have developed different bearing surfaces. For this purpose, Metal-on-Metal (MoM) surface bearings have a long tradition in THA. The re-introduction of the second generation MoM in the 1990s took place after the first generation of MoM was abandoned due to unacceptable high failure rates and as an answer to “polyethylene disease”, occurring with standard Ultra High Molecular Weight Polyethylene (UHMWPE) bearings. Wear simulation tests of second generation MoM bearings showed that wear rates were 20 to 100 times lower compared to metal-on-conventional polyethylene, and MoM bearing couples started to experience widespread clinical use in both hip resurfacing and total hip arthroplasty. The material properties allowed the use of large heads in thin acetabular shells, promising a reduced incidence of hip dislocation in younger and more active patients.

Despite the biomechanical advantages of MoM bearings, metal ion release over time and the potential detrimental effects of accumulated metal ions in the body remained a concern, and research started to identify implant failure modes in reaction to metal wear particles. The terms ALVAL [2005], pseudotumor [2008] and metallosis were used, together with a new umbrella term for these modes of failure: Adverse Reaction to Metal Debris (“ARMD”, 2010). These unforeseen complications revealed serious shortcomings in how orthopaedic innovations are introduced into clinical practice. The conflicting interests of making promising new hip implant materials and designs available so patients can benefit as soon as possible, and the fact that these same joint replacement devices have to perform well for over more than 10 years and preferably more than 20 years after implantation in the patient, make it difficult to design a model for market introduction that effectively and safely guards all these requirements. In comparison to pharmaceuticals which require multiple controlled clinical trials
prior to approval, which take a mean of nine years and cost an average of 800 million U.S. dollars, medical devices such as a new hip implant design can be released onto the market after in vitro testing and very limited clinical trials. As witnessed with the re-introduction of MoM bearings in THA, serious complications which were unforeseen at the time of introduction became only became known after a large number of patients (worldwide an estimated one million patients) had become at risk.

Chapter 3
Before MoM hip arthroplasty became available for clinical use in the Netherlands, uncemented THA with standard UHMWPE was the gold standard for younger patients. With this prosthesis design, PE wear remained an important clinical observation and to evaluate implant performance, we retrospectively measured radiographic wear and implant survival of the first 200 consecutive uncemented hip arthroplasties with standard UHMWPE used in our clinic. In this series we found a high proportion (53.4%) of implants with a wear rate of >0.2 mm per year, which is considered a threshold for accelerated wear. This was after a mean follow-up of 8.3 years. Somewhat in contrast, implant survival at a maximum of 12 years was acceptable (Kaplan-Meier survival probability 90.1%), and compliant to international guidelines such as the NICE criteria.

Chapter 4
With the re-introduction of MoM arthroplasty, all major orthopedic device manufacturers designed and introduced, sometimes slightly, different hip resurfacing implants. Individual studies using different resurfacing designs reported marked differences in short term implant performance, so we decided to systematically review the peer-reviewed literature on implant survival of all contemporary MoM hybrid hip resurfacing designs. A total of 29 studies, compromising 10,621 patients, were included. All but one of the implants studied had insufficient follow up to be compliant with the NICE benchmark, of a revision rate of less than 10% at ten years, for choosing a prosthesis for primary THR. The study reporting a follow-up of longer than ten years had a revision rate of 16%, mainly for aseptic loosening of the implant. This high failure rate was attributed to the double-heat-treatment manufacturing process which is no longer in use. The prosthesis was superseded by the Cormet 2000 implant in 1996. Compared with
the three-year NICE entry-benchmark of implant survival ≥97%, 13 studies (44.8%) showed satisfactory survival. Eight used the BHR implant, two the Conserve plus, one the Durom, one the Cormet 2000 and one both the McMinn and BHR implants. Based on the results of this review we concluded that aseptic loosening was the most frequent failure mode and that none of the contemporary hip resurfacing designs met the full 10 year NICE benchmark for survival. In this systematic review we were unable to include studies on the resurfacing implant used in our clinic.

Chapter 5
In this chapter we present the results of data prospectively collected in a series of 280 consecutive hip resurfacing procedures (ReCap, Biomet, Warsaw, USA) performed in our clinic. Mean follow up was 3.3 years (range: 1.0 to 6.3) and four patients were lost to follow-up. All patients were diagnosed with end-stage hip osteoarthritis, their mean age was 54 years and 76.4% of all patients were male. All were evaluated with standard radiographic imaging and clinical outcome scores before surgery and yearly after the index surgery. There were 16 revisions and four patients reported a Harris Hip Score <70 points at their latest follow up. Kaplan-Meier implant survival probability, with revision for any reason as endpoint, was 93.5% at six years follow-up (95%-CI: 88.8-95.3). There were no revisions for Adverse Reactions to Metal Debris (ARMD) and no indications of ARMD in symptomatic non-revised patients, although diagnostics were limited to ultrasound scans. We concluded that hip resurfacing is a demanding procedure, and that implant survival of the ReCap hip resurfacing system is on a critical level in our series. However, in non-revised patients, reported outcomes are generally excellent.

Chapter 6
In chapter six we presented the results of a pilot study in which we used an intensified screening protocol to detect pseudotumor formation after MoM hip resurfacing in three selected groups of patients: a group with a theoretically high risk for pseudotumor formation, a group with a very low risk for pseudotumor formation and a group scheduled for routine follow up with a mix of risk factors present. Risk factors were based on component size and orientation, gender, bilateral or unilateral MoM surgery and clinical symptoms. All selected patients
underwent blood metal ion level analysis and cross-sectional imaging using MARS-MRI. In this study we used a pseudotumor classification system devised by Anderson et al to grade pseudotumor severity. Pseudotumor formation was observed in all three groups, even in asymptomatic patients with normal blood metal ion levels. In 15 out of 44 MRI scans pseudotumors were observed (34.1%), of which six were graded with mild (13.6%), eight with moderate (18.2%) and one with severe MoM disease (2.3%). Twelve pseudotumors were present in asymptomatic patients (27.3%). Metal ion levels were normal in 80% of the MARS-MRI screened patients. As a consequence to our intensified screening protocol, one patient was revised for pseudotumor occurrence and another patient scheduled for revision. Asymptomatic pseudotumors were observed in all three groups. We concluded that clinical outcomes and plain radiographs for screening MoM patients severely underestimated the presence of pseudotumors in MoM patients.

Chapter 7
Different pseudotumor grading systems had been described in the scientific literature, but no studies had compared these different systems for use in clinical practice and only limited data on the reliability of these grading systems was available. In chapter 7 we investigated the influence of using these different pseudotumor grading systems on how severe pseudotumors were classified. For this study we evaluated a cohort of 42 THA patients (49 MoM hips) using three different pseudotumor grading systems designed respectively by Anderson et al, by Matthies et al and by Hauptfleisch et al. Two experienced musculoskeletal radiologists evaluated all MARS-MRI scans with these systems, allowing us to calculate the interobserver reliability for each system. Our results showed that, regardless of the classification system used, grading pseudotumor severity on MARS-MRI had only a moderate interobserver reliability (ICC 0.65 to 0.68). The reliability of pseudotumor severity grading was high between the Matthies and Hauptfleisch system but low between the Anderson and the other two systems. We concluded that a more succinct pseudotumor severity grading system is needed for clinical use.
Chapter 8
Since we demonstrated in chapter 6 that standard radiographic follow up combined with clinical outcomes was not sensitive enough to detect pseudotumor formation after MoM hip arthroplasty, we extended our intensitied screening protocol to our complete cohort of MoM hip resurfacing patients. This study is presented in chapter 8. At the time this study was started, 248 MoM hip resurfacing procedures (214 patients, mean follow-up 4.6 years, range: 1 to 8.2) were available for follow up. Again the Anderson classification for pseudotumors was used. We found a pseudotumor prevalence of 36.3%: 61 pseudotumors were graded mild, 25 moderate and four were graded severe. Five revisions followed, all in symptomatic patients with elevated metal ion levels. Since the natural course of pseudotumors is largely unknown, and no validated treatment regime for pseudotumors after MoM hip arthroplasty exists, we suggested to repeat MARS-MRI in asymptomatic patients with mild to moderately severe pseudotumors combined with normal metal ion levels, rather than to immediately revise these cases. The use of this screening protocol and this pseudotumor grading system allowed us to be conservative with revision surgery for mild and moderate MoM disease. Of course patients with non-revised pseudotumors were kept under increased surveillance. These results could be used as a clinical guideline for management of observed pseudotumor after MoM hip resurfacing.

Chapter 9
As stated in the previous chapter, intensified follow up of cases with non-revised pseudotumors was needed to validate a more conservative approach in the management of observed pseudotumors. In chapter 9 we present the results of repeated MARS-MRI’s which were used to follow up on identified pseudotumors from our previous studies. To monitor how pseudotumors developed in time, we repeated cross-sectional imaging 6 to 12 months after the initial MARS-MRI, together with repeated metal ion analysis and clinical examination. In this study, 14 unrevised cases with pseudotumour and a control group of 23 cases without pseudotumour on the first MARS-MRI were evaluated. The mean postoperative time to the first MARS-MRI was 4.3 years (range: 2.2 to 8.3) and mean time between first and second MARS-MRI was 8 months (range: 6 to 12). The majority of patients (35/37) showed no change in pseudotumor severity with the second MRI, one new pseudotumour was observed (Anderson C2 score, moderate) and
one pseudotumour was downgraded from C2 (moderate) to C1 (mild). We concluded that repeating of MARS-MRI within one year, in unrevised patients with asymptomatic pseudotumours after MoM hip resurfacing, was of limited use. But, since this was the first longitudinal study on pseudotumours using MARS-MRI, our findings need to be interpreted with caution.

Chapter 10
Since management of non-revised pseudotumors depends on both severity (based on location, type of content, growth rate) and on pseudotumor dimensions, it is relevant to have an accurate clinical measurement method of pseudotumor dimensions. In this chapter our objective was to validate clinical measurements of (change in) pseudotumor dimensions (maximum diameter and estimated volume) against three-dimensional region-of-interest (3-D ROI) volume measurements. Therefore, we had MARS-MRI scans available for 13 cases of non-revised pseudotumors after Metal-on-Metal hip resurfacing. Mean follow-up at the first MARS-MRI was 5.3 years (range: 2.4 to 7.5), a second MARS-MRI was acquired after a mean of 7.5 months (range: 6 to 12). On all scans pseudotumor dimensions were measured by (1) maximum diameter in one plane (MD) and (2) by estimating pseudotumor volume based on maximum diameter in three different planes (EV). (3) For validation, a 3-D ROI based volume (V) was calculated by the summation of all pseudotumor areas in each slice and multiplication by the slice profile. Correlations between MD, EV and V were calculated. Correlation was high between all three measurement methods, but the correlation was strongest between EV and V. EV overestimated V with a mean of 72.6%, and more so in non-ellipsoid pseudotumors than in ellipsoid pseudotumors. Median values for MD, EV or V were not significantly different between first and second MARS-MRI. Median change for MD was 0.0cm (range: -1.5 to 3.4), -0.5ml for EV (range: -16.4 to 45.5) and 0.5ml for V (range: -7.7 to 5.2). Percent change in pseudotumor dimensions was not significantly different between MD, EV and V.

We concluded that estimating pseudotumor volume in clinical practice using maximum pseudotumor diameter in three different planes has a strong correlation with a more elaborate 3D-ROI method. This method of estimating volume is easily attainable in clinical practice and can be used for monitoring change in pseudotumor volume over time.
Chapter 11

In this chapter, the findings of all studies conducted for this thesis are synthesized and discussed in their context, resulting in answers to the main study aims and propositions for future research. The first aim of this thesis was to report the implant survival of the current THA bearing options seen as gold standard for the young and more active patients. We concluded that for these patients, hip resurfacing with MoM bearing surfaces was not compliant with the international benchmarks for 10 year implant survival and uncemented, standard UHMWPE hip prostheses just barely reached this benchmark. For the UHMWPE prostheses, the high amount of wear was noticed as the biggest downside with a potential accelerated wear in the second decade after implantation.

The second aim of this thesis was to review all the different resurfacing systems on the market for implant survival results. After systematically reviewing the literature, we concluded that all reviewed hip resurfacing systems did not meet the international benchmark, and that there were hip resurfacing systems on the market of which no clinical studies were available for our review. It is noteworthy that the data presented in the studies we reviewed were collected before the unexpected adverse reactions to metal debris released by the MoM bearing surfaces were investigated. Therefore our conclusion that aseptic loosening was the main failure mode of MoM hip resurfacings needs to be seen in this context, and a future update of this review based on current knowledge might change the view on failure of reasons for current MoM systems.

The third aim of this thesis was to investigate complications after MoM due to soft tissue reactions to metal wear debris. We concluded that the incidence of these complications, diagnosed as pseudotumors, was higher than expected, that risk factors were difficult to interpret, and that cross-sectional imaging is necessary to find the true incidence, since many patients have these reactions without being symptomatic. We also found that the use of a pseudotumor classification system was helpful in managing treatment of these complications.

The fourth aim of this thesis was to study the most used diagnostic tool, MRI, and the classification systems used to find and rate these complications. Based on our validation studies we concluded that using these systems observers were able to identify pseudotumors but that the classification of pseudotumors severity needed more refinement. Therefore future studies need to validate the treatment which was chosen upon the pseudotumor severity grade that was seen with
MARS-MRI. For clinical practice, we found that a simple box model for estimating pseudotumor volume correlated well with a more elaborate three-dimensional region-of-interest system and was easily used in clinical practice, although clinicians using this method have to take some overestimation of pseudotumor volume into account.
CERAMIC-ON-POLYETHELENE DUAL MOBILITY HIP PROSTHESIS
Chapter 13

Samenvatting (summary in Dutch)
De levensduur van heupprotheses wordt beperkt door slijtage van de contactoppervlakken van de femorale en de acetabulaire component. Er zijn dan ook verschillende materialen toegepast om deze slijtage te minimaliseren, en daarmee de overleving van het implantaat te verlengen. Polyethyleen (PE) is het meest gebruikte materiaal als oppervlak aan de acetabulaire zijde, maar de slijtage deeltjes die daarvan vrijkomen kunnen osteolyse induteren, waardoor uiteindelijk de prothese los laat en daarmee het implantaat faalt. In het voortdurend streven naar verbeteringen in het ontwerp en de duurzaamheid van de gebruikte materialen, werd rond 1990 een tweede generatie van Metaal-op-Metaal (MoM) implantaten geïntroduceerd, met daarbij de belofte van minder slijtage en een verbeterde lange termijn overleving van het implantaat. De belangrijkste onderbouwing hiervoor werd geleverd door de resultaten van in-vitro testen, uitgevoerd met behulp van slijtage-simulatoren. Deze herintroductie van MoM protheses vond plaats nadat het gebruik van de eerste generatie MoM protheses was gestaakt na onaanvaardbaar hoge revisiepercentages, en waarbij de regelgeving maar beperkt eisen stelde aan de introductie van nieuwe prothese-ontwerpen. Nadat van deze tweede generatie MoM protheses wereldwijd ongeveer een miljoen operaties waren gedaan (zowel in de vorm van een totale heupprothese (THP), als in de vorm van een resurfacing prothese), werd het langzaam duidelijk dat er onverwachte complicaties konden optreden in de weefsels rondom deze implantaten. Deze reacties waren het gevolg van vrijkomende slijtagedeeltjes van de metalen oppervlakken. Deze reacties bleken voor specifieke MoM heupprotheses zo vaak voor te komen, dat in 2010 een wereldwijde terugroep actie (‘recall’) werd uitgevaardigd voor een van deze MoM heup implantaten. Dit werd later gevolgd door een verbod in verschillende landen, waaronder Nederland, op het gebruik van MoM heupprotheses met een grote diameter, ongeacht het merk van de prothese. Aangezien deze reacties onvoorzien waren, waren er geen richtlijnen over hoe het beste deze afwijkingen te diagnosticeren en te behandelen. Nog steeds is er geen consensus over de verhouding tussen het risico van een MoM prothese en de mogelijke voordelen, waardoor het gebruik van grote diameter MoM heup implantaten momenteel in enkele landen verboden is terwijl dit type prothese nog wordt gebruikt in andere landen.
Hoofdstuk 1
In het eerste hoofdstuk wordt de toepassing van heupprotheses bij relatief jonge patiënten met ernstige arthrose van de heup besproken, met als belangrijkste aandachtspunt de overleving van het implantaat bij deze patiënten. Doordat het implantaat bij deze jongere patiënt zwaarder wordt belast dan bij de veel oudere patiënt, is de kans op slijtage van de gewrichtsoppervlakken groter. Daarnaast moet de prothese bij deze relatief jonge patiënten, vanwege hun langere levensverwachting, ook langer in situ kunnen blijven. Door deze zwaardere eisen die hiermee worden gesteld aan de prothese, is er veel aandacht voor de slijtage van de gewrichtsoppervlakken van deze protheses. Het gebrek aan onderbouwing van welk type prothese voor deze relatief jonge patiënten het meest geschikt zou zijn, leidde dan ook tot de vier belangrijkste doelstellingen van dit proefschrift, namelijk; (1) Te analyseren wat de slijtage en overleving was van de ongecementeerde THP, die voor het gebruik van MoM als de gouden standaard werd gezien voor de jonge, actieve patiënten met invaliderende heuparthrose; (2) De beschikbare wetenschappelijke literatuur over de resultaten van MoM heup resurfacing implantaten systematisch te analyseren; (3) Het optreden van complicaties in de weke delen rondom de MoM prothese te bestuderen, en (4) de toepassing en resultaten van ‘Metal-Artefact Reducing Sequence - Magnetic Resonance Imaging’ (MARS-MRI) als diagnostische instrument voor deze complicaties na het plaatsen van een MoM heupprothese te bestuderen.

Hoofdstuk 2
In hoofdstuk twee wordt de geschiedenis van MoM heupprothese besproken. Aangezien jongere patiënten in het algemeen fysiek actiever zijn dan de meeste oudere patiënten, moeten hun prothese gedurende langere tijd meer biomechanische stress doorstaan, waardoor er versnelde slijtage van de gewrichtsoppervlakken kan optreden. Om deze slijtage te verminderen, hebben orthopeden, ingenieurs en wetenschappers verschillende kunstmatige gewrichtsoppervlakken ontwikkeld, waaronder MoM. De introductie van de tweede generatie MoM in de jaren 1990 vond plaats nadat het gebruik van de eerste generatie MoM was gestaakt vanwege onaanvaardbaar hoge revisiepercentages, en als antwoord op de zogenaamde “polyethyleen ziekte”, die op kon treden bij het gebruik van standaard Ultra Hoog Moleculair Polyethyleen (UHMWPE) gewrichtsoppervlakken. Na laboratorium tests van de tweede
generatie MoM gewrichtsoppervlakken bleek de slijtage 20 tot 100 keer minder te zijn dan bij UHMWPE. Vanaf 2000 nam het gebruik van MoM wereldwijd sterk toe. Dit gebeurde zowel als heup resurfacing en als totale heupvervanging. Door de materiaaleigenschappen kon ook een grote diameter heupkop worden gebruikt, waardoor in theorie de kans op heupluxatie sterk afnam, hetgeen dit type prothese met name geschikt maakte voor de jongere, actieve patiënt. Ondanks deze biomechanische voordelen van MoM gewrichtsoppervlakken, bleef het vrijkomen van metaalionen na implantatie en de mogelijke schadelijke effecten hiervan een punt van zorg. Geleidelijk aan werden meer wetenschappelijke resultaten gepubliceerd die reacties op vrijgekomen metaaldeeltjes beschreven. De termen ALVAL (2005), pseudotumor (2008) en metallosis werden hiervoor geïntroduceerd, met daarbij een nieuwe, overkoepelende term: “Adverse Reactions to Metal Debris” (ARMD, 2010). Deze onvoorziene complicaties toonden tevens aan dat er tekortkomingen waren in de wijze waarop orthopedische innovaties werden geïntroduceerd in de klinische praktijk. De tegenstrijdige belangen bij dergelijke introducties van een nieuw ontwerp heupprothese zijn evident: het introduceren van betere ontwerpen en materialen laat patiënten direct profiteren van deze innovaties; Daarentegen moet ook op langer termijn worden aangetoond dat nieuwe implantaten goed presteren. Bij voorkeur langer dan 10 jaar en liefst zelfs meer dan 20 jaar na implantatie. Deze tegenstrijdigheid maakt het moeilijk om een model voor marktintroductie te hanteren dat tegelijkertijd innovaties zo snel mogelijk toelaat tot de klinische praktijk om zoveel mogelijk patiënten te laten profiteren, waarbij tegelijkertijd de veiligheid van de patiënten optimaal wordt bewaakt. In vergelijking met de introductie van nieuwe geneesmiddelen, waarbij een strikte regelgeving met meerdere gecontroleerde klinische studies vóór toelating tot de markt er toe leidt dat dit proces gemiddeld negen jaar duurt en $800.000.000 kost kunnen medische hulpmiddelen, zoals een nieuwe heupprothese, worden vrijgegeven in de klinische markt na een beperkt aantal klinische trials, vaak van geringe omvang. Met de herintroductie van MoM heupprotheses konden ernstige complicaties ontstaan die op het moment van introductie niet waren voorzien, waardoor een groot aantal patiënten, wereldwijd naar schatting 1 miljoen, ‘at risk’ kwam voor deze complicatie.
Hoofdstuk 3
Voordat MoM heupprothese werd geïntroduceerd in de kliniek was de gecementeerde THA met UHMWPE de “gouden standaard” voor jongere patiënten met invaliderende heuparthrose. PE slijtage was hierbij de belangrijke klinische observatie in lange termijn studies. Om deze resultaten te evalueren in de eigen kliniek, hebben we retrospectief onderzocht wat de radiologische slijtage en de implantaatoverleving was van de eerste 200 achtereenvolgende ongecementeerde heuprothesen met standaard UHMWPE geplaatst in onze kliniek. In deze serie vonden we dat 53% van de geplaatste protheses een versnelde slijtage (>0,2 mm per jaar) liet zien, bij een gemiddelde “follow-up” van 8.3 jaar. Enigszins in tegenstelling tot deze resultaten bleek de overleving van dit type implantaat acceptabel te zijn volgens de internationale richtlijnen van de (National Institute for Clinical Excellence) NICE- criteria, zelfs bij de maximale “follow-up” duur van 12 jaar (Kaplan-Meier overlevingskans 90.1%)

Hoofdstuk 4
Met de herintroductie van MoM heupresurfacing protheses, brachten alle grote orthopedische fabrikanten een eigen variant hiervan op de markt. In wetenschappelijk onderzoek naar de resultaten van de verschillende ontwerpen, leek het er op dat er tussen de verschillende merken verschillen waren op te merken in de korte termijn overlevings resultaten. Hierop besloten we om een systematisch literatuuronderzoek te doen naar de “peer-reviewed” literatuur over implantaatoverleving van alle hedendaagse MoM heupresurfacings. In totaal includeerden we 29 studies, die gezamenlijk de resultaten van 10.621 patiënten beschreven. Alle merken heupresurfacing, op één na, bleken onvoldoende lange termijn resultaten te presenteren om te kunnen voldoen aan het NICE criterium (maximaal 10% revisie in 10 jaar). Het merk heupresurfacing met wel een “follow-up” van meer dan tien jaar had een revisiepercentage van 16%, hoofdzakelijk door aseptische loslating van het implantaat. Dit hoge percentage werd toegeschreven aan het specifieke productieproces (dubbele warmtebehandeling) en deze werkwijze werd dan ook niet meer toegepast. Naast de 10 jaar benchmark kent NICE ook nog de 3-jaar benchmark. Vergeleken met dit criterium (maximaal 3% revisie na drie jaar), bleken er 13 studies (44.8 %) een positief resultaat te halen: Acht gebruikten het BHR implantaat, twee de Conserve plus, een de Durom, een de Cormet 2000 en een zowel het McMinn als het BHR implantaat.
Op basis van deze resultaten konden we concluderen dat geen van de huidige heup resurfacing ontwerpen voldeed aan het 10 jaar NICE criterium, en de minderheid voldeed aan het 3 jaar criterium. Aseptische loslating van de componenten was daar de meest frequente reden van falen Daarnaast konden we met deze systematische review geen studies includeren die rapporteerden over het specifieke resurfacing implantaat dat in onze kliniek werd gebruikt.

Hoofdstuk 5
In hoofdstuk vijf presenteren we de resultaten van prospectief verzamelde data van een reeks van 280 achtereenvolgende heupresurfacing operaties in onze eigen kliniek, waarbij in alle procedures het ReCap heup resurfacing systeem (Biomet, Warsaw, USA) was gebruikt. De gemiddelde “follow-up” was 3.3 jaar (range: 1.0 tot 6.3) waarbij vier patiënten “lost to follow-up” waren. Alle patiënten waren voor de operatie gediagnosticeerd met ernstige heuparthrose (gemiddelde leeftijd: 54 jaar, 76.4% mannen). Alle patiënten waren voor de operatie onderzocht met standaard radiologische onderzoek en de afname van gevalideerde klinische meetinstrumenten, en dit werd jaarlijks herhaald na operatie. Uiteindelijk waren er 16 revisies. Vier ongereviseerde patiënten scoorden slecht op de klinische uitkomstmaten tijdens hun laatste follow up (o.a. Harris Hip Score <70 punten). Dit resulteerde in een Kaplan-Meier overlevingskans voor het implantaat, met revisie om welke reden dan ook als eindpunt, van 93.5% na zes jaar “follow-up” (95%-BI: 88.8-95.3). Ten tijde van dit onderzoek werd er geen ARMD geconstateerd, waarbij moet worden opgemerkt dat aanvullende diagnostiek voor deze problematiek bij symptomatische patiënten beperkt bleef tot echografie. De conclusie van dit onderzoek was dat heupresurfacing een veeleisende procedure is, en dat implantaatoverleving van de ReCap heup resurfacing in onze kliniek op een kritisch niveau bleek te zijn. Echter, in niet gereviseerde patiënten bleken de gerapporteerde resultaten veelal uitstekend te zijn.

Hoofdstuk 6
In hoofdstuk zes presenteren we de resultaten van een pilotstudie met een intensief screening protocol naar pseudotumorvorming als gevolg van een MoM heupresurfacing. Daarbij vergeleken we de resultaten van drie groepen: (1) patiënten met een theoretisch hoog risico op pseudotumorvorming, (2) patiënten
met een zeer laag risico en (3) patiënten die voor routine “follow-up” in de kliniek kwamen, met daardoor een mix van risicofactoren. Risicofactoren waren gebaseerd op positionering en grootte van de gebruikte componenten, geslacht van de patiënt, het uni of bilateraal hebben van een MoM heupprothese en klinische symptomen. Alle deelnemende patiënten kregen bloedonderzoek om de concentratie metaalionen te bepalen en een Metal-Artefact Reducing Sequence (MARS)-MRI. In dit onderzoek gebruikten we een pseudotumor classificatiesysteem dat door Anderson et al was beschreven waarmee de ernst van de pseudotumor kon worden vastgesteld. Tot onze verrassing werden in alle drie de groepen pseudotumoren waargenomen, ook bij asymptomatische patiënten met normale concentraties metaalionen en in het bloed. In 15 van de 44 MRI-scans werden pseudotumoren waargenomen (34.1%), waarvan er zes werden beoordeeld als niet ernstig (13.6%), acht als matig ernstig (18.2%) en één als ernstig (2.3%). Van de waargenomen pseudotumoren waren er 12 aanwezig bij asymptomatische patiënten (27.3%) en deze “asymptomatische pseudotumoren” werden waargenomen in alle drie de groepen. Bij 80% van de onderzochte patiënten waren de concentraties metaalionen niet afwijkend. Ten gevolge van deze observaties werd bij één patiënt een revisie van de MoM heupresurfacing gedaan. We concludeerden dat klinisch onderzoek met daarbij alleen het maken van een röntgenfoto de aanwezigheid van pseudotumoren in MoM patiënten ernstig onderschat.

Hoofdstuk 7
In de wetenschappelijk literatuur werden ondertussen verschillende systemen beschreven om waargenomen pseudotumoren na een MOM heupprothese te classificeren, maar er was nog geen onderzoek gedaan waarin deze systemen werden vergeleken. Daardoor was er voor de kliniek maar beperkt informatie voorhanden over de betrouwbaarheid van deze classificatiesystemen. In hoofdstuk zeven hebben we onderzocht hoe betrouwbaar deze systemen waren in het classificeren van de ernst van de geobserveerde pseudotumoren, en wat de invloed was van de keuze voor een bepaald systeem. Voor deze studie hebben we een cohort van 42 THA patiënten (49 MoM heupen) onderzocht met behulp van drie verschillende pseudotumor graderingsystemen: (1) het systeem van Anderson et al, (2) Matthies et al en (3) Hauptfleisch et al.
De resultaten van twee ervaren musculoskeletale radiologen werden vergeleken (interbeoordelaarsbetrouwbaarheid) waarbij zij de MARS-MRI beelden steeds met deze drie systemen beoordeelden Daarnaast herhaalden zij hun beoordelingen enige tijd later, zodat we ook de intrabeoordelaarsbetrouwbaarheid konden berekenen. Onze resultaten toonden aan dat, ongeacht het gebruikte classificatiesysteem, de betrouwbaarheid van de beoordeling van de pseudotumor matig was. De uiteindelijke conclusie kon dan ook niet anders zijn dan dat er behoefte is aan een beknopt pseudotumor graderings systeem voor klinisch gebruik dat een hogere betrouwbaarheid laat zien.

**Hoofdstuk 8**

In hoofdstuk zes toonden we in een kleine groep MoM patiënten aan dat standaard röntgenologische follow-up in combinatie met klinisch onderzoek niet voldoende gevoelig was om pseudotumorvorming op te sporen. Daarop hebben we het intensieve screenings protocol toegepast op het volledige cohort van MoM heup resurfacing patiënten in onze praktijk. De resultaten van dit onderzoek worden beschreven in hoofdstuk acht. Bij de start van dit onderzoek, waren er 248 MoM heup resurfacing beschikbaar (214 patiënten, gemiddelde follow-up 4.6 jaar, range: 1 tot 8.2). In dit onderzoek vonden we een pseudotumor prevalentie van 36.3%; hierbij werd de ernst van 61 pseudotumoren beoordeeld als zijdne mild, 25 als zijnde matig ernstig en vier pseudotumoren als ernstig (Anderson classificatie). Vijf revisie-operaties volgden, allemaal bij symptomatische patiënten met verhoogde concentraties metalionen. Aangezien het natuurlijke beloop van pseudotumoren nog grotendeels onbekend is, en er geen consensus over de optimale behandeling van pseudotumoren na MoM heupprothese bestaat, werd er bij symptomatische patiënten met een milde tot matig ernstige pseudotumor en normale concentraties metaalionen gekozen voor een voorlopig conservatief beleid. Het gebruik van dit intensieve screeningsprotocol en het indelen van de ernst van de pseudotumor liet ons vooralsnog toe om dit conservatieve beleid te voeren. De patiënten met niet-gereviseerde pseudotumoren bleven onder verscherpt toezicht, inclusief herhalen van MARS-MRI.
Hoofdstuk 9
Zoals gesteld in hoofdstuk acht was een intensieve controle van de patiënten met niet-gereviseerde pseudotumoren nodig om de, voorlopig, conservatieve behandeling te valideren. Om het natuurlijke beloop van pseudotumoren in de tijd te volgen, herhaalden we deze onderzoeken zes tot 12 maanden na de eerste MARS-MRI, waarbij ook de metaal-ionen concentraties opnieuw werden gemeten en het klinisch onderzoek werd herhaald. In deze studie werden 14 niet-gereviseerde pseudotumoren bestudeerd evenals een controlegroep van 23 patiënten waarbij er geen pseudotumor aanwezig was op de eerste MARS-MRI. De gemiddelde postoperatieve tijd tot de eerste MARS-MRI was 4.3 jaar (range: 2.2 tot 8.3), en de gemiddelde tijd tussen de eerste en tweede MARS-MRI was acht maanden (range: 6 tot 12). Bij de meerderheid van de patiënten (35/37) was er geen verschil te zien in de ernst van de pseudotumor, eenmaal werd een nieuwe pseudotumor waargenomen (Anderson C2 score, matig ernstig) en eenmaal was de ernst van de pseudotumor afgenomen (van C2, matig ernstig, naar tot C1, mild). De conclusie van dit onderzoek was dat het herhalen van een MARS-MRI binnen een jaar bij niet-gereviseerde MoM patiënten met milde, tot matig ernstige asymptomatische pseudotumor na MoM resurfacing heup, van weinig nut is. Tegelijkertijd is het belangrijk deze conclusie voorzichtig te interpreteren omdat dit de eerste longitudinale MARS-MRI studie is naar het natuurlijk beloop van pseudotumoren.

Hoofdstuk 10
Aangezien de behandeling van de niet-gereviseerde pseudotumoren afhankelijk is van zowel de ernst (gebaseerd op locatie, inhoud en groeisnelheid) als van de afmetingen van de pseudotumor, is het relevant om een nauwkeurige klinische meetmethode beschikbaar te hebben die de dimensies van een pseudotumor goed kan meten. Het doel van de studie beschreven in hoofdstuk 10 was om klinische methoden waarmee (veranderingen in) afmetingen van pseudotumor kunnen worden gemeten te valideren met een laboratorium methode die als gouden standaard kan worden gebruikt. Deze laboratorium methode was een drie-dimensionale, region- of-interest (3-D ROI) volume meting. Voor deze studie waren er 13 MARS-MRI scans beschikbaar van niet-gereviseerde pseudotumoren, ontstaan na MoM heup resurfacing. Van deze patiënten was ook een tweede MARS-MRI beschikbaar voor de metingen. De gemiddelde follow-up bij de eerste
MARS-MRI was 5.3 jaar (range: 2.4 tot 7.5), en de tweede MARS-MRI was verkregen na gemiddeld 7.5 maanden (range: 6 tot 12). Op alle beschikbare scans werden de pseudotumor afmetingen gemeten middels: (1) maximale diameter in één vlak (MD), (2) door een schatting van het pseudotumor volume gebaseerd op de gemeten maximale diameter in drie verschillende vlakken (“Estimated volume, EV”). (3) Ter validatie werd van elke scan een 3-D ROI volume (V) berekend door de som van de oppervlakte van het pseudotumor gebied per MRI-segment te vermenigvuldigen met de MRI-segment dikte. De correlatie was het sterkst tussen EV en V, maar EV overschatte V gemiddeld met 72.6 %, vooral in niet-ellipsoïïd gevormde pseudotumoren. De mediane waarden voor MD, EV of V waren niet significant verschillend tussen de eerste en tweede MARS-MRI. De mediane verandering voor MD was 0 cm (range: -1.5 tot 3.4), 0.5 ml voor EV (range: -16.4 tot 45.5) en ook 0.5 ml voor V (range: -7.7 tot 5.2). Dit leidde tot de conclusie dat in de klinische praktijk de methode waarbij het pseudotumor volume (EV) werd geschat aan de hand van de maximale diameter gemeten in drie vlakken, beter kan worden gebruikt dan het meten van de maximale pseudotumor diameter in één vlak, en dat deze methode een sterke correlatie heeft met een meer complexere 3D-ROI methode. Klinisch moeten we daarbij wel rekening houden met een overschatting van de grootte van de pseudotumor, vooral in niet-ellipsoïde pseudotumoren. Daarnaast lijkt deze methode gemakkelijk toepasbaar in de klinische praktijk en kan dan ook worden gebruikt voor het monitoren van verandering in de grootte van geobserveerde pseudotumoren.

Hoofdstuk 11

Hoofdstuk 11 vat de resultaten van alle studies beschreven in dit proefschrift samen, en worden daarnaast de vraagstellingen beantwoord en worden voorstellen voor toekomstig onderzoek worden gedaan. De eerste doelstelling van dit proefschrift was om de implantaatoverleving van heupprotheses gebruikt voor de behandeling van ernstige heuparthrose bij jonge en actieve patiënten te onderzoeken. We concludeerden dat voor deze patiënten, de overleving van MoM heupresurfacing niet voldeed aan de internationale richtlijn, terwijl de ongecementeerde heupprothese met UHMWPE net voldeed aan deze richtlijn. Echter, bij deze laatste werd wel een hoge mate van slijtage opgemerkt in de eerste 10 jaar, waarbij er mogelijk nog een potentiële versnelde slijtage in het tweede decennium na implantatie zou kunnen optreden.
Het tweede doel van dit proefschrift was om de overleving van alle bekende heupresurfacing systemen te beoordelen. Na het systematisch analyseren van de wetenschappelijke literatuur, was de conclusie dat geen van deze implantaten aan de internationale richtlijn voor implantaatoverleving voldeed, en dat er heupresurfacing systemen in de kliniek werden gebruikt zonder dat daarvoor adequate klinische studies beschikbaar waren. Tevens kon worden vastgesteld dat, ten tijde van dit onderzoek, aseptische loslating de belangrijkste reden van falen was voor de verschillende MoM heupresurfacing systemen. De derde doelstelling van dit proefschrift was om onderzoek te doen naar complicaties die optraden na het plaatsen van een MoM heupprothese en die te wijten waren aan reacties op metalen slijtagedeeltjes. Na intensieve controle van alle patiënten met een dergelijke prothese in onze kliniek, konden wij concluderen dat de prevalentie van deze complicaties hoger was dan verwacht, en dat er geen eenduidige risicofactoren voor het ontstaan van deze complicaties konden worden vastgesteld. Ook concludeerden we dat het gebruik van MARS-MRI belangrijk was aangezien bij een aanzienlijk deel van deze complicaties werd waargenomen bij asymptomatische patiënten. Daarnaast bleek het toepassen van een pseudotumor classificatiesysteem een waardevolle aanvulling in de klinische praktijk. De vierde en laatste doelstelling van dit proefschrift was om de toepassing van MARS-MRI, het diagnostisch instrument voor deze complicaties, en de daarvoor beschikbare pseudotumor classificatiesystemen te onderzoeken. Op basis van onze studies concludeerden we dat met MARS-MRI en bij behorende classificatiesystemen de pseudotumoren goed konden worden geïdentificeerd maar dat er voor het vaststellen van de ernst van de pseudotumor nog een verdere verfijning nodig is. Toekomstige studies zouden de behandeling van conservatief behandelde pseudotumoren verder moeten valideren.
Disclosure

S. Austen: No conflict of interest

K. Brakel: No conflict of interest

H.J. Hoekstra: Institutional grant for research received from Biomet

R.J. Horn: No conflict of interest

B. Kaptein: No conflict of interest

R.G. Nelissen: Research support from Stryker, support from Springer, board member of ERASS (European Rheumatism and Arthritis Surgical Society) LROI (Dutch Arthroplasty Register)

P. Pilot: Paid employee of Biomet

T. Sijbesma: Institutional grant for research received from Biomet

J.M.H. Smolders: No conflict of interest

J.L.C. van Susante: No conflict of interest

R.W. Poolman: Consultant for Amgen Co; Link Orthopaedics; unpaid consultant for Biomet; Zimmer; Research funding by Amgen Co, AstraTech, ; Zimmer; Link Orthopaedics, editorial board of BMC Musculoskeletal Disorders Nederlands Tijdschrift voor Orthopedie; board member of Dutch Orthopaedic Trauma Association

R. Raidou: No conflict of interest
E.H. Schemitsch: Royalties from Stryker, consultant for Amgen Co; Stryker; Smith & Nephew; Wright Medical Technology, Inc.; Kuros, research funding by Smith & Nephew, Canadian Institutes of Health Research (CIHR); OMEGA; Smith & Nephew; Zimmer; Stryker; Synthes, support from Saunders/Mosby-Elsevier, editor for Journal of Orthopaedic Trauma, board member of Orthopaedic Trauma Association; Canadian Orthopaedic Association; Osteosynthesis and Trauma Care Foundation

W. van der Weegen: Institutional grant for research received from Biomet. Consultant for Biomet, board member of Werkgroep Orthopedie en Wetenschap.

J.A. Wullems: Institutional grant for research received from Biomet. Consultant for Biomet.
Dankwoord

Dit proefschrift is het resultaat van de inzet van velen. Als eerste wil ik mijn promotor, prof. Dr. Rob Nelissen bedanken. Jouw goede begeleiding heeft mijn boekje mogelijk gemaakt. Veel dank ben ik verschuldigd aan mijn beide copromotoren, dr. Thea Sijbesma, orthopedisch chirurg in het St. Anna ziekenhuis, en dr. Peter Pilot. Thea, samen hebben we het Kenniscentrum Orthopedie in Geldrop gebracht tot iets moois. Peter, jij bent de copromotor die elke promovendus naast zich wil hebben, het was fijn om met je te werken.

De beoordelingscommissie: prof. Dr. Bloem, Prof. Dr. Verhaar, dr. Verheyen, Prof. dr. Wilkinson, it was an honor to have you on the committee.

Henk Hoekstra, Wart van Zoest, Dirk Das en Rob Bogie, allen orthopedisch chirurg in het St. Anna ziekenhuis: we zijn een mooi team, dank voor al het werkplezier. Rogier van Drumpt, ook orthopedisch chirurg in het Anna, als paranimf verdien je een bijzondere vermelding. Dick Ernest, mijn andere paranimf, dank voor de ondersteuning.

Ellis Bos, physician assistant in het St. Anna ziekenhuis, je was en bent onmisbaar om onze studies te kunnen doen. Jur Vellema, niet langer in “het Anna”, maar dank voor jouw bijdrage aan het kenniscentrum. Jorgen Wullems, nu nog collega onderzoeker in het Anna: je hulp bij onze studies was onmisbaar. Alle co-auteurs die zo veel hebben bijgedragen: Shennah Austen, Koen Brakel, Roelof Horn, Bart Kaptein, Rudolf Poolman, Renate Raidou, Jose Smolders, Job van Susante, Emil Schemitsch: thank you!

Al dit onderzoek was niet mogelijk geweest zonder de inzet van mijn collega’s van de polikliniek, de verpleegafdeling, de operatiekamers en de MRI-kamer: dank jullie wel. Francine Teng, hartelijk dank voor het ondersteunen van het promotieproces, voor een buitenpromovendus ben je onmisbaar. John Mathijs, dank voor het proeflezen. Hans van de Berg, Roland van Esch, Jim Lancaster, Jing Xie, Han Geurts en Michiel Wouters: your relationship with our research is much appreciated.

Uiteraard ook dank aan alle patiënten die hebben deelgenomen aan de studies.

Mijn trainingsgroep: Godi, Jan, Peter, Peter en Mark, bedankt voor de afleiding en belangstelling.
Dit werk had ik niet kunnen doen zonder de niet aflatende steun van mijn ouders en mijn zus. Zonder jullie was dit proefschrift er nooit gekomen. Ook mijn schoonouders hebben daar aan bijgedragen, dank! Sylvie, jouw liefde, steun en vrolijkheid zijn de zaken die echt belangrijk voor mij zijn. Marte, Marit en Karst: liefs van papa!
List of publications


Submitted
About the author

Adrianus Philomena Catharina (Walter) van der Weegen was born in Hoogerheide on November 18th, 1965. After completing his physiotherapy education in 1989 he worked as a physiotherapist in the Netherlands, England and Canada. After returning to the Netherlands in 1997, he completed his Master’s degree in Health Sciences at Maastricht University in 2002 (specialization: Movement Sciences) under the supervision of dr H.A. Keizer. He started his scientific research career in sports medicine, but switched to orthopaedic research in 2006, working with the orthopaedic surgeons of the St. Anna hospital in Geldrop. He lives in Geldrop and is very, very happily married to Sylvie Hartman with whom he has two daughters, Marte and Marit, and a son, Karst.