Metal Debris Release From Metal-on-Metal Hip Arthroplasty: Mechanism, Quantification and Clinical Effects

David Langton
BEST COPY AVAILABLE.

VARIABLE PRINT QUALITY
TEXT BOUND CLOSE TO THE SPINE IN THE ORIGINAL THESIS
Metal Debris Release From Metal-on-Metal Hip Arthroplasty: Mechanism, Quantification and Clinical Effects

David Langton

Metal on metal (MoM) hip replacements consist of a cobalt-chromium-molybdenum alloy femoral head which articulates against an acetabular cup manufactured from similar material.

MoM hip replacements were introduced in the 1980s. It was thought that the overall reduction in volumetric wear as well as the avoidance of polyethylene would lead to greater longevity of these prostheses. There had been isolated reports of adverse tissue reactions with previous generations of MoM devices but it was thought that improved manufacturing technology would eliminate these problems.

In the 1990s, the Birmingham Hip Resurfacing (BHR) was developed. The positive mid-term results of this device led to a rapid increase in the use of the BHR throughout the world. For obvious reasons, the enhanced stability large diameter bearings provided proved extremely attractive to surgeons and patients. Manufacturers therefore began to develop total hip replacement systems for patients unsuitable for the resurfacing procedure. These systems used bearings of size 36mm and greater, in contrast to the existing 28mm Metasul device.

From 2005 onwards there began to emerge increasing numbers of reports of local complications in the tissues adjacent to MoM prostheses. These reactions included sterile masses, tissue destruction and osteolysis. The incidence of these tissue reactions was unknown, as were the risk factors for their development.

This piece of work sought to quantity the volumetric and linear wear rates of failed MoM hips and to investigate the relationship these wear rates and a number of clinical parameters. These parameters included blood, serum and hip fluid chromium and cobalt concentrations, and the macro and microscopic appearance of periprosthetic tissue at revision surgery. In this way it was hoped that component design, host and surgical factors leading to adverse tissue reactions could be identified and potentially eliminated.
Metal Debris Release From Metal-on-Metal Hip Arthroplasty: Mechanism, Quantification and Clinical Effects

David Langton

Background

Hip arthroplasty was developed in order to address the needs of patients disabled with pain and stiffness due to degenerative joint disease. The earliest recorded attempts at hip replacement occurred in Germany in 1891, when Professor Themistocles Glück used ivory to replace femoral heads of patients whose hip joints had been destroyed by tuberculosis[1]. In 1925, the American surgeon Marius Smith-Petersen created a mould arthroplasty out of glass. The glass was too brittle to survive the forces transmitted to it in vivo and he subsequently replaced them with a Vitallium (cobalt chromium molybdenum alloy) component in 1938. These Smith Peterson designs are generally thought of as the first resurfacings[2] though they were actually not intended as hip replacements. They were designed to be temporary moulds for cartilage regeneration, with the intention of removing the mould when the femoral head and acetabulum would have become smooth and congruent[2].

The 1940s, 50s and 60s saw three other innovators begin to develop hip prostheses that would have continued clinical impact up to the present day. Sir John Charnley initially suffered set-backs using a Teflon cup as part of a hip resurfacing device. In the immediate post-operative period patients were highly satisfied. But in the course of time, the plastic wore and the products of wear produced an intense inflammatory reaction. He found long term success by substituting high density polyethylene for Teflon to articulate against a stainless steel femoral head as part of a conventional THR[3][4]. Charnley's fundamental principle was that of a low friction arthroplasty system which he believed could be achieved by selecting acceptable materials with the lowest coefficient of friction and by minimizing the torque arm or head diameter[5]. In 1951 George McKee in Norwich introduced a total hip replacement made of stainless steel, but his first three appliances failed in less than a year due to loosening. He subsequently redesigned this apparatus using cobalt chromium and the results were undoubtedly improved, but loosening remained a problem until 1961 when he turned to a cobalt chromium prosthesis in which both components were fixed in position with bone cement[6]. This design was called the Mckee-Farrar - named with his colleague Mr John Watson-Farrar. At around the same time, Peter Ring developed another hip arthroplasty system which employed cobalt chromium alloy heads to articulate with cobalt chromium cups[6]. The Ring and Mckee designs were the so called “first generation” metal on metal (MoM) hip prostheses.

The principles of evidence based medicine were not fully embraced at this point in time. It is difficult therefore to satisfactorily assess the relative success or failure of these MoM prostheses. Irrespective of this, there was an irresistible shift to the global adoption of the Charnley metal on polyethylene (MoP) hip, much to the dissatisfaction of the MoM proponents. Charnley had convincingly argued that MoM bearings would seize up or give rise to high frictional torque and lead to early failures[5]. Peter Ring conceded in 1971: “There is no doubt that to date the high density polyethylene wears exceptionally well and forms an articulation with a very low coefficient of
friction. Although long-term results are not yet available, it seems likely that the plastic will have at least a ten year working life.” He went further: “There is, at the moment, no evidence that this is occurring, but there is no doubt that the metallic elements from which the alloy is created, cobalt, chromium and nickel will circulate within the blood and be removed from the region of the joint and can be found in tissues at some distance. There has been a suggestion recently that some patients might be sensitive to nickel and that this might produce a local reaction. Whether there is any substance in this argument, time must be left to tell.”[6] Despite these reservations, Ring continued to believe in MoM technology, arguing, in the same year, that “Our experience with this type of articulation now extends over a period of seven years; we have in all undertaken almost 600 replacements and have maintained these patients under continuous review. I think it is important at the outset to state that about 90% of these replacements can be rated as excellent results, that is to say these hips are pain free, have more than a right angle of flexion and the patients who own them are able to walk independently as far as they wish. Indeed, it can be calculated that after 50 years use the loss of substance in terms of loss of metal on the surface, will be no more than one or two thousandths of an inch.”[6]

There was a specific design feature of the Ring acetabular component which would significantly affect the joint’s performance. The acetabular component was fixed using a screw, the length of which dictated a particular orientation. “The mouth of the cup faced about 20° forwards and 20° laterally.... I regard this as the optimum position for the acetabular component,” Ring commented[6]. Clearly this would have conferred a great deal of protection from the effects of supero-lateral cup edge wear - a factor which would turn out to be something of an Achilles’ Heel for future generations of MoM hips. This is described in detail in chapters three to nine.

To this day, the reports detailing the clinical outcomes of the first generation MoM prostheses provide conflicting evidence. In general, the quality and number of studies available to the reader is limited. Brown et al reported a 28 year survival rate of 74%[7] in a retrospective analysis of 153 consecutive McKee-Farrar hips implanted between 1969 and 1973. The authors concluded in 2002 that “given the inherent problems associated with implant wear debris, especially polyethylene wear particles, second generation metal-on-metal bearing implants may offer a viable alternative to current designs. Their excellent long-term survival may infer particular suitability for use in younger patients.” The average (range) age of the patients in this study was 61 (21 - 85), meaning that the survival of the patients would significantly affect the analysis of the true longevity of the joints.

Wahlstrom et al’s prospective comparison of the Charnley and Mc Kee-Farrar devices “disclosed no major differences” at five years and concluded that their findings did “not support the hypothesis that the metal-on-metal prosthesis is clinically inferior to the metal-on-polyethylene prosthesis.”[8]

Wahlstrom’s twenty year results further concluded that “the long term results of the McKee-Farrar prosthesis are comparable with those of the low friction arthroplasty in this series. Wear of the polyethylene bearing and accumulation of polyethylene particles in the periprosthetic tissue may become an increasing problem. Second generation all metal implants seem to be worth considering in patients with long life expectancy.”[9] Sixteen McKee-Farrar and eighty Charnley prostheses had been revised for aseptic loosening, giving a 20-year aseptic probability of survival of 77% and 73%, respectively the authors found. However, only 18% of the hips were reviewed up to 21 years and radiographic signs of loosening were present in 52% of the surviving prostheses.
Bryant et al found markedly contrasting results[10]. Two hundred and fifty three Ring mark 2 MoM hip arthroplasties performed between 1968 and 1974 were evaluated using survivorship analysis. Using revision as the criterion for failure, the authors found a cumulative survival rate of 60.4% after 21 years. The results were compared with data from previous studies that had also used survivorship analysis for MoM hip arthroplasties, and it was shown that the Ring hip arthroplasty performed as well as the McKee-Farrar prosthesis. Other authors reported poor survival of the McKee-Farrar compared to modern standards[11] and Wretenberg's case study demonstrated that it was possible that a number of these patients may have been exposed to high concentrations of chromium and cobalt in the absence of local symptoms[12].

In the 1970s and 80s[13] it became increasingly apparent that MoP joints were developing problems after several years in vivo. These problems consisted of bony destruction (osteolysis) and subsequent loosening of components. It was initially thought that the bone cement used to fix the hips in place was the cause of these failures. However, after a number of years of research, the orthopaedic community reached a consensus[14]. Polyethylene wear debris stimulated a macrophage dominated immune response which led to the development of osteolysis[15]. While it was appreciated that some patients may have varying degrees of sensitivity to polyethylene wear debris, it was acknowledged that a reduction in wear was likely to reduce the incidence of osteolysis[16][17]. Younger, active patients were particularly at risk of wear induced osteolysis due to the longer they were likely to live and also due to the higher level of activity they were likely to subject their prostheses to. The Swedish Arthroplasty Joint Registry, one of the oldest of its kind in the world[18], provided evidence of this. For example, the 2002 report showed that the survivorship of conventional MoP THRs in males below the age of 55 years was only 80 per cent and 33 per cent at 10 and 16 years respectively[19].

It was primarily for this reason that manufacturers looked to develop new prostheses. Building on the knowledge and perceived success of the McKee-Farrar and Ring Prostheses, second-generation MoM THRs were developed in the 1980s. The Metasul 28mm bearing was a precisely engineered high carbon-containing wrought-forged cobalt chromium alloy manufactured by Sulzer[20]. This device was used extensively in Europe with good early clinical results. Some clinical studies showed outstanding survival rates of 98.6–99% at ten years with a near absence of osteolysis[21][22]. These findings were supported by findings of low wear of retrieved implants[23][24] which in turn confirmed preliminary hip simulator tests[25] and former clinical data[26]. But as with the previous generation of implants, there was some conflicting evidence. Some early failures were reported with these devices secondary to loosening in cemented versions[27] and there were reports of unusual cellular responses in the periprosthetic tissues of hips revised for unexplained pain[28].

There were also ongoing concerns that while the total volumetric wear of MoM joints was greatly reduced compared to polyethylene, the total number of released particles was of a scale of magnitude higher. The long term health effects of lifelong exposure to increased quantities of chromium and cobalt metal ions in the blood stream had not been investigated in depth and was far from fully understood. There were also concerns of deposition of metallic debris in end organs and possible carcinogenicity[29][30][31]. There appeared to be more immediate local concerns as well. In 1975, a paper in one of the major orthopaedic journals, the Journal of Bone and Joint Surgery of Britain, described a series of seven patients (six females) implanted with McKee-Farrar implants who had undergone revision for unexplained pain and were found to have thick yellow/green hip
effusions in association with soft tissue necrosis. The mean Co concentration in the joint fluid was found to be 250µg/l[32].

Despite these concerns, the Metasul MoM THR was implanted in large numbers of patients across the world over the next two decades. Data regarding blood metal ion concentrations and prospective randomised controlled data to prove its safety and clinical superiority over MoP remain limited even to the present day. Despite this, there was a resurgence of interest in MoM. Harlan Amstutz in America and Derek McMinn in the United Kingdom were convinced that MoM was the answer to solve wear induced osteolysis in younger patients and combined this bearing couple with hip resurfacing[33]. This form of arthroplasty had undergone its own evolution after a number of failures through the years. In hip resurfacing only the diseased articular parts of the femur are removed rather than the proximal femur being sectioned and discarded as in conventional THR. A hollow spherical head is then cemented over the femur, with the bearing diameter of the prosthesis being determined by the patient's existing anatomy. Hip resurfacing therefore provided two options highly attractive to both surgeons and patients: femoral bone conservation and a reduced risk of dislocation[34]. Furthermore, larger components were thought to more easily harness a beneficial fluid film and hence reduce wear, as had been shown in hip simulator studies[35].

There were dissenting voices though, from the past and the present. Many years before, Sir John Charnley, who believed that lubrication was too easily disrupted in vivo, had warned that large diameter bearings would generate larger frictional forces, leading to accelerated wear and loosening. Indeed this had been Charnley's own experience with MoP resurfacing[5], replicated by other pioneers including Amstutz himself in the 1970s[36]. But Amstutz and McMinn believed that if the clearance and surface roughness of the bearing surfaces could be optimised then the thickness of the lubricating film could be sufficient to sustain fluid film lubrication[35]. The production of precision engineered large diameter bearings could only feasibly be carried out using metal bearings at that time.

In 1991, Derek McMinn, working independently in Birmingham in the United Kingdom, began using a new MoM hip resurfacing. The prosthesis was developed in collaboration with Corin (Cirencester, UK) and was based on a cast cobalt chromium alloy[37]. After experiencing some failures with early devices, McMinn produced the Birmingham Hip Resurfacing (BHR) in conjunction with Midland Medical Technologies (MMT). The initial five year results of the BHR were extremely encouraging, with excellent clinical outcome scores, extremely low revision rates and no dislocations reported[38]. Consequently, aided by an aggressive marketing campaign, a number of centres throughout the world began to offer the resurfacing procedure. Smith and Nephew purchased the rights to the BHR from MMT in 2004 and, backed by the financial might of this company, the procedure became even more popular. Rival manufacturers understood the appeal of MoM hip resurfacing and realised that they would lose this market share if they did not have their own devices. In a short period of time, at least ten resurfacing devices appeared on the market, each with their own unique properties[39].

Manufacturers then looked to expand the use of large diameter MoM hip arthroplasty beyond the boundaries of hip resurfacing. A pre requisite for hip resurfacing is a proximal femur which has not been extensively damaged by degenerative disease[40]. Large diameter THR, which had previously been used only for situations in which hip resurfacings failed due to femoral neck fracture, therefore
began to be marketed as a procedure in their own right for patients without sufficient femoral bone quality to support a resurfacing. In this way the popularity of MoM hip arthroplasty increased still further. In 2008, over one third of all hip arthroplasties in the United States, the largest market in the world, utilised a MoM bearing couple[41].

Through the years of expansion of the MoM market, some surgeons continued to express concern over the long term health effects of exposure to increased concentrations of chromium and cobalt metal ions[42][43]. These unknowns remain unknowns. But in 2005, Willert, who decades earlier had described polyethylene induced osteolysis in conventional THRs, described a different, lymphocytic dominated reaction associated with a number of failed second generation Metasul hips[28]. Case reports began to emerge of revision surgeries where the description of large green/yellow sterile fluid effusions in the presence of soft tissue necrosis was strikingly similar to those described in the 1975 paper by Jones et al[32]. The orthopaedic community as a whole struggled to understand why these problems were occurring, how many patients were affected and which patients were at risk[45].

The first report on a series of resurfacing patients with pseudotumours from Oxford[45] in July 2008 summed up the depth of knowledge of the area at the time I commenced my research. The authors stated in their discussion:

“We therefore believe it is a new type of complication (pseudotumours) which is directly related to hip resurfacing. We do not know its precise incidence.......we estimate the incidence to be approximately 1% at five years. We do not know whether this will increase with time or not. If it does, it may present a major problem.

The spectrum of host changes seen in these pseudotumours could represent that of an inflammatory response associated with a delayed hypersensitivity reaction to antigen components such as nickel-chromium or chromium cobalt[45].

The cause of these pseudotumours is not apparent. Metal wear particles were found in every case which was examined histologically. However, in most cases there was no gross clinical metallosis. There is weak evidence of a relationship between the inclination of the acetabular component and the time of onset of symptoms; steeply positioned cups are more likely to lead to edge loading thereby generating a large amount of metal debris. However, there were also well-positioned implants in this series. We suggest that in some cases the pseudotumour may be the result of a toxic effect on cells of a large amount of particulate wear debris resulting from some problem with the articulation such as edge loading, whereas in other cases it is an idiosyncratic response to a moderate release of cobalt-chrome particles.

All our patients were female, which raises the possibility that pre-operative sensitisation to metal may be a factor. Nickel allergy is predominantly a condition of women and is related to the wearing of jewellery and in particular to ear piercing. It is interesting to note that seven of the patients in our series were allergic to antibiotics.”

Given that this was an emerging clinical problem, it was understandable that the authors of the Oxford article did not analyse blood or joint fluid metal ion concentrations. They also did not have
suitable equipment to assess the three dimensional orientation of the implanted acetabular components. No analysis was conducted to assess the wear of explanted prostheses.

At the time the Oxford pseudotumour paper was submitted, my full time research post at University Hospital of North Tees hospital had commenced. The role was intended to audit the clinical, radiological and biochemical performance of MoM hip arthroplasties at the hospital with a specific emphasis on the DePuy ASR hip resurfacing.

The Articular Surface Replacement ((ASR) (DePuy)) had been implanted from April 2004 at North Tees Hospital. It was marketed as a “fourth generation” hip resurfacing, implying that it had certain design modifications compared to the BHR which were intended to improve its tribological and clinical performance[46]. The diurnal clearance was markedly lower than the BHR with the intention being to enhance fluid film lubrication[47]. It had a thinner cup so that less acetabular bone had to be removed prior to implantation and the effective articular coverage angle was reduced in order to allow a greater range of motion[46]. The acetabular cup also underwent post casting heat treatment and hot isostatic pressing and solution annealing[39].

Derek McMinn, the designer surgeon of the BHR expressed concern over certain design features of these “fourth generation” resurfacings, most notably the metallurgy of the bearing surfaces, which remains one of the most controversial issues in contemporary hip resurfacing. Whilst all manufacturers use high carbon containing cobalt chromium alloy, the processing of the alloy differs. The alloy can either be wrought, forged or cast. Following investment casting, high-carbon cobalt chrome forms a typical microstructure with large blocky carbides precipitating in the metal matrix. These $M_{23}C_6$ carbides have the same hardness as alumina ceramic and confer wear resistance on the material when used as a MoM articulation[48]. McMinn believed that heat treatments were detrimental to the wear properties of metal bearings as they depleted or weakened the carbide structures. He and his team had investigated three commonly used heat processes: hot isostatic pressing, solution heat treatment, and sintering, all of which involved heating the metal to approximately 1200°C. McMinn made the conclusions that: “The effects of these processes were essentially twofold. Firstly, the overall carbide volume fraction of the alloy was reduced. Secondly, the original large blocky carbides disintegrate into smaller particles, making them unstable and easily dislodgeable under stress.”[49] Grigoris disagreed, writing in 2006 that: “The significance of these heat treatments has been debated over the last 7 years. Hip simulator studies do not demonstrate significant differences between the wear behaviours of as-cast and heat-treated alloys. Until long-term clinical outcomes and reliable retrieval studies become available, it will not be possible to determine the relevance of surface carbide depletion.”[50]

But McMinn argued that his cohort of heat treated hip resurfacing components showed a significantly increased rate of failure compared to components which had not been heat treated. These failures from McMinn’s 1996 series of double heat-treated components had not become apparent until more than 5.5 years after implantation[51]. Smith and Nephew went on to demonstrate a significantly increased rate of wear in heat treated compared to as cast cobalt chromium alloy using a novel physiological hip simulator protocol[52]. These findings were consistent with the pin-on-plate tests reported by Kinbrum and Unsworth[53] but in contrast to those reported in a previously conducted severe wear hip simulator test[54].
The acetabular components of some hip resurfacing systems, most notably the ASR, were designed with smaller coverage angles[55] (i.e., the articular surfaces subtended much smaller angles) than the true hemispheres of conventional hip designs. This was promoted as a way of reducing impingement and increasing range of motion. However, it left the cups potentially vulnerable to rim loading and edge wear, a known cause of catastrophically increased rates of wear. DePuy’s own design engineers confirmed this in simulator studies three years after the ASR resurfacing had been released[56].

Hip simulator studies had shown that another way of promoting fluid film lubrication, and thus a reduction in wear rates, was by reducing the diametral clearance between the head and cup[47]. The ASR was developed with the lowest diametral clearance of all commercially available devices[39]. This feature, coupled with a thin cup wall led to concerns of cup deformation on impaction of the cup. This could lead to starvation of lubrication and even clamping of the head by the deformed cup[57]. DePuy’s own pre-clinical hip simulator studies had shown significantly increased rates of wear when cups were deformed prior to the tests[58].

Given this background, what was the clinical situation? Three ASR patients at University Hospital of North Tees, all female, had suffered florid periprosthetic tissue reactions by September 2007. The research therefore focused on several key issues:

1. The quantification of the chromium and cobalt blood and serum metal ion exposure in MoM hip prostheses.

2. The identification of key variables associated with ‘excessive’ (and this was to be defined) metal ion release.

3. A comparison of blood/serum metal ion concentrations between commonly used resurfacing prostheses.

4. The investigation of the adverse local effects of accelerated wear in MoM prostheses.

5. The establishment of the rates of failure secondary to adverse reactions to metal debris of commonly used resurfacing prostheses.

6. The investigation of the link between excessive’ (again to be defined) wear and extent of local tissue damage.

7. The identification of trademark pathological appearances of periprosthetic tissues retrieved from patients with failed MoM prostheses.

The ultimate aim was to attempt to answer two basic questions:

*Why do adverse reactions to metal debris (ARMD) occur?*

*Who are the at risk patients?*

Answers to the above were provided and disseminated to the wider orthopaedic and bioengineering community through the following peer-reviewed journal publications.
Chapter One

"The Influence of Age and Sex on Early Clinical Results After Hip Resurfacing: An Independent Center Analysis."

A prospective study of the ASR resurfacing and ASR XL device was commenced at the University Hospital of North Tees in 2004. All patients with devices implanted by Mr Antoni Nargol, consultant orthopaedic surgeon, were included in the study. The patients were assessed preoperatively using internationally accepted scoring systems. They were then regularly reviewed clinically and radiologically. The paper "The Influence of Age and Sex on Early Clinical Results After Hip Resurfacing: An Independent Center Analysis" describes the results in detail. The initial satisfaction scores and functional results of patients implanted with the ASR device were encouraging although the operation appeared to be more successful in males compared to females.

Chapter Two

"Cup Anteversion in Hip Resurfacing: Validation of EBRA and the Presentation of a Simple Clinical Grading System."

Radiological assessment of the ASR cohort was expanded to include analysis of the three dimensional orientation of the acetabular components using EBRA software. As this software had not been used previously to examine resurfacing cups, a radiological study was carried out to establish the reliability of this technique for these devices. The results showed that EBRA was sufficiently accurate for the purposes of the research. EBRA has subsequently been used in at least five other studies[59][60] to record the orientation of MoM resurfacing cups. Hart et al published a study entitled “Large ball metal on metal hips obscure cup angle measurement on plain radiographs” however the findings were not applicable to my results given that the authors had simply measured anteversion on plain lateral radiographs rather than specialised software.[61]

Chapter Three

"The effect of component size and orientation on the concentrations of metal ions after resurfacing arthroplasty of the hip."

As detailed in "The Influence of Age and Sex on Early Clinical Results After Hip Resurfacing: An Independent Center Analysis" (Chapter One), three ASR patients were encountered who had developed worsening pain in association with soft tissue destruction and sterile joint fluid effusions. After the first two of these patients had failed to improve following two stage revision to MoM THRs for assumed infections, it was suspected that the clinical problems may be linked to excessive Cr and Co debris generated from the MoM bearing surfaces. The appearance of the tissues at revision surgery bore striking similarity to the descriptions of failed MoM joints in 1975[32]. For this reason a blood metal ion screening programme was instigated to accompany the prospective review of the patients. In the past Brodner[62], and, more recently Venditolli[63], had shown a link between cup
orientation and blood metal ion concentrations in MoM hips. The three ASR patients with suspected metal reactions were found to have cup inclinations greater than 50 degrees. For this reason the relationship between acetabular cup orientation and blood metal ion concentrations was studied. Established tribological theory suggested that lubrication is increased (and thus wear reduced) when larger bearing surfaces are utilised in joint replacements[64]. Therefore, the relationship between bearing diameter and blood ions was also examined, as well as a host of other variables including patient activity[65].

The results showed that the ASR device appeared to be extremely sensitive to cup orientation in terms of the release of chromium and cobalt ions. Larger diameter ASRs appeared to be protected to some extent from the effects of cup orientation. The results were similar to, although the metal ion concentrations were generally much higher, than those reported by Venditolfi[63]. The clinical and biochemical results needed to be compared to the gold standard implant in this class - the BHR, which was an ODEP 5a rated implant. Mr Nargol had originally used the BHR for all hip resurfacing operations from 2002 to 2004. The patients who received the BHR during this period were reviewed in clinic in a similar manner to the ASR patients in order to audit the results and ensure that they were not developing any unforeseen problems.

Chapter Four

"Blood metal ion concentrations after hip resurfacing arthroplasty: A COMPARATIVE STUDY OF ARTICULAR SURFACE REPLACEMENT AND BIRMINGHAM HIP RESURFACING ARTHROPLASTIES."

At the very beginning of 2008, the patients who had received BHRs were not undergoing regular follow up. Although the complications encountered with the new ASR device at this point in time were relatively few in number, there was a wide variation in the blood tests in terms of chromium and cobalt concentrations. It was therefore deemed prudent to examine the BHR blood metal ion results in the same way that ASR results had been. Bearing diameters and acetabular component orientations were again measured and documented. As with the ASR, there was an inverse relationship between increasing bearing diameter and decreasing metal ion concentrations. This relationship was much weaker, however, than in the case of the ASR. The most significant findings however concerned the metal ion concentrations as a whole, as well as the relationship of the ions to acetabular orientation. There were far fewer patients in the BHR group who were found to have extremely high cobalt or chromium levels. This finding was not satisfactorily explained by a difference in cup orientations between the ASR and BHR patient groups. Patients with sub-optimally positioned BHR components were significantly less likely to develop high blood metal ions compared to the equivalent ASR patients.

Despite the BHR having been implanted since 1997 and tens of thousands of surgeries being carried out subsequently worldwide, published blood metal ion data was extremely limited. Of note Daniel, working with the designer of the BHR, had performed a small study recording the blood metal ion concentrations in 26 patients[66]. The authors stated that there was no association between bearing diameter and metal ions, although only patients with 50mm and 54mm bearings were included and cup orientation was not studied in depth. Their conclusions drew criticism from other authors[67].
Chapter Five

“Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: A CONSEQUENCE OF EXCESS WEAR.”

“Tribological analysis of failed resurfacing hip prostheses and comparison with clinical data.”

“A study of the wear of explanted metal-on-metal resurfacing hip prostheses.”

Through the year 2008, seventeen patients with ASR devices had experienced ARMD. All patients who had received ASRs and BHRs were now under regular review. The vast majority remained asymptomatic at this point in time. We compared a number of parameters between this large pool of asymptomatic patients and those who had developed complications. These parameters included blood/serum Cr and Co concentrations, joint fluid concentrations (using samples obtained from patients whose joints had failed for other reasons, such as avascular necrosis), acetabular orientation and bearing diameter. ASR implants which had been revised underwent analysis using a ZYGO NewView 5000 non-contacting profilometer and a Zeiss TSK Rondcom60A roundness measuring machine. Measurements taken with the non-contacting profilometer showed that roughness values of the articulating surfaces had increased. This implied that the theoretical lubrication regime would have shifted from fluid film to boundary lubrication. In the boundary lubrication regime, with its preponderance of surface-to-surface interaction, increased wear would be expected. In turn, this would result in higher volumes of wear debris. In addition, the relatively large size of the resurfacing hip prostheses meant that, under boundary lubrication conditions, wear would take place over a large sliding distance and volumes of wear debris would be maximized. Measurements taken with the Rondcom showed that ARMD failures were associated with values of out of roundness far higher than at manufacture. Data obtained from the retrieval analysis, combined with the blood and joint metal ion concentrations provided convincing evidence that ARMD failure was associated with abnormal wear of prostheses. It was clear that there was an obvious difference in the clinical performance of the ASR and BHR. A much smaller percentage of the BHR patients were found to have extremely high blood metal ion concentrations and up until this point, no ARMD failures had occurred in the BHR patient group. Furthermore, patients with failed MoM hips who who gave blood for in vitro lymphocyte proliferation tests were not found to have results indicative of existing metal hypersensitivity. These findings were substantiated later by Kwon et al in an investigation involving patients with pseudotumours[68][69]. The implication was that ARMD was directly related to excessive wear of MoM hips rather than an idiopathic host response.

The findings prompted the cessation of the use of the ASR at University Hospital of North Tees.

Chapter Six

“Volumetric wear assessment of failed metal-on-metal hip resurfacing prostheses.”

“Reducing Metal Ion Release Following Hip Resurfacing Arthroplasty.”
In 2008, following ethical approval from County Durham and Tees LREC, a prospective study at North Tees Hospital and Newcastle University was initiated to characterise and quantify material loss from failed MoM hip prostheses. A small Belgian study had previously shown that serum ion concentrations were significantly related to linear wear of retrieved femoral components[70]. Up until this point it had been our working hypothesis that the Cr and Co concentrations observed in the blood of the patients who had received MoM prostheses were directly related to the amount of material liberated from the bearing surfaces. The aim of this study, therefore, was therefore to examine this assumption by comparing calculated volumetric wear rates with various clinical results including blood/serum/joint chromium and cobalt concentrations.

Previously, a number of papers from other centres had used coordinate measuring machine (CMM) technology to calculate volumetric loss from failed devices[71]. With the aid of a grant from the British Orthopaedic Association/Joint Action a state of the art, high accuracy CMM was purchased. A bespoke Matlab programme was designed to allow accurate quantification of volumetric and linear wear. Gravimetric measurements were utilised to validate the analytical techniques using a sterile component. This validation study was subsequently expanded in a more recent paper[72] described in chapter eleven.

Chapter Seven

"Reducing Metal Ion Release Following Hip Resurfacing Arthroplasty."

Up until this point in time, only one surgeon’s results had been analysed in depth. While this was an ideal way to examine the performance of the ASR against the BHR while excluding differences in technique between surgeons, it was not clear whether the results and conclusions could be applied to the wider surgical community. For this reason, our existing data was pooled with the clinical, blood metal ion and radiological results from two other high volume hip resurfacing surgeons in Europe. A study of this type had previously been recommended in the literature[73]. Three devices were studied: the ASR, the BHR and the Conserve Plus. The Conserve Plus was an ideal device to include as it was well known to have one of the largest acetabular coverage angles, and this coverage did not change significantly between cups with different diameters.

This study included the largest analysis of radiological orientation of acetabular cups in existence. The range of acetabular component placement was found to differ very little between surgeons, a finding which was entirely consistent with the existing literature on the subject[74][75].

The results substantiated our previous findings that the ASR was more sensitive to the effects of acetabular component orientation compared to devices with larger acetabular coverage angles. The Conserve Plus, with a large coverage angle that did not vary significantly with bearing diameter was the least sensitive to component position, and blood metal ion concentrations were not found to be significantly related to bearing diameter. These conclusions were in agreement with those made previously by De Smet et al[76].

Chapter Eight
"Adverse reaction to metal debris following hip resurfacing: THE INFLUENCE OF COMPONENT TYPE, ORIENTATION AND VOLUMETRIC WEAR."

"Adverse reactions to metal debris: histopathological features of periprosthetic soft tissue reactions seen in association with failed metal on metal hip arthroplasties."

Following on from the previous study which focused solely on blood metal ion concentrations, we used the same pooled data to analyse the incidence of local ARMD in this large series of 4,226 hip resurfacings. At the time of writing, this study is the largest analysis of clinical, biochemical and radiological results of hip resurfacings in the global published literature. As well as documenting the incidence of ARMD in each surgeon/device group, a Cox proportional hazards model was constructed to examine the risk factors for joint failure secondary to ARMD. The Cox model showed that the most significant variable leading to ARMD related failure was the ASR device itself. These findings were consistent with our earlier single surgeon study which showed a much greater ARMD failure rate in ASR patients compared to BHR patients (chapter 5).

Another aim of this study was to examine the relationship between wear and local tissue reactions. Our previous work had shown for the first time that ARMD had occurred in association with abnormally wearing MoM prostheses. Our clinical experience was that, while the development of bony injury tended to be strongly related to wear rates (and thus, metal ion concentrations), the extent of soft tissue damage seemed to be linked to the intensity of the local cellular response rather than being dose related to wear. In order to investigate this relationship, the macroscopic appearances of bone and soft tissues at revision surgeries were categorised using a novel grading system. Similarly, tissue specimens excised at revision were graded by a consultant histopathologist. These grading scores were then compared to the measured volumetric wear rates of the corresponding explanted components.

There were a number of key findings. No patients experiencing joint failure secondary to ARMD were proven to have "normal" or "expected" wear of their retrieved devices and cup edge wear was identified on all explanted acetabular components. Compatible with this finding was the observation that no patients who had developed ARMD were found to have low blood metal ion concentrations. Neither volumetric wear rates nor joint/serum ion concentrations correlated with microscopic necrosis or the extent of tissue destruction observed at revision surgery. There was, however, a significant relationship between the thickness of perivascular lymphocytic cuffs and surface layer necrosis which suggested that tissue destruction was not a result of toxic concentrations of metal debris, but was more likely to be the result of an immune response provoked by the debris.

Chapter Nine

"Accelerating failure rate of the ASR total hip replacement."

"High failure rates with a large-diameter hybrid metal-on-metal total hip replacement: CLINICAL, RADIOLOGICAL AND RETRIEVAL ANALYSIS."

From 2010 onwards it became increasingly apparent that there was a disproportionate increase in the rates of failure of the ASR THR compared to the resurfacings. We had previously noted this
discrepancy in our earlier work (chapter five), ascribing the difference to the increased number of females who had received THRs compared to ASR resurfacings. Females, due to their smaller anatomy, are more likely to be implanted with smaller diameter components. As we had previously shown that smaller diameter cups were more likely to generate greater concentration of metal ions [77], we believed the difference in failure to be due to increased bearing surface wear rather than any specific difference in the construct of the arthroplasty systems themselves.

However, as more failures were encountered at University Hospital of North Tees, more data became available for analysis. It became clear that the bearing surfaces of ARMD ASR THRs had experienced significantly less wear than the equivalent ASR resurfacings. Correspondingly, the blood cobalt and chromium concentrations were also significantly lower in the THR group. A re-examination of all clinical data showed that the clinical presentation of patients in the two groups was similar, and revision findings were also similar, if not more severe in the THR group. The ratio of cobalt to chromium in the joint fluid of the THR patients was potentially indicative of a different mechanism of metal debris release however. All explants were therefore re-evaluated with a specific emphasis on the junction between the femoral components and femoral stems.

A review of the literature showed that modular junctions in hip arthroplasty had been a source of much debate in the 1990s[78][79] but relatively little research had been focused on this area in the 21st century. Initial investigations consisted of basic out of roundness traces using a coordinate measuring machine to identify gross changes of the surface of the female taper surfaces of the explanted femoral heads.

The initial clinical and analytical results were published in two papers from two centres. The clinical results at these two centres were extremely concerning. Both centres were experiencing very high failure rates with stemmed MoM THR - even when the examination of retrieved explants showed very little bearing surface wear had taken place.

Interestingly however, overall, the total measured volumetric CoCr loss (taper and bearing surface combined) of an explanted THR from a patient suffering ARMD was approximately one third that from a resurfacing device[80]. Coupled with this finding, CoCr ratios in fluids taken from joints with severe taper damage were somewhat different than those with excess bearing surface wear. For example, with severe bearing surface wear the ratio of Cr to Co is frequently more than 3 to 1. Taper wear was often associated with a reversal of this ratio, with Co the dominant element. It should also be noted that a typical CoCrMo alloy used in an artificial hip will consist of approximately 60-65 % Co and 27-30 % Cr.

An examination of published literature shows that the failure rates of ASRs, Duroms and BHRs THRs are greater than the failure rates of the pure resurfacing systems[81]. Furthermore, Garbuz et al [82] Nargol et al[83] and Beaulé et al[84] have shown evidence that median blood metal ion concentrations are elevated in THRs compared with their resurfacing counterparts in studies involving the Durom (Zimmer, Warsaw, Indiana), ASR and Conserve Plus systems (Wright Medical, Memphis, Tennessee), respectively. A prospective study in the United Kingdom comparing the Birmingham Hip Resurfacing with the BHR THR was terminated due to unacceptably high metal ion levels and failures in patients receiving the THRs[85].
It is highly likely that these clinical outcomes are the consequences of taper wear debris. Taper wear debris itself appears to be particularly potent in terms of causing direct or indirect soft tissue damage\[86\]. We have examined this relationship in our ongoing unpublished work\[80\]. When matched with resurfacing patients for total volumetric loss (taper and bearing surface combined), patients with taper damage are more likely to develop moderate to severe soft tissue destruction, are more likely to have moderate/severe ALVAL type reactions and are more likely to develop granulomas\[80\].

Reports of ARMD in the presence of optimally functioning MoM hip implants are rare. However it does not follow that metal ion concentrations (and thus total volumetric wear) are dose related to the amount of tissue damage identified at revision surgery\[87][88\]. Previously, ARMD has been divided into histiocyte dominated response (ARMD-H) and lymphocytic dominated responses (ARMD-L)\[88\]. ARMD-L cases are often found to have the greatest extent of soft tissue destruction. The implication is that soft tissue damage is caused by the immune response rather than a direct toxic effect of the locally elevated Cr and Co concentrations. It is this distinction which may ultimately prove to be the underlying explanation for the difference in clinical performance and failure modes of hip resurfacing and THR systems. It may also explain the confusion which reigns in the literature on the subject of the clinical use of blood metal ion tests. It is interesting to note that when patients with resurfacings are analysed as a group on their own, authors tend to conclude that risk factors for failure are smaller joint sizes, sub optimally positioned acetabular components and elevated metal ion levels/increased wear\[88][89\]. However, when studies examine THR patient groups (or actually combine THR and resurfacing patients into the same group), it is often concluded that ARMD is not related to size, orientation or elevated metal ion concentrations\[90\].

Frequently patients with THRs who develop ARMD have blood or serum Cr or Co lower than the 7µg/l threshold level suggested by the MHRA. This has been noted by a number of authors\[83][91\]. It may also be the reason why one study found that the wear rates of failed ASR devices (a device widely acknowledged to be a catastrophically failing device) were not found to be significantly different to the wear rates of BHRs. Examining the breakdown of explants in this study revealed something quite interesting - the ASR explants were split 1.75 : 1 ASR THRs to ASR resurfacings and the BHR explants 1 : 2.76 BHR THRs to BHRs. The same authors also found that pseudotumours are common in well positioned low wearing devices. Revealingly, there was no mention in either study of the taper junctions\[92][93\].

Given that, as discussed above, soft tissue destruction does not appear to be directly linearly related to wear rates/metal ion concentrations but rather to a negative immune cascade once a threshold has been passed, it is not unreasonable to suggest that taper debris is more immunogenic. Why might this be the case? There are a number of possibilities:

1. Taper debris may be released by different mechanisms and therefore may be of a different morphology to bearing surface debris. The size and shape of particles released from this junction may have a different response on the immune system which may have a greater capacity to cause soft tissue injury. There is a large body of work throughout many fields of medicine and surgery to show that immune responses to foreign materials are very much dictated by the morphology of those materials.
2. Titanium alloy debris may have a significant synergistic effect with CoCr debris. It may be that the different types of ions compete for different metallocprotein binding sites. Injury may be caused indirectly due to certain metallocproteins instigating a negative immune response, or directly, as a result of the concentration of unbound metallic ions in the joint fluid.

3. Co and Cr ions may be released in different relative quantities. For example, with bearing surface wear there is, theoretically, constant repassivation of the CoCr surface by the formation of Cr oxide. At the taper junction there appears to be no such repassivation of CoCr and the alloy itself is released in a proportion which is more consistent with the elemental proportions of the alloy. Failed taper junctions are associated with higher Co concentrations in the joint fluid. It could be that this Co preponderance is key to the process of tissue damage or it may simply be an association. One thing that is clear, however, is that tissue injury is not directly related to overall total wear or joint fluid metal ion concentration. When joint concentrations are at their highest, Cr concentrations dominate (likely due to the sequestration of chromium orthophosphate in the periprosthetic tissue) and it may even be that high concentrations of Cr in some way inhibit a lymphocytic dominant response.

4. Acceleration of debris production rather than total dose or average wear rate may be more significant to the host response. When tapers begin to fail the acceleration in debris production may be far steeper than in a failing bearing surface. It is also possible that debris could be released in bursts as the taper junction opens and seals with various activities of daily living.

5. The location of debris production. Firstly, synovial fluid may have an important chemical effect on the released debris. Furthermore with uncemented stems, debris liberated from the taper junction may find a route to the bone marrow, again with differing immunogenic potential.

6. Taper debris may be an association rather than a root cause. For example it has been suggested fluid ejected when head and cup components separate and relocate have an important role in the development of tissue necrosis[94]. This mechanism may simply be more frequent in THR compared to resurfacings.

Irrespective of the underlying mechanisms, large diameter MoM THRs are failing at a higher rate than their resurfacing counterparts. The ASR THR has been described as having a failure rate as much as 44% at 7 years[95]. I have personally examined 120 failed ASR THRs. In my opinion, taper failure has occurred in over 60% of these cases. The extent of taper damage is not clearly linked to surface wear - the implication being that there is no reason at present to say without doubt that this is uniquely an ASR problem.

Most importantly, at present, Co Cr blood levels, unless they are within the ranges of physiological values described herein (< 2 μg/l) are of no real use for these arthroplasty systems in terms of identifying pathological responses. All patients with these devices must be kept under strict follow up until we know more.

Chapter Ten
"Taper junction failure in large-diameter metal-on-metal bearings."

The previous study identified significant surface changes at the modular junctions of ASR THRs. These changes were associated with negative clinical outcomes indistinguishable from ARMD as seen in MoM hip resurfacings. Furthermore it appeared that THR patients could develop significant soft tissue destruction when exposed to lower concentrations of metal debris as evidenced by a comparison of the two groups in terms of bearing surface material loss and blood and joint metal ion concentrations. A mass recall of all MoM hip patients was commenced at University Hospital of North Tees from 2010 onwards. Blood metal ion testing was offered to all patients. Patients with high blood metal ion concentrations and/or pain were referred for cross sectional imaging by an experienced musculoskeletal radiologist. A number of ARMD failures were identified in patients who had received the Pinnacle MoM THR. This THR system, also manufactured by DePuy was implanted in approximately 1000 patients at the hospital by different consultant orthopaedic surgeons. The bearing surfaces are manufactured from wrought CoCr and the heads are implanted onto titanium alloy stems, identical to the larger diameter ASR THRs. The taper junction specifications are identical to the ASR systems. Unsurprisingly, a number of explanted Pinnacle devices showed patterns of taper surface change visual to the naked eye identical to those identified on the retrieved ASRs.

It was obviously desirable to develop a method to quantify and characterise the material loss from the taper junctions. Preliminary investigations had shown that the most significant damage was occurring on the female taper (femoral head) side. We therefore developed and validated a method to quantify material loss from the female tapers in ASR and Pinnacle systems. This is described in the paper in this chapter.

The results of the analysis were extremely concerning. They may, in fact, prove to have huge implications for contemporary hip arthroplasty. We showed clear evidence that the titanium alloy stems (trunnions) were causing changes at the female CoCr alloy taper surface and leaving behind an imprint of the trunnion grooves. Following this imprinting, preferential wear of the CoCr was taking place. Variables associated with increased wear rates were larger head diameters, increased head offsets and lower femoral stem shaft (varus) angles. In a number of cases, volumetric loss at the taper junction exceeded that at the bearing surface.

Chapter Eleven

"Blood metal ion testing is an effective screening tool to identify poorly performing metal on metal bearing surfaces."

Our previous work had suggested that the rate of ARMD in low wearing MoM hip resurfacings was low. We had also demonstrated a highly significant relationship between blood/serum metal ion concentrations and the wear of hip resurfacings as measured post explantation. Additionally we had presented evidence to show that cobalt was the most reliable element to test between different laboratories[96]. The next most obvious next step was to determine whether a reliable threshold concentration of cobalt existed which could be used to identify abnormally wearing implants while they remained in vivo. We did this by comparing pre revision blood metal ion concentrations to the wear rates of retrieved MoM components. The process is described in depth in the paper in this
chapter. To allow legitimate conclusions to be drawn from this data, we recognised the need for a large study of background Cr Co concentrations in a healthy population. This data was provided via collaboration with the local regional toxicology centre and is included in the aforementioned paper. Using the background concentrations in combination with pre revision blood samples it was clear that a blood Co concentration above 5µg/l was highly specific for the identification of abnormal wear of a MoM hip.

Throughout the acquisition and analysis of data, it became apparent that there was a phenomenon occurring where, as overall metal ion concentrations increased, serum concentrations of Cr became greater than the whole blood concentrations. Given that concerns already existed as to the possible generation of hexavalent Cr species by MoM joints, we investigated the distribution of hexavalent and trivalent Cr species in blood and serum with an in vitro study. This investigation is described in detail in this chapter.

Chapter Twelve

“The clinical implications of elevated blood metal ion concentrations in asymptomatic patients with MoM hip resurfacings: a cohort study.”

In 2012, the Medicines and Healthcare products Regulatory Agency (MHRA) of the United Kingdom released the latest guidelines on the management of patients with MoM hip joints. These guidelines did not recommend that patients with MoM hip resurfacing arthroplasties should undergo routine blood metal ion testing in the absence of symptoms ‘unless the patient cohort is of concern’. They also advised that a blood or serum Cr or Co concentration of 7µg/l indicates the potential for a soft tissue reaction. It was not clear how this value had been determined but appeared to be from the same study that the previous guidance from 2011 had drawn upon. The source data was generated from a cross-sectional study[97] which compared patients with failed MoM hips to control patients with ‘well-functioning’ MoM hips. Several patients assigned to the ‘well-functioning’ hip group had blood Co concentrations in excess of 5 µg/l and one was in excess of 50 µg/l. All our work so far had shown that it was clearly undesirable to designate patients with high Cr Co levels as having “well-functioning” hips.

Tribological theory states that as articulating surfaces become rougher they are less likely to harness a sufficiently thick fluid film[98]. If surfaces become rougher, therefore, they are less likely to develop a beneficial lubricating film and wear accelerates. With this in mind, we did not feel it appropriate to label a patient’s hip as “well-functioning” simply because of the absence of gross clinical symptoms. Our clinical experience had suggested to us that it was quite likely that patients exposed to increased metal debris may show temporary ‘tolerance’ to the stimulus and a certain time period must elapse or a threshold exposure be reached before an immune response is established and symptoms develop. Several publications had in fact already included descriptions of patients with high metal ion loads who were initially pain-free but went on to develop pain a number of years later[99][43]. Furthermore, increased metal ion levels in asymptomatic patients may be associated with underlying pathology, including osteolysis[100]. It is well recognised that osteolysis can be silent, often only manifesting in the form of radiographic changes or pain secondary to loosening of components.
For the above reasons, we conducted a study to document the clinical course of asymptomatic patients with certain blood metal ion concentrations. We found that blood tests are a clinically useful indicator of the risk of the development of ARMD in asymptomatic patients. Given our clinical experience this was not an unexpected finding. Compatible with our experience and those reported from other centres, the results also showed that women are more likely than men to develop ARMD, even when exposed to the same concentrations of blood metal ions. Furthermore, while blood metal ion concentrations were invariably associated with the development of osteolysis, it appeared that ASR patients were more likely to develop lymphocytic dominant ARMD as opposed to macrophage dominant cellular reactions.

Conclusions

MoM hip arthroplasty has a long history. Modern MoM hips were developed on the founding belief that certain first generation designs had shown extremely good survival rates. An examination of the evidence supporting this belief reveals insufficient data to make firm conclusions.

The MoM bearing surface has long been a source of anxiety for those who believe that long term exposure to metal debris may be detrimental to the systemic health of recipients of MoM hip arthroplasties[42]. There remain a limited number of case reports of neurological and cardiological complications secondary to metal ion intoxication at present although this situation may change in the next few years as more physicians investigate the issue.

It is becoming clear however that local reactions to metal debris currently present a huge problem to patients and healthcare providers throughout the world. The aim of this research was to investigate why ARMD occurs and to identify at risk patients.

ARMD occurs more commonly when MoM prostheses are wearing at greater than expected rates. There remains no convincing evidence in the literature describing severe soft tissue injury or osteolysis in association with a low wearing implant (a wear rate of less than 1mm$^3$ per year of bearing surface wear with no contributing debris from a taper junction).

The extent of osteolysis appears to be strongly related to volumetric wear. In this investigation, all patients with bearing surface wear rates greater than 20mm$^3$ per year were found to have gross metallosis and acetabular and or femoral osteolysis. A number of these patients described only minor discomfort or no pain at all. Some hip resurfacing patients presented to the emergency department with acute femoral neck fractures. Periprosthetic tissues of patients suffering osteolysis as opposed to soft tissue injury are more likely to exhibit histiocyte dominant responses.

Soft tissue damage appears to be more dependent on host factors. In patients where there is extensive destruction of the abductor musculature copious sterile fluid effusions are frequently found. Histological examination often identifies an intense lymphocytic dominated response. It seems that once a certain threshold is breached in a certain patient an immune cascade develops and this response is strongly associated with soft tissue damage. As the extent of tissue injury does not appear to be dose related to metal debris (aside from a categorical relationship ie abnormal
versus expected wear), conclusions drawn from different research centres are not in harmony. The situation is further complicated when hip resurfacings and THRs are not analysed separately.

Excessive wear and, by extension ARMD, is relatively well understood in MoM hip resurfacings. Any factors that predispose a hip resurfacing to rim loading and edge wear can be considered risk factors for ARMD. Smaller diameter resurfacings, cups placed in sub optimal orientations and designs with low cup coverage angles greatly increase the risk of increased local and systemic chromium and cobalt exposure as well as the development of soft tissue and bony injury. Women are at greater risk of excessive metal debris exposure than males due to the fact that on average they have smaller diameter cups which are more likely to be implanted at higher angles of anteversion and inclination. Furthermore, females appear to more readily mount a lymphocyte dominated response and are at greater risk of ARMD when exposed to similar amounts of debris as males.

Metal ion testing appears to be a useful clinical tool in the follow up of patients with hip resurfacings. In the United Kingdom, the following guide appears to be appropriate in the consideration of blood Co results one year post implantation:

< 2µg/l Co concentrations close to normal background levels
2 - 5µg/l Equivocal test
5 - 10µg/l High sensitivity and specificity for abnormal wear process
10 - 20µg/l 100% specificity for abnormal wear and patient at greatly increased risk for development of ARMD even if asymptomatic
> 20µg/l Macroscopic metallosis of periprosthetic tissues and patient is likely developing osteolysis

ARMD is more common in patients with resurfacing THRs than in patients with pure resurfacings. THR patients who develop ARMD are on average exposed to a smaller total amount of wear debris in volumetric terms than resurfacing patients who develop ARMD. This makes the interpretation of metal ion results more difficult. Updated Kaplan Meier analysis of the ASR THR population at University Hospital of North Tees indicates 45% failure at 8 years. At present the simplest approach to the management of patients with large diameter THRs appears to be close surveillance and to categorise blood Co results into "normal" (< 2µg/l) and "abnormal" (≥ 2µg/l). There appears to be a common mechanism of taper failure between various manufacturers' stems which is linked to the use of larger heads in combination with trunnions optimised to accommodate ceramic heads.[101] When one also considers that there are well documented reports of ARMD in asymptomatic patients[100], my personal recommendation is that all patients with "abnormal" blood Co concentrations should undergo baseline cross sectional imaging. Unfortunately, ongoing investigations suggest that there will be a catastrophic number of ARMD failures worldwide over the coming years.
References


58. DePuy ASR XL 510k FDA application.


72. Sidaginamale RP, Joyce TJ, Lord JK, Nargol AVF, Langton DJ. Blood metal ion testing is an effective screening tool to identify poorly performing metal on metal bearing surfaces. Bone and Joint Research in press.


82. Garbuz DS, Tanzer M, Greidanus NV, Masri BA, Duncan CP. The John Charnley Award. Metal-on-metal hip resurfacing versus large-diameter head metal-on-metal total hip arthroplasty: a randomized clinical trial. Clin Orthop Relat Res 2010;468:318-325


Chapter One
The Influence of Age and Sex on Early Clinical Results After Hip Resurfacing
An Independent Center Analysis

Simon S. Jameson, MRCS, David J. Langton, MRCS, Sonali Natu, FRCPath, and Tony V.F. Nargol, FRCS (Tr&Orth)

Abstract: Patient selection is critical to the excellent medium-term clinical results after hip resurfacing. We assessed the influence of age and sex on early survivorship and functional outcome by comparing 100 female hips resurfaced with male hips resurfaced for the same period. In patients older than 55 years, Harris hip score improved to 97.4 in males compared with 91.2 (P < .01) in females with a revision rate of 2.2% and 7.4%, respectively. There was no correlation between age and functional score. Three percent of females and 1.3% of males sustained a femoral neck fracture. Hip resurfacing provides excellent early functional recovery in males and females. However, the revision rate in older females is high. Changes to surgical technique may minimize the risk of early failure in this group. Key words: hip resurfacing, metal-on-metal.

© 2008 Published by Elsevier Inc.

Total hip arthroplasty (THA) is a successful operation for middle to late age with more than 90% to 95% survival rates at 10 years [1,2]. Implant survival after THA in younger patients is lower [1,2]. Despite high-wear characteristics with earlier generations, metal-on-metal (MoM) hip resurfacing has become popular for the treatment of younger patients [3]. Resurfacing is recommended for active people who would otherwise receive a conventional THA but who are likely to outlive all currently available THA implants [4].

The advantages of hip resurfacing over THA include proximal femoral bone preservation, optimization of stress transfer to the proximal femur and reproduction of normal hip biomechanics, greater implant stability, and perceived ease of revision [5,6]. Clinical function is thought to be better with a large femoral head [7] and patients report less limp [8]. Large diameter bearings also reduce the risk of dislocation, with reports of less than 1% [9]. Medium-term results are encouraging, with 94.4% to 99.8% survivorship at 2 to 6 years for the treatment of young patients [10-12]. In a retrospective age and sex-matched series, resurfaced patients had significantly better activity levels and quality of life than THA patients [13].

The ideal candidate for hip resurfacing is a young active male with normal proximal femoral bone geometry and quality. Absolute contraindications include elderly patients with osteoporotic bone and impaired renal function [6]. Patient selection appears to be critical to the success of these implants. Most patients who undergo THA are female, in all age groups [1]. However, female numbers are low in the published resurfacing literature (21%-27%) [10-12].

Fractured femoral neck is the most common complication of resurfacing with reported rates of 0% to 6% [10-15]. Concerns regarding female hip resurfacing have focused on the risk of fracture as a result of poorer bone density. Some groups suggest a lower age ceiling for
female resurfacing [11]. Beaulé et al [16] showed that females may be at higher risk of early failure.

The influence of age and sex on clinical outcome and survivorship after hip resurfacing is still not clear. With the exception of Siebel et al [14], who compared early revision rates, to our knowledge there are no reported comparisons of hip resurfacing in males and females. Much of the resurfacing literature principally documents the outcome and survival of young men. Ng et al [17] showed that females have significantly lower functional scores after THA when compared with men. This study aims to give a clearer picture of the effects of age and sex on outcome after hip resurfacing.

Materials and Methods

Between April 2004 and January 2007, the first 254 consecutive hips in 231 patients were resurfaced by the senior surgeon (AVFN) at an independent center. AVFN is an experienced hip resurfacing surgeon with a previous series of more than 150 Birmingham hip resurfacings (BHR).

All patients were followed up prospectively. The mean age at the time of surgery was 56 years (range, 28-74 years). There were 154 male and 100 female hips. Mean follow-up time was 28 months (range, 12-45 months). The diagnosis was primary osteoarthritis in 164 patients (67%). There was no history of osteoporosis or renal failure in this cohort. Patient demographics, diagnosis, American Society of Anesthesiologists (ASA) grade, and Charney walking grade were recorded.

There was no specific selection criteria, but young and active patients with end-stage primary and secondary osteoarthritis were considered for resurfacing. Bone quality was assessed radiographically, and patients offered a resurfacing were told that a final decision on the type of implant (resurfacing or THA) would be made based on the intraoperative findings. Large cysts and avascular necrosis of the femoral head were not contraindications. Patients were not routinely scanned for osteoporosis. No patients reported previous metal sensitivity reactions.

The articular surface replacement (ASR) resurfacing prosthesis (Deputy International Ltd, Leeds, United Kingdom) was used for the duration of the study. Design changes over earlier designs (such as the BHR) include a lower radial clearance to reduce wear and a shallower cup to preserve acetabular bone stock [18]. Early outcome data from the inventors' series of 300 has been reported [14]. The design of the prosthesis and the surgical technique did not change throughout the study period.

At the time the patient was listed for theater and during the consenting process, patients were made aware of the limited knowledge regarding the implant and the results of contemporary MoM articulations. Potential problems and theoretical risks associated with metal ions were discussed. Femoral neck fracture was quoted as 1.5% to 2%. All patients agreed to long-term clinical and radiologic follow-up.

All procedures were performed by the senior author. Patients were given a second generation cephalosporin antibiotic peroperatively and 2 further doses postoperatively as a prophylactic measure. Ten patients had bilateral simultaneous resurfacings. All other patients underwent a unilateral resurfacing procedure. The posterior approach was used. The external rotators were detached. In all cases, the acetabular component was implanted first with a press fit by underreaming the acetabulum by 1 mm, according to manufacturer's recommendations. The femoral head was sized according to neck diameter and then carefully reamed to avoid neck notching. The femoral head was then prepared. Any cysts were curettaged and grafted (using reamed bone), and the head cleaned with jet lavage. Low-viscosity vacuum-mixed cement was poured into the femoral component that was then firmly tapped onto the reamed femoral head. The head was then reduced, and the external rotators reattached. The wound was closed in layers without a drain.

To reduce the risk of a thromboembolic event, patients were given calf pumps during their inpatient stay and compression stockings from the day of surgery for 6 weeks. High-risk patients were given subcutaneous low-molecular-weight heparin for 6 weeks, as per hospital guidelines. Postoperatively, patients were allowed to fully weight bear immediately unless there was an intraoperative concern. For example, in patients with large cysts or if the femoral neck was notched, touch weight bearing was advised for 6 weeks. All patients were told to avoid high-impact sports such as jogging, tennis, and squash for 6 months. After this period, all restrictions were lifted.

Clinical assessment was made preoperatively and at 1 year postoperatively using the Harris hip score (HHS) and the University of California Los Angeles (UCLA) activity score. Hips revised before the 1 year review were excluded from postoperative functional assessment. At follow-up, patients were also asked to rate their satisfaction with the procedure (numerical scale 0-4). Patients scoring 3 or 4 were deemed "highly satisfied." Complications were recorded. Acetabular inclination angle was measured on postoperative radiographs.

Patients were divided into 4 groups based on age (<55 and >55 years) and sex. Of the younger patients, 45 hips were female and 59 were male. The older groups consisted of 55 females and 95 males (Table 1).

The Mann-Whitney U test and analysis of variance testing was used to calculate differences between groups. Spearman rank correlation coefficients between HHS and various clinical parameters were calculated. Kaplan-Meier implant survivorship was plotted for each group. All analyses were performed by the SPSS for Windows (version 16.0) statistical package (SPSS Inc, Chicago, Ill).

Results

There were no patients lost to follow-up, and there were no deaths. Hips of 8 patients have been revised
Five patients sustained femoral neck fractures (3.3%). Five patients sustained femoral neck fractures (2.0%). Hips of 3 patients were revised for persisting severe pain (1.2%). There was one deep wound infection that was successfully treated with a wound debridement and intravenous antibiotics. There were no known thromboembolic events.

At 1 year review, overall mean HHS for males was 97.0 and for females, 90.7. Mean acetabular inclination angle in males was 47.4° and in females, 48.9° (overall, 48.0°; range, 31°-63°).

Harris hip score in young males was 96.4, and UCLA activity score was 7.9. There was a 100% implant survival rate (Fig. 1). All patients reported high satisfaction. In the older males, HHS improved to 97.4 and UCLA activity score to 7.5. Ninety-seven percent were highly satisfied. There were 2 femoral neck fractures. One patient was left with a superior neck notch after an intraoperative femoral jag failure. Despite limited weight bearing, he sustained a neck fracture at 3 months postoperatively. The second male had a spontaneous fracture at 12 months. Both were revised before 1 year review.

In young females, mean HHS improved to 95.6 and UCLA activity score to 6.9. There was one spontaneous neck fracture at 12 months. Eighty-two percent were highly satisfied. In the older females, mean HHS was 91.2 and UCLA activity score 6.3 at the 1 year review. Eighty-four percent were highly satisfied. There were 2 spontaneous fractures at 3 and 12 months. The second patient was reviewed at 1 week before their fracture and reported no problems and a HHS of 100. Both patients had a femoral neck notch intraoperatively. Fig. 1 shows a Kaplan-Meier survival plot for the 4 patient groups. Results are summarized in Table 1.

The older female group had significantly lower HHS when compared with the 2 male groups (P < .01). However, there were no significant differences between the groups when improvement in HHS was calculated. In all patients, there was no significant correlation between HHS and age (r = 0.042; P = .29). This was also true when females were analyzed separately (r = 0.041; P = .36). Young females had a significantly higher body mass index (BMI) than older females (P = .04), but there was no significant correlation between HHS and BMI in all females (r = 0.190; P = .08). Older females had significantly lower preoperative UCLA activity scores compared with males (P = .01 for both groups) and significantly lower postoperative scores when compared with all groups (all P < .05). There were no significant differences in functional outcome between patients with osteoarthritis and those with other diagnoses (P = .84) and between Charnley walking grade A patients and other grades (P = .22). Except for sex and implant size (a function of sex), there were no other significant differences in HHS between the groups, and no other significant correlations.

Four patients had a HHS less than 50 at follow-up (all female). One patient was satisfied with the resurfacing but was restricted by bilateral knee osteoarthritis. Hips of the other 3 patients were revised because of persisting severe groin pain. The original indication for resurfacing was...
osteoarthritis in all 3 patients. These patients classically improved after surgery but then presented complaining of difficulty in getting into and out of their car and pain restricting flexion beyond 30° on straight leg raise. The first patient was 56 years old with a BMI of 31 kg/m² and an ASA grade of 1. She underwent local anesthetic injection of her psoas tendon for a possible tendonitis, but despite an initial improvement, symptoms subsequently worsened. Hip aspiration was performed to rule out infection. This revealed copious amounts of green-gray colored aseptic fluid. A decision was taken to explore and revise the implant because of possible infection. At revision, there was gross swelling of the pseudocapsule, and a similar green-gray fluid was found surrounding the implant and tracking into the psoas tendon (Fig. 2). The other 2 patients were 36 and 61 years old, both ASA 1, with BMIs of 26 kg/m² and 38 kg/m², respectively. They both had similar findings at aspiration and revision, with varying degrees of tissue necrosis found around the implant. The acetabular cup was found to be loose in one patient, and radiographic review subsequently identified subtle cup migration. Histopathologic analysis of tissue samples from these patients consistently showed tissue necrosis, cuffs of lymphocytes (Fig. 3), and sheets of histiocytes containing large unidentified black particles, probably metal particulate debris. In addition, the tissue from around the loose implant showed high numbers of lymphocytes and vascular tissue ulceration. There were no organisms grown from any of the fluid or tissue samples in any of the patients.

All patients with neck fractures had successful femoral revision to an uncemented THA comprising a modular ASR XL head on an S-ROM (Depuy International Ltd, Leeds, United Kingdom) femoral stem. The acetabular cup was left in situ. Hips of patients revised for pain received a THA with an S-ROM stem and a ceramic Pinnacle (Depuy International Ltd, Leeds, United Kingdom) head and cup. At latest follow-up review, all 3 patients had a HHS greater than 80.

Discussion

Despite lower postoperative HHS and satisfaction scores in females compared with male patients, improvements in functional scores after resurfacing do not significantly differ between the sexes. Ng et al [17] has

---

**Fig. 1.** A Kaplan-Meier survival plot of all groups.

**Fig. 2.** Aseptic fluid released from the pseudocapsule at revision of the MonM ASR resurfacing implant in patient with ALVAL.

**Fig. 3.** Photomicrography of tissue samples from revision of a loose MonM ASR (hematoxylin and eosin). A. The thick cuff of perivascular lymphocytes that is typically seen in layer 3 or the vascular layer of neocapsules formed around metal on metal prostheses (×20 magnification). B. Large, dense perivascular cuffs of lymphocytes (×40).
previously reported similar significant differences in postoperative scores after THA. The postoperative HHS in both the male groups was extremely encouraging and justifies its continued use. The results in our male cohort mirrors the published data for hip resurfacing (in which there are only small numbers of females). In 154 hip resurfacings performed on male hips, there were only 2 fractures in the older group and 100% survivorship in the younger group. However, as expected, older females (who have not traditionally been candidates for resurfacing) did have a higher complication rate and lower mean HHS than the other groups. It must be stressed that, despite this, 83% of this group have a good or excellent HHS (>80), and 84% were highly satisfied with the outcome.

There were no statistically significant associations with age, ASA status, Charnley grade, diagnosis, and the functional outcome. Poorer bone quality in the older female patient has been cited as a cause of fracture in this group. In our study, the rate of fracture was 3.6% in the older female group. However, the overall fracture rate was 2.0%. These figures are comparable to reports in the literature from independent units with other resurfacing implants and similar to the inventors’ experience of 300 ASRs (1.7%) [14]. Shimmim et al [15] showed that female patients had twice the risk of fracture as males, and notching of the neck was a risk factor for fracture. In 3 cases (of 5 fractures), the neck had been notched at the time of surgery. We question the rationale of continuing with a resurfacing procedure once a notch to the femoral neck has been identified and suggest that surgeons should consider conversion to THA at this point.

Three female patients were revised because of persisting undiagnosed pain. Although an improvement was apparent after treatment of tendonitis in one, this was short-lived. Microbiological analysis of tissue samples was negative. In the 1970s, there was a high incidence of prosthesis failure after implantation with MonM bearings [19]. Obliterative vascular changes, local bone necrosis, and loosening of the metal prosthesis have previously been described around metal implants [20]. More recently, Jacobs and Hallab [21] described an aseptic lymphocyte-dominated, vasculitis-associated lesion (ALVAL) in the tissue around MonM bearings. Characteristic microscopic appearances have been seen around many different MonM implants and may represent a natural reaction to metal wear debris [21]. The histologic changes identified on samples taken from around the loose implant in our series showed ALVAL. The other 2 patients with revised hips had well-fixed cups, and tissue analysis revealed changes similar to ALVAL. These changes may be attributable to high levels of metal wear debris. Running-in wear can produce a gray fluid [22], and this may explain the macroscopic findings, but not pain. It is more likely that symptoms are the result of an excess of inflammatory fluid around the hip joint resulting in tissues stretched under tension.

The failures in our group caused by a hypersensitivity reaction to “normal” levels of metal wear particles, or is this a local reaction to excess metal debris from a high-wear bearing, eventually resulting in osteolysis? The answer to this is not clear. Vendittoli et al [23] analyzed cobalt and chromium levels in patients resurfaced with an implant similar in design to the ASR. Their results showed that inclination of the acetabular component correlated significantly with metal ion levels at 1 year postoperatively. High cup inclination angles result in component malalignment and possibly accelerated wear. An open cup angle of more than 50° has been associated with dramatically elevated metal ion blood levels in patients with MonM bearings [24]. Given the spread of inclination angles seen in the literature (23°-63°), many surgeons may inadvertently place their cups at angles more than 50°. The range of inclination angles was similar in our patients. Inclination angles were 50°, 54°, and 55° in our problem patient group. However, 98 patients had a cup inclination angle greater than 50° without symptoms.

The population studied in this article represents a normal district hospital workload rather than the specially selected patient groups described in much of the resurfacing literature. The senior surgeon was prepared to offer resurfacing to patients who may fall outside the inclusion criteria of specialist centers because of his experience and the perceived benefits. An upper age limit was not set, but the decision to offer resurfacing was based on physiologic age. Moreover, this is a consecutive series of all ages, all diagnoses, and all activity levels included rather than a specific subset of patients from a larger study. To this end, we feel the early results of this new prosthesis are extremely encouraging in male patients. Our fracture rates across all groups are comparable to the published data. However, we have some reservations about resurfacing hips of older females until there is clearer data on risk factors for implant failure. Ultimately, these problems may be technique-specific rather than sex-specific.

Implant survival in the female group may be compromised by inadequate patient selection, inaccurate surgical technique, implant problems, or a combination of these factors. Surgical inexperience and higher failure rates in resurfaced hips of females may explain some disappointing joint register results [25]. Awareness of notched necks and early conversion to THA and further investigations of cup positioning may ultimately reduce the failure rates. Our data help to understand the excellent results published in the literature in principally younger male cohorts. We are committed to long-term follow-up of our patients to enhance our knowledge of these implants, improve our selection process, reduce failure rates, and ultimately perform the most suitable operation for each individual patient.

References

Chapter Two
Cup Anteversion in Hip Resurfacing: Validation of EBRA and the Presentation of a Simple Clinical Grading System

David J. Langton, MRCS,* Andrew P. Sprowson, MRCS,† Dhirendra Mahadeva, MRCS,† Sharad Bhatnagar, MRCS,† James P. Holland, FRCS (Tr & Orth),‡ and Antoni V.F. Nargol, FRCS (Tr & Orth)*

Abstract: The use of large metal on metal bearings has led to a reduction in the risk of dislocation post hip arthroplasty. Because of this, and also because of the technical difficulties associated with resurfacing surgery in particular, it could be argued that a less meticulous approach to acetabular cup placement has developed in comparison with conventional metal on polyethylene arthroplasty. Resurfacing cups may produce significant clinical problems when placed at the extremes of version, including increased production of metal debris and psoas tendinitis. Presented in this article is evidence that EBRA software (Einzel-Bild-Roentgen-Analysis, University of Innsbruck, Austria) can be used to reliably assess the version of resurfacing cups, when radiographs are of sufficient quality. The cups have characteristic appearances when placed at the extremes of version. These characteristics can allow the surgeon to identify poorly positioned cups without the use of software. Keywords: hip, arthroplasty, resurfacing, acetabular, anteversion, EBRA.

© 2010 Elsevier Inc. All rights reserved.

Acetabular component orientation during conventional total hip arthroplasty is critical for stability, post operative range of motion, polyethylene wear, pelvic osteolysis, and component migration [1-5]. Much attention has already been focused on the anteversion of the acetabular cup with the primary aim being to define the most stable mechanical situation in order to reduce the risk of dislocation. No absolute consensus has been achieved, with suggested angles ranging from 0° to 30°, although most surgeons regard 10° of radiological anteversion to be insufficient and 30° to be excessive [6-11].

The greater stability afforded by larger components [12-14] has anecdotally led to a misplaced acceptance of less meticulous cup orientation during resurfacing arthroplasty. Recently, however, the importance of resurfacing cup position has been established with the recognition that higher inclination angles are associated with an increase in blood metal ion levels [15]. Recently published work has identified a positive relationship between increasing anteversion and blood metal ion levels [16]. Cups with little or no anteversion may also be associated with increased metal ion levels and other problems including psoas irritation [17,18] and reduced flexion due to anterior impingement [19]. A grading system which would allow the surgeon to recognize a cup at the extremes of version may provide guidance for utilization of further diagnostic tests or therapeutic intervention.

EBRA (Einzel-Bild-Roentgen-Analysis, University of Innsbruck, Austria) is a method of assessing radiological anteversion from pelvic anteroposterior (AP) radiographs using a computer software package [20]. EBRA allows the user to input pelvic/spinal landmarks to compensate for changes in the position of the pelvis. The software has previously been validated in numerous hip arthroplasty studies [20-21]. EBRA has also been used to assess femoral component migration in resurfacing arthroplasty [22].

We sought to determine whether EBRA can accurately assess radiological anteversion of commonly used resurfacing implants. We also wanted to investigate the possibility that certain cup orientations have hallmark appearances allowing a clinically adequate assessment of version to be made by eye.
Methods

We obtained the unused femoral and acetabular components of one ASR (Articular surface Replacement, DePuy, Leeds, United Kingdom) and one Birmingham hip resurfacing (BHR) device (Smith and Nephew, Warwick, United Kingdom) (cup sizes, 56 and 52 mm, respectively). Plasticine moulds were fashioned to imbed the acetabular components in varying inclinations and versions on a radiolucent table. A horizontal grid was marked using the horizontal reference laser of the x-ray apparatus. Radio opaque markers were placed around the implant to simulate position of the pubic rami, pelvic brim, obturator foramen and pubic symphysis. The x-ray beam was center over the pubic symphysis before each exposure [23]. Because of the absence of the soft tissue envelope, the x-ray intensity was reduced to 60 kV 12.5 mA/s in order to recreate the sharpness of the implant images obtained in vivo. After each AP exposure, a lateral image was captured with the x-ray beam directed perpendicular to the axis of the cup. From this second image, the version of the cup could be determined accurately using standard digital IMPAX software (IMPAX AGFA IMPAX ES Web 1000 version 5.1, Agfa-Gevaert Group, Mortsel, Belgium) by measuring the angle subtended at the intersection of the vertical and the cup rim (see Fig. 1). To recreate the joint in vivo, the femoral component was placed in the cup and fixed in the mould at an anteversion angle of 10° (measured relative to the plane of the radiolucent table) and 140° in the anteroposterior plane relative to the horizontal reference line.

Twenty-five radiographs were obtained with the ASR components in non sequential degrees of inclination and versions. The radiographs were labelled A1 to A25. This process was then repeated using the Birmingham implant with the radiographs being labelled B1 to B25.

Five of the authors then used EBRA individually to determine angles of anteversion from 15 ASR and 15 Birmingham images randomly selected from the pool of 50 images. One author is highly experienced with the use of the software while the remaining four were first time users. All authors were blinded to the angles of version obtained from the lateral views. The measurements were recorded and compared to the angles of version recorded by a separate observer from the secondary images of the implants. Paired t tests were used to identify significant differences between observer EBRA measurements and those obtained from the lateral images. Significance was taken as \( P < .05 \). Statistical analysis was performed using Windows SPSS software, version 16.0. Each individual observer measurement was plotted against values obtained from the lateral radiographs on Bland Altman charts to assess the suitability of the software for clinical use [24]. Repeated measurements to assess intraobserver error were not performed.

A second study was carried out to assess the effectiveness of a clinical grading system. The grading system was developed by performing a series of radiographs of the resurfacing cups positioned in increasing degrees of anteversion and observing the difference in the appearance of the cup vertices. "Vertices" refer to the protuberant apices of the cup above the femoral component. In a neutral/near neutral version, the vertices appear sharp and clearly defined (see Fig. 2). As can be seen from Figs. 2 and 3A, the ASR has a double step rim design. Between 10° and 20° of anteversion, the vertices take on a more rounded appearance (Fig. 3B). As version increases, they lose their definition to an increasing extent (Fig. 3C) until they become completely obscured at anteversions greater than approximately 30° (Fig. 3D).

All the ASR and Birmingham radiographs were shown to eight registrars and two consultants in the orthopedic department. They were asked to grade the extent of version using the following system:

1. Cup vertices clearly visible and sharp. Lines joining the superior and inferior cup vertices with the respective cup head vertices (Fig. 3A) intersect within the femoral head (-10 to 10° of version). For the double stepped appearance of the ASR rim, the inferior vertices are selected for this process.
Fig. 3. (A) Grade 1 anteversion (−10° to 10°). A line connecting the cup vertex to the cup head junction is continued inferiorly to intersect the opposing line created by joining the same points on the other side of the cup. (B) Grade 2 anteversion (10°-20°). (C) Grade 3 anteversion (20-30°). (D) Grade 4 anteversion (>30°).

2. Cup vertices clearly visible but slightly rounded. Vertex lines intersect between the femoral head and the mid point of the femoral stem (10°-20° of anteversion) (Fig. 3B).

3. Cup vertices obviously rounded but just identifiable. Vertex lines intersect distal to the stem of the femoral component (20°-30° of anteversion) (Fig. 3C).

4. Cup vertices impossible to identify (>30° of anteversion) (Fig. 3D).

Given that in this particular study there is a "gold standard" measure of radiological anteversion, interobserver reliability was assessed by comparison of the grading of the cups by each observer with the grades obtained from the lateral images. A chance-corrected, weighted agreement index $k$ was calculated for each observer [25]. The $k$ coefficient adjusts for the proportion of agreement between observations by correcting for the proportion of agreement that could have occurred by chance. $k$ Coefficients range from +1.0 (complete agreement) to less than 0.0 (less agreement than that expected by chance).

**Results**

**ASR Analysis**

The observer measurement errors can be seen in Fig. 4. Paired $t$ tests identified no significant differences between each observer’s EBRA measurements and those obtained from the lateral images (Table 1). All but one of the retroverted cups were not recognized, leading to large maximal errors. Given the relative paucity of retroverted cups in clinical practice, we felt it reasonable to exclude errors arising from retroversion in the calculation of the maximum, mean and standard errors of each observer. These values can be seen in Table 2. A Bland Altman plot is shown in Fig. 5.

**BHR analysis**

The individual observer measurement errors can be seen in Fig. 6. Paired $t$ tests identified no significant differences between each observer’s EBRA measurements and those obtained from the lateral images. As with the ASR, the majority of the retroverted cups were not recognized as such. Therefore we have again excluded these measurements from the calculation of the observers’ maximum, mean and standard errors (see Table 3). A Bland Altman plot is shown in Fig. 7.

**Clinical Grading System**

The sensitivity and specificity of the clinical grading system can be seen in Table 4. Cups with version ranging from 5° of retroversion to 5° of anteversion were graded "1" accurately on every occasion. This was the same case with cups placed in greater than 35° of anteversion (grade 4). However, grading errors occurred when cups were placed with versions near the boundaries of the grades. As a result, grades of 1 and 4 showed high specificity for both the ASR and BHR cups but reduced sensitivity. The observers found it more difficult to accurately designate cups a grade of 2 or 3, reflecting the more subtle changes in the appearance of the vertices between these grades compared to grades 1 and 4 (Fig. 8). The reliability $\kappa$ coefficients for each observer’s set of grades can be seen in Table 5. Svanholm et al [26] defined $\kappa$ coefficients of 0.50 to 0.75 as "fair" acceptability and coefficients above 0.75 as having good to excellent acceptability.

**Discussion**

Although inclination may be reliably assessed using a simple AP pelvic radiograph [27], researchers have

![Fig. 4. Observer errors for all the ASR EBRA measurements.](attachment:image.png)
struggled to adequately quantify cup anteversion using simple, cost-effective methods. A recent study illustrated the inaccuracies occurring in calculation of anteversion from plain radiographs due to variation in pelvic tilt, rotation, and obliquity [27]. Even the concept of anteversion itself has caused confusion [28]. Murray made the point that when discussing “anteversion,” it must always be made clear whether this is a reference to anatomical, operative, or radiological anteversion [29]. Radiographic anteversion is measured by projecting the acetabulum axis to the plane perpendicular to the coronal plane and the plane of the acetabulum. Operative anteversion is measured by projecting the same vector on to the sagittal plane [30].

Our results show that EBRA can reliably determine the radiological anteversion of resurfacing cups to a clinically acceptable level of accuracy. It does this by creating a best fit ellipse from the landmarks inputted by the user. Because it relies on the ellipse rather than the cup itself to perform the calculations, EBRA can still be used for the ASR cup, which is subhemispheric in design. However, the software is not capable of distinguishing between ante and retroversion itself but provides 2 possible solutions, one positive and one negative (positive for anteversion, negative for retroversion) for the clinician to decide. As the results illustrate, it can frequently be difficult to correctly identify a retroverted cup. This is clearly the Achilles heel of the software. We believe this should only present a problem should the cup be placed in more than 10° of retroversion, when a malpositioned cup may be mistaken for one in an acceptable position. Cups in such poor position are rare in our experience and are likely to manifest themselves clinically with an obvious restriction of hip flexion. Cups placed within 10° either side of the neutral position (which many authors would agree is an insufficient amount of version) are identified with high sensitivity and specificity by EBRA and the clinical grading system.

A limitation of this study is the absence of inter and intra observer reliability measures. In defense of this, the object of this investigation was to determine the accuracy of the software in a controlled environment where the degree of anteversion was known. The interobserver reliability of EBRA has already been published [17,20,21].

The clinical grading system presented in this article cannot determine cup version precisely. It should be noted also that occasionally it proves impossible to create the intersecting lines due to a gross increase in cup inclination angle or excessive femoral component valgus/anteversion. However, the strength of the grading system lies in its ability to identify, with a high degree of sensitivity, resurfacing cups placed at the extremes of anteversion. The grading of these cups is independent of femoral orientation and cup inclination. Only at these extremes does cup version per se cause problems such as bone on prosthesis impingement, psoas irritation, and increased generation of metal debris.

Table 1. Significance Values for Paired t tests Between Each Observer’s EBRA Measurements and Those Obtained From the Lateral Radiographs

<table>
<thead>
<tr>
<th>Observer</th>
<th>ASR</th>
<th>BHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.443</td>
<td>0.698</td>
</tr>
<tr>
<td>2</td>
<td>0.641</td>
<td>0.823</td>
</tr>
<tr>
<td>3</td>
<td>0.885</td>
<td>0.801</td>
</tr>
<tr>
<td>4</td>
<td>0.754</td>
<td>0.621</td>
</tr>
<tr>
<td>5</td>
<td>0.796</td>
<td>0.777</td>
</tr>
</tbody>
</table>

Fig. 5. Bland Altman plot of the ASR EBRA observer measurements against those obtained from the lateral images. Horizontal broken lines denote overall mean error ± 1.96 SD (these errors exclude retroverted cups).

Fig. 6. Observer errors for all the BHR EBRA measurements.

Table 2. Observer Error Measurements for the ASR Group

<table>
<thead>
<tr>
<th>Observer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range (°)</td>
<td>9.48</td>
<td>10.45</td>
<td>8.26</td>
<td>12.34</td>
<td>9.56</td>
</tr>
<tr>
<td>Median (°)</td>
<td>-1.04</td>
<td>1.09</td>
<td>-0.19</td>
<td>-1</td>
<td>0.46</td>
</tr>
<tr>
<td>Mean (°)</td>
<td>-0.78</td>
<td>0.72</td>
<td>-0.27</td>
<td>-0.27</td>
<td>-0.01</td>
</tr>
<tr>
<td>SD (°)</td>
<td>2.34</td>
<td>3.03</td>
<td>2.28</td>
<td>3.16</td>
<td>2.47</td>
</tr>
</tbody>
</table>

These errors exclude retroverted cups.
These errors exclude retroverted cups.

[17-19]. The grading system proved to be more reliable overall for the BHR cup due to the radiographic appearance of the ASR cup rim hindering the ability to distinguish between grades 2 and 3, in particular. \( \kappa \) Coefficients showed fair to excellent reliability for all observers for both cups, however.

It has been recognized that there are major limitations in the assessment of cup version using plain film radiography. The position of the patient's pelvis in terms of pelvic tilt, rotation, and obliquity can lead to unacceptable inaccuracy if not taken into account [23]. EBRA allows the user to place guidelines over various landmarks in order to compensate for these problems. We do not, however, use obviously rotated films in the assessment of cup positions in our clinical practice and research work. Accuracy of observations (including recognition of version) increases with the number of standardized AP films available for interpretation [21].

Computed tomography (CT) has been advocated as the only reliable determinant of cup anteversion [31]. Although we agree that CT with 3D reconstruction is the most accurate method to assess cup orientation in terms of anatomical values referenced from the anterior pelvic plane, we believe that it has a number of drawbacks. Clearly, CT exposes the patients to a greater dose of radiation, it is more expensive and time consuming in comparison to plain radiographs and is not always routinely available. We also believe that, in terms of impingement, dislocation and even production of metal wear debris, assessment of the patient in the standing, weight-bearing position is preferable to images generated while the patient is supine. There is considerable variation in pelvic anteversion in the lying and standing positions of individual patients. One study found a mean difference in pelvic anteversion of 6.7° in the standing positions of 24 patients compared to their lying positions [32]. A CT examination is always performed with the patient supine, whereas impingement processes, which may be accompanied by frank dislocation, are dynamic processes which occur during changes of position. Lewinnek [10] famously described the "safe zone" for stability but later studies have suggested that dislocation is not always a complication of a poorly orientated joint [33]. Pierchon et al found the same rate of dislocation in hips in which cups were "well positioned" as oppose to those described as "poorly positioned." Cuscks which appear well positioned at the time of surgery and, also, on the CT images, may become excessively anteverted or even retroverted when the patient is standing [32].

**Conclusion**

Although dislocation does not appear to be the same threat as found in conventional hip arthroplasty, extremes of version may present new problems in large bearing metal joints. Excessive anteversion may cause reduced head coverage, posterior impingement and increased production of metal debris along with the associated theoretical health hazards. Insufficient anteversion may also lead to increased bearing surface wear.

**Table 3.** Observer Error Measurements for the BHR Group

<table>
<thead>
<tr>
<th>Observer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range (*)</td>
<td>11.83</td>
<td>11.59</td>
<td>10.68</td>
<td>11.99</td>
<td>13.1</td>
</tr>
<tr>
<td>Median (*)</td>
<td>1.8</td>
<td>2.26</td>
<td>2.76</td>
<td>1.02</td>
<td>1.9</td>
</tr>
<tr>
<td>Mean (*)</td>
<td>1.11</td>
<td>0.78</td>
<td>1.83</td>
<td>0.97</td>
<td>1.62</td>
</tr>
<tr>
<td>SD (*)</td>
<td>2.97</td>
<td>3.68</td>
<td>2.45</td>
<td>2.61</td>
<td>2.52</td>
</tr>
</tbody>
</table>

These errors exclude retroverted cups.

**Table 4.** The Clinical Grading System Results

<table>
<thead>
<tr>
<th></th>
<th>ASR</th>
<th>BHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>100</td>
<td>98.3</td>
</tr>
<tr>
<td>Grade 3</td>
<td>71.1</td>
<td>90.0</td>
</tr>
<tr>
<td>Grade 4</td>
<td>82.9</td>
<td>77.0</td>
</tr>
<tr>
<td></td>
<td>96.0</td>
<td>94.2</td>
</tr>
</tbody>
</table>

Total ASR gradings (n = 250). Total BHR gradings (n = 250).

**Fig. 7.** Bland Altman plot of the BHR EBRA observer measurements against those obtained from the lateral images. Horizontal broken lines denote overall mean error ± 1.96 SD (these errors exclude retroverted cups).

**Fig. 8.** Each observer grade plotted against "true grades" determined from the lateral radiographs.
and lack of bony coverage of the anterior lip of the cup may result in psoas tendonitis. The clinical grading system presented in this paper can provide a useful rapid screening tool for potential problems caused by cup malalignment and aid the surgeon's decision making process in regard to further investigations. EBRA can provide a reliable quantifiable assessment of radiological anteversion in the standing position. The main drawback is that inexperienced users may fail to identify cup retroversion. Good quality, standardized radiographs are essential for accurate measurements to be obtained.

References


Chapter Three
The effect of component size and orientation on the concentrations of metal ions after resurfacing arthroplasty of the hip


From the University Hospital of North Tees, Stockton-on-Tees, England

Increased concentrations of metal ions after metal-on-metal resurfacing arthroplasty of the hip remain a concern. Although there has been no proven link to long-term health problems or early prosthetic failure, variables associated with high metal ion concentrations should be identified and, if possible, corrected. Our study provides data on metal ion levels from a series of 76 consecutive patients (76 hips) after resurfacing arthroplasty with the Articular Surface Replacement. Chromium and cobalt ion concentrations in the whole blood of patients with smaller (≤ 51 mm) femoral components were significantly higher than in those with the larger (≥ 53 mm) components (p < 0.01). Ion concentrations in the former group were significantly related to the inclination (p = 0.01) and anteversion (p = 0.01) of the acetabular component. The same relationships were not significant in the patients with larger femoral components (p = 0.61 and p = 0.49, respectively). Accurate positioning of the acetabular component intra-operatively is essential in order to reduce the concentration of metal ions in the blood after hip resurfacing arthroplasty with the Articular Surface Replacement.

The use of metal-on-metal hip resurfacing arthroplasty continues to grow despite concerns about the potential consequences of chronic exposure to metal ions. Currently, there is no firm evidence to suggest that there is an accumulation of metal particles or an increased risk of systemic or end-organ disease in patients with metal implants. However, the potential of metal-on-metal arthroplasty to trigger mutations leading to increased susceptibility to malignant transformation has been suggested. Pre-malignant changes in the bone marrow adjacent to total hip replacement (THR) implants have been reported. Further concerns include the possibility of teratogenicity. It is likely that the systemic concentrations of metal ions are directly related to the rate of wear of the components. To date, results from tribological studies have proposed that wear rates could be reduced by decreasing the clearance and surface roughness, and by increasing the sphericity and carbon content of the components. An experimental study has suggested that larger components are more likely to produce a thicker fluid film between the articular surfaces. They also provide larger surface areas over which contact stresses are distributed.

A meaningful comparison of the results reported so far is difficult because of the different types of sample and analytical methods used. Daniel et al concluded that measurements in whole blood samples using inductively-coupled plasma mass spectrometry were more sensitive than those obtained from serum samples using graphite furnace atomic absorption spectrometry. The former has the ability to overcome any interference caused by the complex matrix in whole blood, making it a more reliable method for measuring levels of metal ions. The concentrations of metal ions in serum samples, measured with graphite furnace atomic absorption spectrometry, may fall outside the limits of detection, with values often assigned arbitrary figures.

Furthermore, analysis of serum is not sensitive enough to detect placental transmission of metal ions. This type of transmission has been proven to occur when high-resolution inductively-coupled plasma-mass spectrometry was used to analyse whole blood samples.
loading, sliding distance and the thickness of the lubricating fluid film can all influence the extent of a metal ion generation.23 We believe that the alignment of both the femoral and acetabular components is critical to this process. We expect larger implants to have improved wear characteristics and to be associated with reduced systemic ion concentrations because of improved fluid-film lubrication.

We have therefore analysed the effect of these variables in a large series of consecutive patients with a wide range of sizes of the femoral component.

**Patients and Methods**

Between April 2004 and August 2006, the senior author (AVFN) performed the first consecutive series of hip resurfacing arthroplasties using the Articular Surface Replacement (ASR; DePuy International Ltd, Leeds, United Kingdom). These patients were enrolled in a prospective, single-surgeon, independent study. Informed consent was obtained from the patients at the time of their operations. All the patients agreed to long-term clinical and radiological follow-up and to post-operative measurement of metal ion levels in whole blood. Specific exclusion criteria were bilateral resurfacing arthroplasties and other metal implants at the time of blood sampling. Patients with abnormal serum urea/creatinine concentrations were not considered for hip resurfacing, since impaired renal function can result in dramatically elevated serum chromium (Cr) and Co concentrations.24 Blood samples were obtained at more than 12 months after operation when the implants were anticipated to have passed the 'bedding-in' phase and, theoretically, had reached a steady-state wear.25,26 This series represents the first 76 patients to meet the above criteria. Their clinical details are summarised in Table I. The mean follow-up was 26 months (13 to 44).

**The ASR prosthesis.** The femoral and acetabular components are made of cast high-carbon-Cr-Co-alloy. The surface roughness (Ra) should be no greater than 0.05 μm, the deviation from full sphericity less than 10 μm and the radial clearance less than 10 μm (manufacturer’s data; DePuy International). The prosthesis is intended for hybrid fixation, with a cemented femoral component and a press-fit cementless acetabular component with an external porous coating (Porocoat, Depuy, Warsaw, Indiana) with a volume of porosity of between 40% and 50%. The acetabular component subtends less than a hemisphere to allow for shallower seating and potential preservation of acetabular bone stock. The design of the implant and the surgical technique (posterior surgical approach) did not change throughout the study period. The median diameter of the femoral components used was 46 mm in women and 51 mm in men (41 to 59).

**Metal ion sampling and analysis.** Venous cannulation was performed with a 21-gauge stainless-steel needle (Venflon, Becton Dickinson, Helsingborg, Sweden), with disposal of the first 5 ml of blood to avoid contamination. All samples were frozen and sent for blinded trace-element analysis to the Biochemistry Department of the Royal Surrey County Hospital, Guildford, United Kingdom.
The concentrations of Cr and Co in serum, whole blood and erythrocytes were measured using inductively-coupled mass spectrometry. The measured isotopes were $^{52}$Cr and $^{59}$Co. Collision cell mode, using 7% hydrogen in helium, was adopted to eliminate isotopic interferences. Blood samples were diluted 50-fold in 1% v/v nitric acid containing germanium as an internal standard. The analysis was calibrated using matrix-matched standard solutions prepared in horse blood. A series of reference materials was analysed within the same analytical run to demonstrate the accuracy of the results. The concentrations of metal ions was expressed as micrograms per litre (μg/l).

**Radiological measurements.** These were obtained from digital post-operative standing anteroposterior pelvic radiographs, with the patients in their natural weight-bearing position. Care was taken to minimise pelvic malrotation. The AGFA IMPAX ES Web version 5.1 (Afga-Gevaert Group, Mortsel, Belgium) and Einzel-Bild-Roentgen-Analyse (EBRA, University of Innsbruck, Innsbruck, Austria) software were used for analysis. All the angles described were radiological. These may have differed from anatomical angles and angles measured intra-operatively. All the radiological measurements were performed by two of the authors (D.L. and S.J.) and the mean values were used. They preceded the metal ion analyses. The following measurements were made (Fig. 1):

**Inclination of the acetabular component.** The ASR acetabular component has an obvious change in contour at the outermost limits of the convex surface. The superolateral and inferomedial apices were identified and joined. This line was extended inferiorly to subtend an angle with the inter-teardrop line. Plain radiography has been shown to be a reliable method of measuring the inclination of acetabular components. Measurements performed using IMPAX software were compared with those obtained with EBRA. Acetabular version. This was analysed using EBRA software, a method previously validated in the literature.\(^29,30\)

**The stem-shaft angle.** In order to determine the angle at which the femoral component was positioned, relative to the femoral shaft in the coronal plane, a line was drawn perpendicular to the base of the component. The angle subtended by this line and the anatomical axis of the proximal femur was defined as the stem-shaft angle. This was preferred to measurements using the prosthetic stem as a reference because the ASR stem has a tapered configuration.

**Stem-neck alignment.** This was calculated by subtracting the pre-operative femoral neck-shaft angle from the stem-shaft angle. The resulting figures were assigned positive values when the femoral component was aligned in relative valgus, and negative values if aligned in relative varus.  

**Statistical analysis.** The results were analysed using Spearman’s correlation analysis by calculating the correlation coefficient ($r$) between the independent variables (radiological measurements, gender, age, the duration of follow-up, the University of California, Los Angeles (UCLA) activity score\(^31\) at most recent follow-up and the dependent variables (Cr and Co concentrations)). Statistical significance was defined as $p \leq 0.05$. Once statistically-significant correlations had been identified, the respective variables were put forward for multiple linear regression analysis. Statistical analysis was carried out using SPSS version 16.0 (SPSS Inc., Chicago, Illinois).

**Results** The concentration of metal ions in whole blood is regarded as the most accurate representation of systemic exposure to metal.\(^11,25\) Serum and whole blood concentrations were found to correlate well with each other for both Cr and Co ($r = 0.917$, $p < 0.001$, $r = 0.851$, $p < 0.001$, $r = 0.890$, $p < 0.001$ and $r = 0.965$, $p < 0.001$, respectively; Table II). Thus, we used the whole blood ion concentrations throughout the study. The concentration of Cr in whole blood strongly correlated with that of Co ($r = 0.890$, $p < 0.001$). For convenience, therefore, the results for Cr were used in the graphical analyses. One female patient had revision at 14 months post-operatively because of aseptic loosening. Her whole blood Cr and Co concentrations immediately before revision were 5.25 μg/l and 7.82 μg/l, respectively. No other patient in the series has undergone or is awaiting revision.

**Reliability of measurements.** Measurements of the acetabular component using EBRA showed a mean inter-observer difference of 1.4° (maximum 7.6°, SD 2.6). For the inclination angle, the mean inter-observer error was 0.2° (maximum 2.4°, SD 1.3) using EBRA, and 0.4° (maximum 6.8°, SD 2.4) using the standard radiological method (IMPAX). Measurements of the inclination angle performed with the IMPAX software showed a mean difference of 0.6° (maximum 8.8°, SD 2.4) compared with those obtained using EBRA. A Bland-Altman plot of these results showed adequate inter-observer reliability. In addition, sufficient reliability
Table II. Results of analysis, using whole blood Chromium (Cr) and Cobalt (Co) levels as the dependent variables

<table>
<thead>
<tr>
<th>Independent variable*</th>
<th>Whole blood Cr levels</th>
<th>Whole blood Co levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation coefficient (r)</td>
<td>p-value</td>
</tr>
<tr>
<td>Age</td>
<td>-0.184</td>
<td>0.190</td>
</tr>
<tr>
<td>BMI</td>
<td>0.089</td>
<td>0.517</td>
</tr>
<tr>
<td>ASA score</td>
<td>-0.056</td>
<td>0.676</td>
</tr>
<tr>
<td>Duration of follow-up</td>
<td>0.024</td>
<td>0.838</td>
</tr>
<tr>
<td>Diameter of femoral component</td>
<td>-0.307</td>
<td>0.007</td>
</tr>
<tr>
<td>Inclination of acetabular component</td>
<td>0.297</td>
<td>0.011</td>
</tr>
<tr>
<td>Anteversion of acetabular component</td>
<td>0.132</td>
<td>0.258</td>
</tr>
<tr>
<td>Stem-neck angle</td>
<td>-0.159</td>
<td>0.001</td>
</tr>
<tr>
<td>Whole blood Co level</td>
<td>0.890</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Serum Co level</td>
<td>0.851</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Serum Cr level</td>
<td>0.917</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

* BMI, body mass index; ASA, American Society of Anesthesiologists

---

was found between the two techniques for measurement of the inclination angle with regard to the clinical implications. Throughout the text therefore our results are presented using the mean of the two observers (DJL, SSJ) measurements obtained by EBRA.

Size of the femoral component. In 21 patients, a femoral component larger than 51 mm in diameter was used. The remaining 55 patients received implants with diameters of between 41 mm and 51 mm. Correlational analysis did not establish a significant relationship between femoral size and blood metal ion concentrations. However, excessively high values were obtained from two patients with femoral components of 51 mm. It is not clear whether these values represented a threshold value of the diameter of the femoral components above which metal ion concentrations tended to be low. With these values treated as statistical outliers, femoral size showed an inverse relationship to Cr ($r = -0.328$, $p = 0.004$) and Co ($r = -0.315$, $p = 0.006$) concentrations. As Figures 2 and 3 show, all patients who received a 'large' ($\geq 53$ mm) femoral implant had lower concentrations of metal ions, compared with those with a 'small' ($\leq 51$ mm) femoral implant. A Mann-Whitney U test showed that there was a significant difference in the concentrations of both Cr (large 1.9, small 9.8; $p = 0.007$) and Co (large 3.0, small 18.1; $p = 0.004$) in these two groups.
The metal ion concentrations in the blood of patients with large femoral implants showed no significant correlation with any of the variables investigated (Table III). All other significant variables between the two groups were comparable except for the gender discrepancy (Table IV). For the purposes of further presentation and analysis of data, the groups remained separated.

Patients with a small femoral component. There were 28 men and 27 women who received femoral implants with a diameter of ≤ 51 mm. In men the median diameter of the femoral component was 49 mm (47 to 59) and in women it was 45 mm (41 to 51) (Mann-Whitney U test, p < 0.001). In the men and women of the small head group, significant differences were found between the mean femoral component size, angle of anteversion and time to follow-up. (Table V). Ion concentrations in the blood of these patients correlated significantly with acetabular inclination and anteversion (Table III).

Acetabular inclination. There was a positive correlation between the inclination of the acetabular component and the concentrations of Co (r = 0.439, p < 0.001) and Cr (r = 0.372, p = 0.011). This relationship was main-

**Table III.** Spearman's rank order correlation coefficient of chromium ion levels in whole blood with various clinical and radiological parameters in patients with small (≤ 51 mm) and large (≥ 53 mm) femoral implants

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Small (≤ 51 mm)</th>
<th>Large (≥ 53 mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation coefficient (r)</td>
<td>p-value</td>
</tr>
<tr>
<td>Age</td>
<td>-0.194</td>
<td>0.190</td>
</tr>
<tr>
<td>BMI</td>
<td>0.089</td>
<td>0.517</td>
</tr>
<tr>
<td>Duration of follow-up</td>
<td>0.024</td>
<td>0.838</td>
</tr>
<tr>
<td>Diameter of femoral component</td>
<td>-0.189</td>
<td>0.844</td>
</tr>
<tr>
<td>Inclination of acetabular component</td>
<td>0.372</td>
<td>0.011</td>
</tr>
<tr>
<td>Anteversion of acetabular component</td>
<td>0.326</td>
<td>0.010</td>
</tr>
<tr>
<td>Stem-neck angle</td>
<td>-0.206</td>
<td>0.631</td>
</tr>
<tr>
<td>Stem-shaft angle</td>
<td>-0.011</td>
<td>0.764</td>
</tr>
<tr>
<td>UCLA score at final follow-up activity score</td>
<td>-0.190</td>
<td>0.844</td>
</tr>
</tbody>
</table>

* BMI, body mass index; UCLA, University of California, Los Angeles

**Table IV.** Comparison of clinical data, radiological measurements and metal ion levels in whole blood in patients with small (≤ 51 mm) and large (≥ 53 mm) femoral implants

<table>
<thead>
<tr>
<th></th>
<th>Small (≤ 51 mm)</th>
<th>Large (≥ 53 mm)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>28</td>
<td>21</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Male:female</td>
<td>55</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Mean age in years (range)</td>
<td>57.6 (35 to 68)</td>
<td>54.5 (42 to 74)</td>
<td>0.149</td>
</tr>
<tr>
<td>Mean duration of follow-up in months (range)</td>
<td>26 (13 to 43)</td>
<td>26 (13 to 41)</td>
<td>0.995</td>
</tr>
<tr>
<td>Mean inclination of acetabular component (°) (range)</td>
<td>48.9 (32 to 61)</td>
<td>48.2 (36 to 62)</td>
<td>0.604</td>
</tr>
<tr>
<td>Mean antversion of acetabular component (°) (range)</td>
<td>20.2 (176 to 23)</td>
<td>18.8 (9.9 to 31.2)</td>
<td>0.454</td>
</tr>
<tr>
<td>Median whole blood concentration of Cr (range)</td>
<td>4.12 (1.5 to 69.8)</td>
<td>3.04 (1.5 to 5.5)</td>
<td>0.004</td>
</tr>
<tr>
<td>Median concentration of Co (range)</td>
<td>2.43 (0.48 to 271)</td>
<td>1.48 (0.4 to 5.6)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

* Fisher's exact test
† independent samples t-test
‡ Mann-Whitney U test

**Table V.** Comparison of clinical data, radiological measurements and metal ion levels in whole blood in men and women with small (≤ 51 mm) femoral components

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>28</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Mean age in years (range)</td>
<td>56 (35 to 66)</td>
<td>53 (36 to 64)</td>
<td>0.270</td>
</tr>
<tr>
<td>Mean duration of follow-up in months (range)</td>
<td>29 (14 to 43)</td>
<td>18 (13 to 41)</td>
<td>0.005</td>
</tr>
<tr>
<td>Mean inclination of acetabular component (°) (range)</td>
<td>47.4 (39 to 60)</td>
<td>40.1 (32 to 61)</td>
<td>0.329</td>
</tr>
<tr>
<td>Mean antversion of acetabular component (°) (range)</td>
<td>176.8 (8 to 31)</td>
<td>23.0 (7 to 37)</td>
<td>0.005</td>
</tr>
<tr>
<td>Mean diameter of the femoral component in mm (range)</td>
<td>49 (47 to 51)</td>
<td>45 (41 to 51)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Median whole blood concentration of Cr (range)</td>
<td>3.25 (1.5 to 69.8)</td>
<td>4.51 (2.29 to 35.9)</td>
<td>0.200</td>
</tr>
<tr>
<td>Median concentration of Co (range)</td>
<td>1.71 (0.59 to 271)</td>
<td>3.54 (0.48 to 155)</td>
<td>0.297</td>
</tr>
</tbody>
</table>

* independent samples t-test
† Mann-Whitney U test
tained when the genders were analysed separately (Figs 4 and 5). Figure 6 shows the non-significant relationship between the inclination of the cup and Cr ion levels in the patients with large implants, demonstrating the distinctly different behaviour between implants of small and large diameter.

*Acetabular anteversion.* There was a positive correlation between the anteversion of the acetabular component and the concentrations of Cr \( (r = 0.330, p = 0.01) \) and Co \( (r = 0.338, p = 0.008) \) in whole blood (Fig. 7). These findings should be interpreted with caution, because anteversion of the acetabular components also correlated positively with inclination \( (r = 0.290, p = 0.018) \).

*Multifactorial analysis.* The inclination and anteversion angles of the acetabular components were included in a multiple linear regression analysis model involving all patients. The correlation coefficient was 0.15 \( (p = 0.05) \).

Acetabular inclination was identified as the variable which best accounted for the variation seen in the concentration of metal ions. These results reflect the complex interaction between the dependent (ion concentrations) and independent (inclination/anteversion angles) variables. Most importantly, it is clear that none of the relationships were linear, thereby reducing the value of the correlation coefficient.
Discussion

Our results identified three variables which were associated with the concentration of metal ions in whole blood, namely the size of the femoral component and the inclination and anteversion of the acetabular component. None of the relationships was straightforward. It was clear that the larger femoral implants did not ‘behave’ in the same way as the smaller implants. The metal ion concentrations in the blood of patients receiving a large femoral component (diameter ≥ 53 mm) were significantly lower than those in patients with a smaller implant (diameter ≤ 51 mm). Overall, there was no significant difference in the alignment of the component between these two groups of patients.

A logical explanation to these findings may be the development of a thicker fluid film between the articulating surfaces of larger implants. A favourable film-thickness-to-surface-roughness ratio is desirable in order to maintain low frictional forces between the articulating surfaces and to achieve full fluid-film lubrication. This ratio is known as the lambda ratio and is proportional to the femoral diameter and clearance. Fluid-film lubrication is enhanced by making the femoral head as large and the clearance as small as practically possible without causing jamming of the articulation. Altering these variables in this direction has been shown to reduce wear in metal-on-metal bench studies. However, the smaller the clearance, the more the implants are susceptible to increased wear, if the acetabular component were to be deformed during implantation. Larger acetabular components may be more resistant to deformation, thereby maintaining adequate clearance and low-wear characteristics.

Our data suggest that there is a critical size of the femoral component above which the articular surfaces are likely to be protected from the adverse effects of malalignment of the component. Daniel et al. found lower metal ion concentrations in patients whose hips had been resurfaced with the Birmingham Hip Resurfacing (BHR) (Smith & Nephew, London, United Kingdom) arthroplasty, and no significant difference in ion levels was observed in regard to the size of the implant. However, only two femoral sizes were examined (50 mm and 54 mm), the study included fewer subjects (26) and the orientation of the acetabular component was not taken into account. A comparison of the results is given in Table VI. If a well-aligned acetabular component is defined as one with 40° (SD 5) of inclination and 15° of anteverision the metal ion concentrations in the well-aligned ASR group are comparable with those of the BHR group. The remaining 26 patients not included in Table VI fall in between having unsatisfactory anteverision or vice versa. Inclination and There are no reports of metal ion concentrations in patients with smaller BHR implants or malaligned cups.

Venditelli et al. found a weak inverse correlation between metal ion concentrations and femoral size, and a direct correlation with the inclination angle of the acetabular component when the Durom resurfacing arthroplasty (Zimmer, Winterthur, Switzerland) was analysed. The Durom is a sub-hemispherical resurfacing device with low clearance, similar to the ASR. In the small (≤ 51 mm diameter) implant group in our study, there was a trend towards higher blood levels of metal ions with inclination angles of the acetabular component above 45° and anteverision angles above 20°. Gross increases in anteverision and inclination may result in poor cover of the femoral component, thus decreasing the area available for generation of a fluid film. Articular surfaces which do not function under conditions of full fluid-film lubrication may be prone to more frequent surface-to-surface contact. Furthermore, if contact occurs, stresses may be transmitted over a smaller surface area. The fully hemispherical acetabular component of the BHR may protect against this phenomenon, since the available surface area is larger, thus promoting the generation of a lubricating fluid-film. However, this theoretical advantage of the BHR may be offset by the larger clearance of this implant.

Retrieval studies of metal-on-metal prostheses have confirmed that acetabular components with a high inclination angle demonstrate increased wear secondary to rim loading. Our results are consistent with this observation. We have also shown the importance of anteverision of the acetabular component. It is clear that increased anteverision, especially when combined with increased inclination, will lead to a reduction in the effective joint surface area. A high anteverision angle alone may present major problems in resurfacing surgery of the hip. As the head-neck unit is largely restored, a decreased head-neck ratio could lead to impingement of the femoral neck on the pos-
teroinferior part of the rim of the acetabular component during hip extension. D’Lima et al.\(^{39}\) using a computer model of total hip replacement (although we recognise the difference in the head-neck ratio in comparison with resurfacing surgery), found that at an acetabular component anteverision of 30° and femoral anteverision of between 0° and 30°, posterior impingement could occur at hip extension of 20° to 50°. An increase in the inclination angle further reduced the range of extension before impingement occurred.\(^{39}\) During the gait cycle the hip reaches extension of up to 20.4° (SD 4) in young adults and of 14.3° (SD 4.4) in healthy elderly individuals.\(^{40}\) Therefore, some patients are at a higher risk of impingement, which may culminate in microseparation,\(^{41,42}\) rim damage\(^{43}\) and possibly even subluxation. Some larger implants may be more forgiving in that respect.

Metal ions are not generated solely by mechanical wear. Another factor, which is extremely difficult to measure in vivo, is the effect of corrosion which is proportional to the surface area of the components.\(^{44}\) As the concentrations of metal ions in our study were lower with larger components, it would appear that mechanical wear is the predominant process driving the generation of metal ions.

Our study has some limitations. Our analysis incorporated the patients body mass index, rather than weight. It has been reported that the type of lubrication of a prosthetic joint is affected by joint loading and thus patient weight.\(^{45}\) Another variable which can influence lubrication, although omitted from our study for obvious practical reasons, is the viscosity of the synovial fluid.\(^{46}\) Other limitations of our study included the absence of pre-operative Cr and Co blood concentrations and information on smoking history and any dietary supplements taken by the patients. The pre-operative concentrations of metal ions are known to correlate positively with post-operative values.\(^{38}\) Finally, while the three-dimensional orientation of the acetabular component was investigated, this was not possible for the femoral component. Although the femoral stem-shaft and neck-shaft angles were measured, the radiographs were not standardised in terms of rotation of the limb. We were also unable to measure the anteverision of the femoral component.

In conclusion, the orientation of the acetabular component in both planes appears to be critical in the generation of metal ions after ASR resurfacing arthroplasty of the hip. Inclination angles greater than 45° and anteverision angles exceeding 20° were associated with higher metal ion concentrations in our study. Larger acetabular components appeared to be more resistant to the effects of malalignment of the acetabular component (Table VI). The median (and maximum) values of Cr and Co levels associated with well-positioned cups (acetabular components placed with 20° or less of version and 45° or less of inclination) were 3.0 (8.0) and 1.1 (6.5), respectively. When cups were placed beyond these boundaries, median Cr concentration was 5.9 (maximum value 69.8) and median Co was 6.7 (maximum value 271). The acetabular orientations associated with low metal ion concentrations in our study were within the limits recommended in previous studies on total hip replacement.\(^{45,46}\)

Based on the results of our study, we recommend that attention should be directed towards optimising the orientation of the acetabular component in order to reduce the metal ion load to which the patient is exposed. Serious consideration should be given to the development of reproducible techniques which will improve the reliability of correct alignment of the acetabular component during resurfacing arthroplasty of the hip.

Although none of the authors has received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article, benefits have been or will be received but will be directed solely to a research fund, foundation educational institution, or other non-profit organisation with which one or more or the authors is associated.

---

**References**


24. Daniel J, Ziaee H, Pradhan C, McMinn DJW. The definition and measurement of the component size and orientation on the concentrations of metal ions after resurfacing arthroplasty of the hip.


Chapter Four
Blood metal ion concentrations after hip resurfacing arthroplasty

A COMPARATIVE STUDY OF ARTICULAR SURFACE REPLACEMENT AND BIRMINGHAM HIP RESURFACING ARTHROPLASTIES

There have been no large comparative studies of the blood levels of metal ions after implantation of commercially available hip resurfacing devices which have taken into account the effects of femoral size and inclination and anteversion of the acetabular component. We present the results in 90 patients with unilateral articular surface replacement (ASR) hip resurfacings (mean time to blood sampling 26 months) and 70 patients with unilateral Birmingham Hip Resurfacing (BHR) implants (mean time 47 months).

The whole blood and serum chromium (Cr) and cobalt (Co) concentrations were inversely related to the size of the femoral component in both groups (p < 0.05). Cr and Co were more strongly influenced by the position of the acetabular component in the case of the ASR, with an increase in metal ions observed at inclinations > 45° and anteversion angles of < 10° and > 20°. These levels were only increased in the BHR group when the acetabular component was implanted with an inclination > 55°.

A significant relationship was identified between the anteversion of the BHR acetabular component and the levels of Cr and Co (p < 0.05 for Co), with an increase observed at anteversion angles < 10° and > 20°. The median whole blood and serum Cr concentrations of the male ASR patients were significantly lower than those of the BHR men (p < 0.001). This indicates that reduced diametral clearance may equate to a reduction in metal ion concentrations in larger joints with satisfactory orientation of the acetabular component.

Hip resurfacing using cementless acetabular components with metal-on-metal bearings for the treatment of second-stage hip osteoarthritis was introduced approximately ten years ago. The implants in use currently are popularly referred to as the ‘third-generation’ and include the Birmingham Hip Resurfacing (BHR) (Smith and Nephew, Warwick, United Kingdom). Independent intermediate clinical results have shown the BHR to be a highly successful procedure when performed to a satisfactory technical standard after appropriate patient selection.

De Smet et al. recently demonstrated that serum metal ion concentrations can be used as a surrogate marker of articular wear, owing to the positive correlation with metal ion concentrations in the joint fluid. Increased wear has been implicated in the early failure of metal-on-metal joints secondary to metal sensitivity and the development of pseudotumours. With the added theoretical concern of carcinogenicity, low levels of metal ions are desirable in resurfacing arthroplasty.

It has not yet been determined in vivo whether the cumulative effects of the macro- and microscopic differences in the design of resurfacing devices from competing manufacturers will lead to an increase or reduction in wear rates or, more importantly, a change in the clinical outcome of the patient receiving the implant. No large-scale independent comparative study of the different types of hip resurfacing has been carried out using the same laboratory for metal ion analysis and taking into account the anteversion and inclination of the acetabular component and the implant size. We sought to determine whether there is a significant difference in the chromium (Cr) and cobalt (Co) concentrations in the blood of patients surfaced with two common types of resurfacing arthroplasty. We also sought to identify the orientations of the component most strongly associated with the lowest blood metal ion levels and, in so doing, test our previous recommendations for the orientation of the acetabular component of the articular surface replacement (ASR) (DePuy, Johnson and Johnson, Leeds, United Kingdom) using a larger group of patients.
Table 1. Subtended articular surface angles increase with increasing acetabular component diameter in both devices (source: manufacturers' details and independent testing [11])

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtended articular surface angle (°)</td>
<td>148 to 160</td>
<td>158 to 166</td>
</tr>
<tr>
<td>Mean radial clearance (μm)</td>
<td>50</td>
<td>100 to 150</td>
</tr>
<tr>
<td>Wall thickness at rim (mm)</td>
<td>3.1</td>
<td>3.6/4.6</td>
</tr>
<tr>
<td>Manufacturing method of head</td>
<td>As cast</td>
<td>As cast</td>
</tr>
<tr>
<td>Surface roughness (μm)</td>
<td>0.025</td>
<td>0.029</td>
</tr>
<tr>
<td>Deviation of roundness head (μm)</td>
<td>3.4</td>
<td>0.9</td>
</tr>
<tr>
<td>Deviation of roundness acetabular component (μm)</td>
<td>3.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Carbon content</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

* ASR, articular surface replacement
† BHR, Birmingham hip resurfacing
‡ Hip/SA, cast process and heat treatment by hot isostatic pressure/surface annealed
§ high carbon content defined as ≥ 0.20%

Patients and Methods

Implants. In vitro studies [8,9] have suggested that full fluid film lubrication can be promoted by reducing the radial clearance between the articulating surfaces of the resurfacing components. It was on this basis that the 'fourth-generation' ASR was introduced with a radial clearance of 50 μm. This nominal clearance is significantly lower than the 100 μm radial clearance of the BHR.

Failure analyses and metal ion studies have shown that edge loading is an important factor leading to increased wear secondary to increased angles of acetabular component inclination [10-12]. The ASR acetabular component is sub-hemispherical in design, as opposed to the near full hemisphere of the BHR implant. De Haan et al. [13] recently showed that reduced component cover can render a joint more susceptible to the negative effects of edge loading at lower angles of inclination than expected.

Although all manufacturers use high-carbon-containing cobalt-chromium alloy, the processing of the alloy differs. Cast components may undergo post-casting heat treatments such as hot isostatic pressing and/or solution heat treatment. The significance of these treatments has been debated over the last decade. Annealing results in depletion of the surface carbides, but hip simulator studies do not demonstrate significant differences between the wear behaviour of as-cast arthroplasty and heat-treated alloys [14]. Table I summarises the important similarities and differences of the ASR and the BHR.

Patients. From April 2004 a prospective trial of the ASR was undertaken. The clinical results of the first 200 patients involved in this independent trial have already been published [15]. Metal ion analysis has been carried out at our unit on a routine basis since June 2007 for patients under the care of the senior author (AVFN). This paper presents data from patients attending for routine follow-up during the period from June 2007 to June 2008. The results of the first 76 of the 90 ASR patients included in this study have been published previously [7].

Owing to the disparity in patient numbers between the ASR and the BHR groups, 12 BHR patients attending review clinics at a second local centre who met the inclusion criteria gave consent to recruitment into this study. Each operating surgeon at this centre is a lower limb arthroplasty specialist. The posterior approach was used in each case. At this centre, all blood samples were obtained by the same orthopaedic registrar involved in the main centre study, using identical equipment and the same laboratory for analysis. For both groups of patients, the only exclusion criteria were the presence of another joint replacement and the resurfacing having been performed within 12 months of blood sampling. Patient demographics are shown in Table II.

Methods. Blood samples are obtained via Venflon, with the first 5 ml being discarded before the definitive sample is drawn. All samples are frozen and sent to the same laboratory for blinded whole blood and serum Cr and Co analysis using inductively coupled plasma mass spectrometry. The same technique of EBRA (Einzel-Bild-Röntgen-Analyse) [7,16] (University of Innsbruck, Innsbruck, Austria) analysis of standing radiographs was used to assess the BHR group and the larger ASR patient group. As well as the orientation of the acetabular component, we measured the femoral stem/shaft angle relative to the anatomical axis of the femur as previously described [7]. The orientation of the components was assessed and recorded prior to the metal ion results becoming available. For all measured parameters, all technically satisfactory weight-bearing radiographs available for each patient were analysed and the mean values calculated and used in the final analysis. An overview of joint sizes and orientations of the acetabular component is shown in Table II.

Statistical analysis. Spearman's rank correlation was used to identify any significant relationships between the independent variables (as above, as well as the Harris Hip Score [17] and the University of California, Los Angeles activity score [18] at the time of blood sampling, the time from surgery and age) and the dependent variables (whole blood and serum values...
Table II. Patient demographics, including joint sizes and acetabular component orientations

<table>
<thead>
<tr>
<th>Demographics</th>
<th>ASR</th>
<th>BHR</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>90</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Age in yrs (range)</td>
<td>55 (28 to 77)</td>
<td>51 (32 to 67)</td>
<td>0.501</td>
</tr>
<tr>
<td>M:F (% female)</td>
<td>42:38 (42)</td>
<td>44.26 (37)</td>
<td></td>
</tr>
<tr>
<td>Post-operative time in mths (range)</td>
<td>26 (12 to 44)</td>
<td>47 (14 to 75)</td>
<td>0.001</td>
</tr>
<tr>
<td>American society of anesthesiologists score</td>
<td>1.6 (1 to 3)</td>
<td>1.39 (1 to 2)</td>
<td>0.630</td>
</tr>
<tr>
<td>Femoral size in mm (range)</td>
<td>49.0 (41 to 59)</td>
<td>48.8 (38 to 58)</td>
<td>0.437</td>
</tr>
<tr>
<td>Inclination angle in ° (range)</td>
<td>49.9 (32 to 85)</td>
<td>47.0 (32 to 65)</td>
<td>0.150</td>
</tr>
<tr>
<td>Anteversion angle in ° (range)</td>
<td>20.6 (4 to 39)</td>
<td>17.4 (6 to 39)</td>
<td>0.041</td>
</tr>
<tr>
<td>Outcome scores (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harris Hip Score</td>
<td>94 (35 to 100)</td>
<td>97 (51 to 100)</td>
<td>0.076</td>
</tr>
<tr>
<td>UCLA* activity score</td>
<td>7.3 (3 to 10)</td>
<td>7.5 (3 to 10)</td>
<td>0.633</td>
</tr>
<tr>
<td>Serum metal ion levels (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr (µg/L)²</td>
<td>3.99 (0.58 to 115.0)</td>
<td>3.55 (0.65 to 190)</td>
<td>0.723</td>
</tr>
<tr>
<td>Co (µg/L)²</td>
<td>2.30 (0.38 to 228.0)</td>
<td>1.65 (1.8 to 76)</td>
<td>0.082</td>
</tr>
<tr>
<td>Whole blood metal ion levels (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr (µg/L)²</td>
<td>3.60 (1.5 to 69.8)</td>
<td>3.95 (2.37 to 39.8)</td>
<td>0.213</td>
</tr>
<tr>
<td>Co (µg/L)²</td>
<td>2.08 (0.38 to 271.0)</td>
<td>1.43 (0.63 to 147)</td>
<td>0.037</td>
</tr>
</tbody>
</table>

* ASR, articular surface replacement  
† BHR, Birmingham hip resurfacing  
* UCLA, University of California, Los Angeles  
§ median values. All other values are means. Cr, chromium; Co, cobalt

In order to determine the joint orientations most strongly associated with the lowest blood metal ion levels, the angles of inclination and anteversion were divided into subgroups. Mann-Whitney U tests were then used to identify significant differences (p < 0.05) between these subgroups. Two-sample t-tests were used to compare values from groups with normal distributions. The same methods were used to determine significant differences in independent and dependent variables between the ASR and BHR groups. Windows SPSS version 16.0 (SPSS Inc., Chicago, Illinois) was used for statistical analysis.

Results

An overview of the Cr ion data in whole blood can be seen in the box and whisker plots in Figure 1. There was a highly significant positive correlation between whole blood/serum Cr and Co concentrations (Tables III and IV), and we therefore feel it appropriate to use Cr and Co interchangeably for graphical representation.

**ASR analysis.** The results from this larger group of patients showed a significant inverse relationship between femoral size and Cr and Co concentrations (Fig. 2 and Tables II and III). In order to investigate the effect of the interaction between size and orientation of the acetabular component on the ion levels, the ASR patients were split into groups according to individual femoral sizes. A multiple regression model was then constructed for each group size, with Co as the dependent variable and acetabular component inclination and anteversion as the explanatory variables. As illustrated in Figure 3, as the size of the ASR implants
increased the corresponding $R^2$ value decreased (i.e. acetabular component orientation had less influence on Co as the femoral size increased). The relationship between acetabular component inclination/anteversion and metal ions was negligible in ASRs with femoral components > 51 mm. For the purposes of this study, ASRs with femoral components > 51 mm are referred to as large ASRs. Below this size there was a significant increase in metal ions when acetabular components were positioned at > 45° of inclination and/or > 20° of anteversion. In this larger series of patients it also became apparent that acetabular component anteversion < 10° was also associated with an increase in metal ion concentrations (Fig. 4). ASRs with femoral components ≤ 51 mm are henceforth referred to as small.

**BHR analysis.** There was a significant inverse relationship between femoral size and metal ion concentrations in the BHR patients (Spearman's rank correlation ($r$) = -0.265, $p = 0.038$, Cr) (Fig. 2). The same multiple regression model described above was used to examine the interaction between femoral size and orientation of the acetabular component on metal ion levels. As with the ASR, increasing femoral size reduced the effect of orientation of the acetabular component on metal ion concentrations (Fig. 3). Acetabular component orientation had no significant effect on BHRs with femoral components in the fourth quartile of the size range (~54 mm). For the rest of the paper, the implants in the fourth quartile of the size range are referred to as large and the remaining implants are referred to as small. Table V shows the significant differences in Cr and Co concentrations between the large and small BHR groups.

**Small BHR implants.**

*Inclination of the acetabular component.* We identified no linear correlation between the inclination of the acetabular
component and the concentration of Cr or Co ions. However, the largest three Co and Cr readings were found in the blood of patients with acetabular components implanted with > 55° of inclination. For acetabular components placed below 55° (32.5° to 54.9°), a non-significant inverse relationship to Cr and Co was observed (p = 0.08 and 0.120, respectively) (see Fig. 5).

Anteversion of the acetabular component. There was a significant positive correlation between anteversion of the acetabular component and Co (r = 0.309, p = 0.047). A similar, but slightly weaker relationship was identified between Cr and anteversion (r = 0.274, p = 0.079). The lowest median values of whole blood Cr and Co were associated with acetabular components positioned in 10° to 20° of anteversion (3.46 and 1.44 µg/L, respectively). Acetabular components placed in versions above or beyond these boundaries were associated with higher Co and Cr concentrations (p = 0.004 and 0.063, respectively) (Fig. 4).

Comparison of devices by acetabular component orientation. Each implant group was divided into sub-groups according to angles of acetabular component inclination and anteversion. Median ion values for each sub group were then calculated and plotted (Figs 4 and 5). Table VI shows the overall relative effects of suboptimal acetabular component position on the two resurfacing systems using patients from the small ASR and BHR implant groups.

Comparison of devices by gender. Men were implanted with significantly larger components than the women and acetabular components with significantly larger angles of anteversion in both the ASR and BHR groups (p < 0.001 for both variables).

Men. There were 52 men with ASR and 44 men with BHR. Whole blood Cr was found to be significantly lower in the ASR men than in the BHR patients (p = 0.012), although Co concentrations were comparable (Table VII).
Table V. Comparison of the mean, median and ranges of whole blood metal ion levels associated with the Birmingham hip resurfacing patient group split by size

<table>
<thead>
<tr>
<th></th>
<th>Large implants (n = 16) mean/median (range)</th>
<th>Small implants (n = 54) mean/median (range)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromium (µg/l)</td>
<td>3.69/3.59 (2.37 to 6.56)</td>
<td>5.92/4.04 (2.41 to 39.8)</td>
<td>0.017</td>
</tr>
<tr>
<td>Cobalt (µg/l)</td>
<td>1.28/1.32 (0.63 to 2.03)</td>
<td>5.31/1.68 (0.77 to 147)</td>
<td>0.023</td>
</tr>
<tr>
<td>Inclination (*)</td>
<td>49.51</td>
<td>470</td>
<td>0.227</td>
</tr>
<tr>
<td>Anteversion (*)</td>
<td>15.02</td>
<td>18.23</td>
<td>0.317</td>
</tr>
</tbody>
</table>

* Mann-Whitney U test for non-parametric data, two-sample t-test

As with the ASR, there was a significant inverse relationship between the size of the femoral component and whole blood Cr in the BHR patients. In the fourth quartile of ASR and BHR implant sizes, the blood metal ion concentrations were significantly lower than those found in the patients with smaller implants. Larger ASR and BHR acetabular components appear more resistant to suboptimal positioning in terms of metal ion generation. This is probably due to the thicker fluid film achieved by the larger implants as well as the greater arc of cover. Both of these factors protect against the increased rates of wear associated with edge loading.

A further variable that has been shown to affect lubrication in vitro is clearance. Rieker et al showed that wear rates can be reduced by reducing clearance as much as manufacturing processes will allow. The improved lubricating film that the reduced clearance of the ASR theoretically provides may be the reason for the significantly lower whole blood Cr values in the male ASR patients. These clearance values are, however, nominal. When the components are implanted they are vulnerable to the effects of distortion, the extent of which is impossible to measure in vivo routinely. By definition, a low-clearance device is at greater risk of accelerated wear should the combined distortion of the components be greater than expected. Whereas the median Cr value of the male ASR patients is lower than that of the male BHR group, the range of the male ASR Cr and Co concentrations is much broader, with several extremely high values.

In order to conserve bone and increase impingement-free range of movement, the ASR acetabular component is subhemispherical in design. For the smallest ASR acetabular component, the arc of cover (the subtended angle to the articular surface) is 148°, increasing to 160° for the largest. This is in sharp contrast to the cover provided by the BHR acetabular component with 158° for the smallest, increasing to 166° for the largest acetabular component. The reduced cover provides a potential explanation for the increased vulnerability of the ASR device to angles of inclination greater than only 45°. This also accounts for the increased sensitivity to suboptimal position of small BHRs compared to their larger counterparts. This is consistent with the results of De Haan et al, who found a highly significant relationship between the arc of cover of the acetabular component with increased metal ion loads observed in components with both excessive (> 20°) and insufficient (< 10°) radiological anteversion.
A theoretical model to explain the variation in Ion levels associated with the two resurfacing systems. Bergmann et al.20 showed that, in the standing position, the average hip joint contact force is directed 14° medially from the longitudinal axis in the anteroposterior plane and 16° anteriorly in the transverse plane. Using these data and the orientations of the acetabular components of the patients in this study, we calculated the distance from the centre of the theoretical contact patch to the superior anterior edge of the articular surface of the acetabular component (or its rim) for each patient in the standing position. We assumed the contact patch to be centred over the hip contact force vector. The resultant distance between the contact patch and the acetabular component rim (CPR distance) is therefore dependent on the diameter and the arc of cover provided by the acetabular component, and also the radiological angles of inclination and anteversion. We found a highly significant inverse correlation between the CPR distance and blood and serum Cr and Co.
Graph showing all articular surface replacement (ASR) and Birmingham hip resurfacing (BHR) patients in the study (Spearman rank correlation -0.49, p < 0.001) Mean ASR male patient acetabular component rim (CPR) distances represented by solid vertical black lines and the mean ASR female patients represented by the broken black line. The mean male BHR patient CPR distance is represented by the solid grey vertical line and the female BHR distance by the broken grey line. CPR values are calculated using the mean acetabular component sizes and inclination/anteversions for each patient group in the study.

There are two main limitations to this study, the first being the difference in time post-operatively at which blood levels in both implant groups, with no large values observed when the contact patch was calculated to be > 10 mm from the rim (Fig. 6). This is consistent with retrieval studies which have demonstrated increased wear secondary to edge-loaded implants. The 'wear scar' on retrieved components is commonly located at 10° to 15° from the vertical, consistent with Bergmann's in vivo joint contact area measurements. In this series a number of patients had acceptable levels of CrCo even with a calculated CPR distance < 10 mm. However, as the time from surgery to blood sampling increased, so did the probability that patients with a CPR distance < 10 mm were found to have grossly increased blood metal ion concentrations (Fig. 7). This observation may represent the temporary tolerance of the hard metal-on-metal surfaces to a suboptimal position until the joint reaches a state of accelerated wear.

We believe that the calculated CPR distance is a reliable indicator of the vulnerability of a joint to the effects of anterior edge loading and anterior subluxation. If subluxation does occur we speculate that the effects of the sharp ASR rim are likely to be particularly damaging to the femoral component. CPR distance does not, however, gauge the vulnerability of the resurfaced joint to the risk of anterior impingement and posterior edge loading/subluxation. The surgeon must strike a balance between attempting to reduce the negative effects of edge loading by implanting an acetabular component with reduced inclination and anteversion, and attempting to maintain a satisfactory range of movement of the hip. From the data available to us, we feel it reasonable to recommend that the ASR acetabular component be placed in 40° (SD 5) of inclination and 15° (± 5) of weight-bearing radiological anteversion in order to reduce levels of metal ions in the blood. We have a smaller amount of ion data for the BHR device. However, the metal ion increase in our series only appears above 55°. When considering the optimal inclination for the BHR, the lowest ion values are associated with acetabular components positioned between 45° and 55°, rather than around the lower limits (≤ 40°). We believe that low inclination angles of near-hemispherical acetabular components with reduced anteversion, coupled with the reduced head to neck ratio inherent to surface arthroplasty, may lead to anterior impingement and posterior edge loading/subluxation during common activities such as stair climbing and rising from a chair. Reduced inclination angles combined with high anteversion may increase the risk of anterior edge wear with or without posterior impingement. For this reason, we feel that the BHR acetabular component should be placed with a target inclination of 45° in mind, allowing a 5° margin of error on either side and remaining within the limits conducive to a satisfactory outcome in terms of function and metal ion levels.

With regard to standing radiological anteversion, we recommend 10° to 20°, with the same justification and allowance for a margin of error. In the senior author's patient group, the minimum Harris Hip Score of the 14 patients with acetabular components in this target area was 97 (mean score 99.0, mean femoral size 48.0 mm (42 to 54)).

Articular surface replacement patients only. Whole blood cobalt values versus distance from the centre of the calculated contact patch to the edge of the articular surface of the cup. 'X' represents samples drawn between 12 and 18 months post resurfacing. * represents blood samples drawn at 18 months and beyond.
samples were collected in the two groups. We do, however, believe that this factor is relatively insignificant compared to implant size and orientation. We believe that comparison between groups is justified as the mean time to sample collection in both groups was well beyond the bedding-in phase, where accelerated wear is said to occur.28,29 Our data suggest that suboptimally placed acetabular components will produce increasing amounts of metal debris with time, and therefore if the difference in follow-up between the ASR and BHR groups were truly significant we would expect the BHR metal ion levels to be higher. Furthermore, the Birmingham group themselves have published convincing evidence that in well-positioned joints, metal ion levels tend to decrease over time.30 This supports the idea that the reduced clearance of the ASR is indeed beneficial, given that Cr levels are lower in the male ASR patients despite the shorter follow-up period.

The second limitation is that we did not collect preoperative blood samples for metal ion analysis, a variable which Vendittoli et al11 showed to correlate with post-operative ion levels. We believe this weakness is offset by the large number of patients and the fact that all those involved in the study came from the same geographical region.

Increased blood Cr and Co concentrations following the Birmingham Hip Resurfacing procedure are associated with smaller components, acetabular component anteversion > 20° and < 10° and inclination > 55°. The female ASR patients in this series had a median Co level three times that of the female BHR group. The male ASR group had a slight but significantly reduced Cr level compared to the BHR group. The overall results suggest that the reduced clearance of the ASR promotes an improved joint lubrication regimen but this positive effect may be overridden by the increased vulnerability of the ASR acetabular component to suboptimal positioning. The BHR acetabular component is tolerant of a wider range of orientations. We have found no evidence in this study to support the idea that inclination angles < 40° equate to a reduction in metal ion concentrations.

Supplementary material
Further information relating to these findings and the measurement of the contact force and contact patch distance is available with the electronic version of this article on our website at www.jbjs.org.uk

Although none of the authors has received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article, benefits have been or will be received but will be directed solely to a research fund, foundation, educational institution, or other nonprofit organization with which one or more of the authors are associated.

References
Chapter Five
Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement

A CONSEQUENCE OF EXCESS WEAR

Early failure associated with adverse reactions to metal debris is an emerging problem after hip resurfacing but the exact mechanism is unclear. We analysed our entire series of 660 metal-on-metal resurfacings (Articular Surface Replacement (ASR) and Birmingham Hip Resurfacing (BHR)) and large-bearing ASR total hip replacements, to establish associations with metal debris-related failures. Clinical and radiological outcomes, metal ion levels, explant studies and lymphocyte transformation tests were performed. A total of 17 patients (3.4%) were identified (all ASR bearings) with adverse reactions to metal debris, for which revision was required. This group had significantly smaller components, significantly higher acetabular component anteversion, and significantly higher whole concentrations of blood and joint chromium and cobalt ions than asymptomatic patients did (all p < 0.001). Post-revision lymphocyte transformation tests on this group showed no reactivity to chromium or cobalt ions. Explants from these revisions had greater surface wear than retrievals for uncomplicated fractures. The absence of adverse reactions to metal debris in patients with well-positioned implants usually implies high component wear.

Surgeons must consider implant design, expected component size and acetabular component positioning in order to reduce early failures when performing large-bearing metal-on-metal hip resurfacing and replacement.

Although debris from metal-on-metal joint replacements may cause local tissue reaction, systemic effects have yet to be demonstrated. In the 1970s there were reports of large joint effusions and widespread tissue destruction following total hip replacement (THR) with McKee-Farrar devices. In 2005 Willert et al described the histological appearance of periprosthetic tissues surrounding failed metal-on-metal joint replacements and suggested a lymphocyte-dominant delayed-type hypersensitivity reaction, aseptic lymphocyte-dominated vasculitis-associated lesion (ALVAL), as a potential explanation. Because the clinical relevance of this testing remains uncertain, this continues to be a concern for surgeons implanting metal-on-metal devices, particularly when combined with the unknown incidence of such tissue reactions.

Pandit et al used the term 'pseudotumour' to describe cystic and solid masses associated with resurfacing devices. They estimated that 1% of patients who have a metal-on-metal resurfacing develop a pseudotumour within five years and speculated that the adverse effects could be mediated by an allergic response to 'normal' levels of metal debris, or could be caused directly by toxic concentrations of chromium (Cr) and cobalt (Co) generated from bearing surfaces experiencing abnormal wear. It remains unclear whether ALVAL is itself an integral factor in the mechanism of failure, or simply an associated observation in the capsular tissues. Without exception, the literature reports an increased incidence of these problems in women. We sought to establish the rate of failure secondary to adverse reactions to metal debris (ARMD) in our patients, to identify relationships between this mode of failure and the wear rate of the prosthetic joint, and to provide a potential explanation for the increased incidence of this in women. There is currently no clear consensus in the literature defining the boundaries of the terms metallosis, ALVAL and pseudotumour. Reports suggest that tissues examined following revision surgery often exhibit a combination of the above pathologies. For the purposes of this paper we therefore use the acronym 'ARMD' as an umbrella term to describe joint failures associated with pain, a large sterile effusion of the hip and/or macroscopic necrosis/metallosis.
Patients and Methods

The senior author (AVFN) used the Birmingham Hip Resurfacing (BHR, Smith & Nephew, Warwick, United Kingdom) implant from 2002 to April 2004, and the Articular Surface Replacement (ASR, DePuy International Ltd, Leeds, United Kingdom) thereafter. All patients were enrolled in a prospective study to monitor clinical, radiological and functional results. The femoral component of the ASR (ASR XL form), can also be used in combination with a stem (Corail or SROM, DePuy International Ltd) to articulate with the standard ASR resurfacing acetabular component. The articulating surfaces of both systems are therefore identical in terms of composition and manufacturing processes. For this reason we also considered patients with ASR THRs in this study. We have already described in full the differences in design of the ASR and BHR acetabular components.6

By January 2009, 505 ASR devices had been implanted (including 87 ASR THRs) and 155 BHRs prior to April 2004. Resurfacing procedures were carried out through a posterior approach and THRs through a direct lateral approach. The technique used for the resurfacing procedure has also been described previously.7 The majority of resurfacings were performed in patients under 65, but older men were included if they were considered biologically younger. Patient demographics can be seen in Table I. Patients were followed up at six weeks, three and six months, and annually thereafter. Harris hip scores (HHS) and University of California, Los Angeles (UCLA) activity scores9 were recorded at one year for all patients.

Radiographic analysis. Weight-bearing, digital anteroposterior (AP) pelvic radiographs were obtained post-operatively and at subsequent visits. Assessment of acetabular component position was carried out using Einzel-Bild-Roentgen-Analyse software (University of Innsbruck, Innsbruck, Austria)10,11 using all available radiographs in order to improve accuracy.12 Radiographic measurements were performed by two of the authors (DJL and SS), with mean values used for analysis.13 The theoretical distance of the articular contact patch to articular rim, which would be present in the standing position, was calculated for all patients using a method previously described.6

Metal ion analysis. Blood. From June 2007, Cr and Co ion levels in whole blood and serum were routinely analysed for each patient with a unilateral implant at a minimum of 12 months after surgery. Beyond this 'running-in' phase, implants are thought to reach a relatively steady state of wear.14,15 The sampling method has been described previously.13 All samples were frozen and sent for blinded trace element analysis to the Biochemistry Department at the Royal Surrey County Hospital, Guildford, United Kingdom. Samples were measured by inductively coupled plasma mass spectrometry.

Joint fluid. During revision surgery, joint fluid was sampled and analysed using the same technique and equipment as in collecting blood specimens. Where the...
joint capsule was still intact, a cannula was inserted through the soft tissue prior to incision and drainage of the effusion.

**Histopathological examination of tissues from revision procedures.** Specimens were received in formalin from various sites within the neocapsule. They were routinely processed, embedded in paraffin and stained with haematoxylin and eosin. Histological analysis was carried out at x2, x4, x10, x25 and x40 magnification using a Nikon Eclipse 80i light microscope (Nikon Instruments Europe BV, Amstelveen, The Netherlands). A calibrated graticule was used for measurements. The various layers of the neocapsule were studied. Layer 1 consists of the surface membrane and was graded as type 1 to 4. Type 1 surfaces have an intact pseudosynovial membrane. Type 2 surfaces show a loss of the pseudosynovial membrane. Type 3 surfaces are associated with fibrin deposition in addition to the loss of the surface membrane, and in type 4 there is gross disruption of the surface membrane with fibrin deposition and fissuring. Layer 2 is the subsynovial layer, which usually contains the inflammatory infiltrate. Layer 3 is the vascular layer, which contains the perivascular lymphocyte cuffing. Layer 4 is the deeper layers containing muscle fat and fibrous tissue. The thickness of the histiocyte band containing particulate matter in Layer 2 was measured using the graticule. The presence of perivascular lymphocytic cuffing and its average thickness in layer 3 was also recorded. Patients who experience failure of their hip resurfacings secondary to a fractured neck of femur have tissue samples sent for routine analysis to exclude pathological causes. All tissue samples from metal-on-metal revision surgery are now graded histologically as described above.

**Explant analysis.** Since October 2007, explants retrieved at revision surgery have undergone analysis using a Zeiss TSK Rondcom60A (Carl Zeiss IMT Group, Minneapolis, Minnesota) roundness measuring machine, which has a resolution >1 μm. Out-of-roundness (deviation in shape from a perfect circle) measurements were taken from three planes for each acetabular and femoral component. For acetabular components these planes were 3 mm, 7 mm and 11 mm below the rim. For the femoral components these planes were 3 mm, 7 mm and 11 mm from the 'pole' of the head. Modern manufacturing of resurfacing hip prostheses allows acetabular and femoral components to be produced with an out-of-roundness <5 μm. Explanted components with greater out-of-roundness values imply that either material has been removed locally (wear) or the component has been distorted.

**Lymphocyte transformation tests.** All patients who had undergone surgery for failure of their metal-on-metal prostheses secondary to ARMD were invited to attend a clinic for metal allergy testing. Blood samples were obtained for metal-lymphocyte transformation tests.

**Patients presenting with pain.** A total of 16 patients presented with groin pain. All had an ASR in situ. Of these, 14 were women. The symptoms and clinical findings were similar, although six patients described deterioration soon after the operation, and the rest had initially been satisfied but developed pain between two and 25 months postoperatively. Of this group, 13 patients have been revised and the remaining three are awaiting revision. This represents a revision rate of 3.2% (35 months, 8 to 57) for our ASR resurfacings and 6% (41 months 10 to 57) for our ASR THRs as a result of ARMD. Pain was localised in the groin and aggravated by straight leg raising, and a 'clicking' or 'clunking' sensation was often described. In each case the WBC, ESR and CRP were normal.

Each patient underwent aspiration of the hip under fluoroscopic guidance, yielding large volumes of yellow/green fluid which was occasionally bloodstained. There was no microbial growth in any of the samples. In one patient, radiographs revealed radiolucent lines in three acetabular zones and the joint was subsequently found to be loose at revision. In the other patients the radiographs showed no evidence of loosening.

In the first patient who underwent revision there was gross swelling of the pseudocapsule and a large volume of the fluid bathing the implant and tracking into the psoas tendon. These findings have been published previously. Similar appearances were seen at the other 12 revision operations, with varying degrees of tissue necrosis around the implant. At one revision of an ASR resurfacing the pressure of the hip fluid had created a fistula through the abductor musculature. In was a further patient a tendinous as described above.

**Statistical analysis.** This was carried out using SPSS version 16 (SPSS Inc., Chicago, Illinois). The Mann-Whitney U test was used to identify significant differences between groups with non-parametric data (Cr and Co concentrations; comparison of ARMD joint orientations with controls) and two sample t-tests were used to compare groups with parametric data. The relationship of the blood and joint fluid to metal ion concentrations was examined with Spearman’s rank correlation. Differences were deemed statistically significant if the p-value was <0.05. Survival was calculated using the Kaplan-Meier method and 95% confidence intervals were calculated.

**Results**

**Patients presenting with pain.** A total of 16 patients presented with groin pain. All had an ASR in situ. Of these, 14 were women. The symptoms and clinical findings were similar, although six patients described deterioration soon after the operation, and the rest had initially been satisfied but developed pain between two and 25 months postoperatively. Of this group, 13 patients have been revised and the remaining three are awaiting revision. This represents a revision rate of 3.2% (35 months, 8 to 57) for our ASR resurfacings and 6% (41 months 10 to 57) for our ASR THRs as a result of ARMD. Pain was localised in the groin and aggravated by straight leg raising, and a 'clicking' or 'clunking' sensation was often described. In each case the WBC, ESR and CRP were normal.

Each patient underwent aspiration of the hip under fluoroscopic guidance, yielding large volumes of yellow/green fluid which was occasionally bloodstained. There was no microbial growth in any of the samples. In one patient, radiographs revealed radiolucent lines in three acetabular zones and the joint was subsequently found to be loose at revision. In the other patients the radiographs showed no evidence of loosening.

In the first patient who underwent revision there was gross swelling of the pseudocapsule and a large volume of the fluid bathing the implant and tracking into the psoas tendon. These findings have been published previously. Similar appearances were seen at the other 12 revision operations, with varying degrees of tissue necrosis around the implant. At one revision of an ASR resurfacing the pressure of the hip fluid had created a fistula through the abductor musculature. In was a further patient a pre-operative ultrasound scan showed a mass consistent with a pseudotumour. It was 11 cm x 3 cm x 4 cm in size, cystic in nature, and found to communicate with the hip joint. At revision an undisplaced fracture of the femoral neck was also noted.
A patient with a 'late' (46 months) fracture associated with gross metallosis had been entirely satisfied at two years post-operatively. He played football on a weekly basis. Routine blood metal ion analysis prior to his fracture had shown whole blood Cr and Co levels to be elevated at 29 μg/l and 69 μg/l, respectively. As he was asymptomatic, a decision was taken to simply observe him, but unfortunately he began to develop acute pain prior to a fracture being identified. A CT scan confirmed a fracture in the presence of a well-placed femoral component, which had been implanted without evidence of notching. At revision, the periprosthetic tissues were stained black, but there was no gross joint effusion. This patient has been included as a failure secondary to ARMD, making a total of 17 from this cause.

**Joint orientations and metal ion concentrations.** The mean anteversion angles of the acetabular component and the joint sizes in the ARMD group were significantly different from those in the remainder of the ASR cohort (Table II). Blood metal ion concentrations in the ARMD group were greatly increased, with the median blood Co ion level 20 times greater than in the asymptomatic group. Levels of Cr and Co in the joint fluid were also significantly increased in the ARMD group compared to those obtained from patients undergoing revision for other reasons (Fig. 1).

**Histopathology.** Histopathological analysis of tissue samples from ARMD patients consistently showed widespread infiltration of histiocytes with areas of tissue necrosis. In many cases the histiocytes were seen to contain small unidentified black particles consistent with the description of metal particulate debris. Most viable tissue samples exhibited areas with high numbers of lymphocytes forming cuffs around vascular tissue to a greater or lesser extent, and synovial ulceration was also frequently seen. These characteristics are consistent with the description of ALVAL. The extent of histiocyte and lymphocytic infiltration was recorded for each patient, along with the circumference of the observed lymphocytic cuffs. Examination of capsular tissues obtained at revision surgery for an uncomplicated fracture did not show lymphocytic cuffing (Table III).

**Explant analysis.** The maximum out-of-roundness value for each available explant can be seen in Table IV. The ARMD explants showed high out-of-roundness values compared to those obtained from explants retrieved for revision for an uncomplicated fracture. Although they can be caused by deformation secondary to the manufacturing and/or implantation processes, the low maximum out-of-roundness values for the fracture explants strongly suggest that an unexpectedly high rate of wear in vivo resulted in the out-of-roundness of the ARMD explants. The out-of-roundness traces of the acetabular components indicated that wear occurred at the edge of the component. The traces of the femoral components indicated a variety of wear patterns. Some showed localised wear on the superior surface, but others implied a more complex pattern, which is the subject of continuing research.

| Table II. Femoral sizes, acetabular component orientation and metal ion concentrations of the patients with adverse reactions to metal debris (ARMD) compared to the remainder of the articular surface replacement (ASR) cohort. Mean values are given for size and angles, with median values given for metal ion concentrations with ranges in parentheses |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | ARMD            | Asymptomatic ASRs | p-value |
| Femoral size (mm) | 44.9 (41 to 51)  | 48.62 (41 to 59) | < 0.001       |
| Inclination angle of acetabular component (°) | 52.5 (42 to 70)  | 48.4 (25 to 68)  | 0.076         |
| Anteversion angle of acetabular component (°) | 27.4 (18 to 39)  | 19.7 (-5 to 40)  | < 0.001       |
| Whole blood Cr (μg/l) (range) | 29.3 (3.89 to 41.8) | 3.89 (1.51 to 69.8) | < 0.001       |
| Whole blood Co (μg/l) (range) | 68 (7.82 to 99.1)  | 2.67 (0.38 to 271) | < 0.001       |
| Serum Cr (μg/l) (range) | 33.6 (3.84 to 67.5) | 4.23 (0.58 to 115) | < 0.001       |
| Serum Co (μg/l) (range) | 29.7 (4.95 to 96.6) | 2.67 (0.38 to 228) | < 0.001       |

Figure showing the comparison of metal ion concentrations in the joint fluid in patients with adverse reactions to metal debris and two controls (two Birmingham hip resurfacing patients revised for severe neck narrowing and another aspirated for suspicion of infection). The inferior, middle and superior horizontal lines of the boxes represent the first quartile, median and third quartiles. The ends of the 'whiskers' correspond to the limits of the data, beyond which values are considered anomalous. The mean is displayed with a +, outliers with *, and upper and lower values with *.
Table III. Histology results and comparison with clinical data. The last three results are given as controls. Zero months indicate that the patient was never free of pain post-operatively

<table>
<thead>
<tr>
<th>Case</th>
<th>Gender</th>
<th>Allergies/ Atopic condition</th>
<th>Onset of pain post initial surgery (mths)</th>
<th>Onset of pain to revision (mths)</th>
<th>Layer 1 surface membrane</th>
<th>Histioocyte sheet thickness (mm)</th>
<th>Lymphocytes</th>
<th>Lymphocyte cuff thickness (mm)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>Codeine</td>
<td>0</td>
<td>19</td>
<td>Type 4 extensive</td>
<td>3</td>
<td>Absent</td>
<td>0.05 to 0.15</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>Nil</td>
<td>3</td>
<td>33</td>
<td>Type 4 extensive</td>
<td>0.5</td>
<td>Marked</td>
<td>0.2 to 0.25</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>Septrin</td>
<td>2</td>
<td>30</td>
<td>Type 3 and type 4</td>
<td>2</td>
<td>Moderate</td>
<td>0.05 to 0.15</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>Morphine</td>
<td>0</td>
<td>36</td>
<td>Type 3 and focal type 4</td>
<td>1</td>
<td>Absent</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>Elasto-plast</td>
<td>7</td>
<td>29</td>
<td>Type 4 massive</td>
<td>NA*</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>Penicillin</td>
<td>0</td>
<td>12</td>
<td>Type 3 and type 4</td>
<td>2</td>
<td>Moderate</td>
<td>0.1 to 0.15</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>Elasto-plast</td>
<td>6</td>
<td>12</td>
<td>Type 4 massive</td>
<td>1.5</td>
<td>Moderate</td>
<td>0.25 to 0.3</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>Asthma, hayfever</td>
<td>25</td>
<td>2</td>
<td>Type 3 and type 4</td>
<td>2.5</td>
<td>Mild</td>
<td>0.2 to 0.35</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>Tramadol, suncream</td>
<td>0</td>
<td>8</td>
<td>Type 3 and type 4</td>
<td>3</td>
<td>Mild/moderate</td>
<td>0.05 to 0.15</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>Asthma</td>
<td>0</td>
<td>17</td>
<td>Type 3 and type 4</td>
<td>0.9 to 2</td>
<td>Mild</td>
<td>0.15 to 0.2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>Sun cream</td>
<td>8</td>
<td>1</td>
<td>Type 4 massive</td>
<td>2.5</td>
<td>Moderate</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>Nil</td>
<td>0</td>
<td>2</td>
<td>Type 4 massive</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>M</td>
<td>Nil</td>
<td>35</td>
<td>3</td>
<td>Type 4 extensive</td>
<td>2</td>
<td>Mild</td>
<td>0.25 to 0.35</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>M</td>
<td>Nil</td>
<td>48</td>
<td>1</td>
<td>Type 4 massive</td>
<td>3</td>
<td>Mild</td>
<td>0.05 to 0.15</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>Nil</td>
<td>48</td>
<td>2</td>
<td>Type 4 massive</td>
<td>3</td>
<td>Mild</td>
<td>0.05 to 0.15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>Nil</td>
<td>No pain</td>
<td>0</td>
<td>Type 2</td>
<td>0.5</td>
<td>Absent</td>
<td>0.15 to 0.2</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>Nil</td>
<td>52</td>
<td>2</td>
<td>Type 2</td>
<td>0.5</td>
<td>Absent</td>
<td>0.15 to 0.2</td>
<td>Sepsis</td>
</tr>
</tbody>
</table>

*NA, not applicable
†Patient sustained a fracture at 46 months
‡ALVAL, aseptic lymphocyte-dominated vasculitis-associated lesions

Metal-lymphocyte transformation tests. None of the ARMD samples showed increased lymphocyte reactivity to Cr or Co in vitro, or demonstrated an elevated incidence of metal sensitivity above that of the general population (10% to 15%). Only one ARMD patient showed metal reactivity, which was a mild response to aluminium, molybdenum and nickel (Table V).

Discussion

In this series we have described a failure rate as a consequence of ARMD of 3.2% in our ARMD resurfacing group, and 6.0% in the ASR THR group. This amounts to a failure rate of 3.4% for all patients with ARMD prostheses in this study. These figures are likely to represent the 'best case scenario'. Based on our experience, which is consistent with the reports in the literature,10 metal debris-related complications may take several years to develop. The absence of ARMD in the BHR group, despite patients being drawn from the same geographical pool, suggests there is no genetic predisposition to this condition.

We have previously shown that smaller ARMD components are particularly sensitive to position in terms of metal ion release.13 Our results have been corroborated by another centre.19 We have now shown that ARMD patients experiencing joint failure as a consequence of ARMD have significantly smaller, suboptimally orientated acetabular components than do asymptomatic patients. Concentrations of blood and joint metal ions in failed cases were significantly increased compared to those in patients who were pain free. Explant analysis of the available failed ARMD joints confirmed greater wear than expected for the duration in situ, as out-of-roundness values are a direct indication of the wear depth on the different components. Maximum out-of-roundness values of ARMD femoral components were markedly elevated compared to those of explants retrieved following fracture, including a femoral component that was retrieved following a fracture two years post-operatively. We conclude that an increase in the production of metal debris increases the risk of adverse local tissue reactions. There are reports of adverse immune responses leading to failure of MoM joints where subsequent retrieval analysis showed only minimal wear on the bearing surfaces.20 The absence of the condition in our study group with BHR implants or larger/well-positioned ASRs suggests that these cases represent a minority group of sensitive patients who go on to develop ARMD despite relatively low levels of metal debris.
Table IV. Comparison of maximum out-of-roundness values for the analysed adverse reaction to metal debris failures with those retrieved following fracture

<table>
<thead>
<tr>
<th>Implant*</th>
<th>Time in situ (mths)</th>
<th>Reasons for revision</th>
<th>Max out of roundness (head) (μm)</th>
<th>Max out of roundness (acetabular component) (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASR</td>
<td>28</td>
<td>Pain, effusion</td>
<td>31.3</td>
<td>13.8</td>
</tr>
<tr>
<td>ASR</td>
<td>18</td>
<td>Pain, effusion</td>
<td>17.7</td>
<td>9.4</td>
</tr>
<tr>
<td>ASR THR</td>
<td>36</td>
<td>Pain, effusion</td>
<td>38.8</td>
<td>22.9</td>
</tr>
<tr>
<td>ASR</td>
<td>8</td>
<td>Pain, effusion</td>
<td>91.8</td>
<td>64</td>
</tr>
<tr>
<td>ASR</td>
<td>27</td>
<td>Pain, effusion</td>
<td>38</td>
<td>28.8</td>
</tr>
<tr>
<td>ASR</td>
<td>35</td>
<td>Pain, pseudotumour</td>
<td>12.9</td>
<td>6.66</td>
</tr>
<tr>
<td>ASR</td>
<td>14</td>
<td>Pain, effusion</td>
<td>3.69</td>
<td>6.86</td>
</tr>
<tr>
<td>ASR</td>
<td>46</td>
<td>Late fracture</td>
<td>66.9</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>17</td>
<td>Pain, effusion</td>
<td>32.9</td>
<td>23.1</td>
</tr>
<tr>
<td>ASR</td>
<td>2</td>
<td>Fracture</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>8</td>
<td>Fracture</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>0</td>
<td>Fracture</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>2</td>
<td>Fracture</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>24</td>
<td>Fracture</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>6</td>
<td>Fracture</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

* ASR, articular surface replacement; THR, total hip replacement

Table V. Metal-lymphocyte transformation (LTT) results. Time indicates number of months after revision surgery when blood samples were taken

<table>
<thead>
<tr>
<th>ARMD* failures</th>
<th>Cr</th>
<th>Co</th>
<th>Time</th>
<th>LTT results</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.9</td>
<td>96</td>
<td>19</td>
<td>No reactivity</td>
<td></td>
</tr>
<tr>
<td>64.8</td>
<td>217</td>
<td>12</td>
<td>No reactivity</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>20.9</td>
<td>7</td>
<td>Mild reactivity to Al, Mo, Ni</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>8.0</td>
<td>1 day</td>
<td>No reactivity</td>
<td></td>
</tr>
<tr>
<td>Asymptomatic high ions</td>
<td>26</td>
<td>46.4</td>
<td>42</td>
<td>Mild reactivity to Ni</td>
</tr>
<tr>
<td>69.8</td>
<td>271</td>
<td>41</td>
<td>No reactivity</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>81</td>
<td>30</td>
<td>Mild reactivity to Ni, Fe</td>
<td></td>
</tr>
<tr>
<td>Asymptomatic low ions</td>
<td>2.89</td>
<td>1.04</td>
<td>38</td>
<td>Mild reactivity to Co, Cr, Ni Moderate to Mo</td>
</tr>
<tr>
<td>3.4</td>
<td>1.5</td>
<td>6</td>
<td>No reactivity</td>
<td></td>
</tr>
</tbody>
</table>

* ARMD, adverse reaction to metal debris; † Al, aluminium; Mo, molybdenum; N, nickel; Fe, iron

Levels of blood metal ions correlate well with articular wear in retrieved components, and can therefore be used as a surrogate marker in vivo wear.21 We can use this information to propose an explanation for the increased incidence of ARMD in the ASR group. BHR acetabular components provide a greater arc of cover than do ASR components. For example, the coverage provided by the average female BHR (size 52 mm) acetabular component in this series is 162°, with the equivalent ASR component (size 52 mm) providing 151° of cover. Therefore, in components matched for size and orientation, articular contact will take place closer to the rim of an ASR acetabular component. This 'edge loading' effect has been shown in numerous retrieval studies to be strongly associated with increased articular wear,5,22 and this is consistent with the results of our explant analyses. We have previously published work proposing that the location of the articular contact patch in the standing position of the patient is crucial in the development of high wear states.6 The closer this patch lies to the rim of the acetabular component the greater the chance of increased blood metal ion levels. Figure 2 shows the calculated mean contact patch to rim distances for ASR and BHR patients (men and women) in our series. The smaller the mean contact patch to rim distance the greater the incidence of ARMD. Women are exposed to higher concentrations of metal ions owing to a combination of factors: increased mean acetabular component inclination and anteversion, and a reduced mean joint size (Table II). Anatomical studies also show that the average female patient has a greater degree of femoral anteversion, which could also lead to posterior impingement and microseparation, a process implicated in increased wear.5,21 It has been reported that high
metal ion concentrations are associated with disturbing amounts of tissue necrosis, even when the patient remains asymptomatic.21

Some individuals and populations (e.g. women) have a greater predisposition to have or develop metal-associated delayed-type metal hypersensitivity, possibly as a consequence of prior metal sensitisation caused by jewellery24 or, environmental exposure. Metallic ions purportedly become bound to plasma proteins, and the resulting complexes may stimulate an immune response.25-27 Furthermore, this reaction is metal specific, and the observed reactivity correlates with serum levels of metal ions in vivo.28 However, the results we obtained from the lymphocyte transformation test studies were not consistent with this idea. The lack of lymphocyte reactivity to Cr and Co ions implies that if an immune response is the pathogenic factor in the development of an effusion, necrosis and pain, then it is likely to be a localised response rather than a pre-existing or systemic sensitivity. The fact that two of the patients with ARMD in this series had bilateral implants but only unilateral symptoms lends support to this theory. However, localised immune responses may act to sequester activated lymphocytes to the site of the joint inflammation and lead to fewer available peripherally primed circulating lymphocytes for lymphocyte transformation testing, and hence a non-elevated systemic response in vitro.29

The unsatisfactory outcome in the two patients described in this series who were revised to metal-on-metal bearings, appear to favour the existence of a localised persisting immune response.7 Newly published results30 have implicated a mediated response to Co-Cr-Mo alloy particles as a potentially important factor in reactions to metal debris. This type of reactivity, which has been shown to be critical to sensitivity responses to other non-metal antigens, would fit with the appearance of the ARMD tissues, which show heavy histiocyte infiltration. However, there is also accumulating evidence to show that massive concentrations of Cr and Co ions in the joint fluid could alone explain the observed micro- and macroscopic tissue necrosis.31,32

The relationship between blood and joint concentrations of metal ions is unexpected and warrants further investigation. In patients with the highest concentrations of joint metal ions, Cr ions predominate, with a concentration around four times that of Co. In the blood of the same patients, the opposite is true, with Co concentrations two to three times those of Cr (Fig. 3). Only when blood metal ions are present in high concentrations do they reflect the ratios of the metal alloy from which they are released (Co:Cr 2.4:1). We acknowledge the small size of our data set, but the results do appear to suggest that Cr is sequestered in and around the joint space, with Co apparently more easily released into the circulation. Transmission electron microscope studies of metal particulate matter from tissues retrieved at revision surgery for failed joints at other centres, have shown that these particles are composed almost entirely of Cr, mostly in the form of a product of corrosion, chromium orthophosphate.33,34

Early failure as a consequence of ARMD is probably multifactorial. Modifiable factors such as the orientation of the acetabular component and implant design conspire
with non-modifiable factors, including component size and host tolerance/response, to determine the success or failure of the resurfacing procedure. We have an unacceptably high failure rate secondary to ARMD in patients with small ASR bearings (Fig. 4). We have previously identified what we believe to be the optimal orientation for the ASR acetabular component and highlighted the relatively small margins of error involved when attempting to reduce blood metal ion concentrations.\(^\text{13}\) This investigation has described the clinical consequences of the resultant increase in articulär wear when these boundaries are exceeded (Fig. 5). Numerous studies have shown that surgeons cannot consistently achieve the desired acetabular component orientation,\(^\text{34,35}\) and our own experience is consistent with the literature. Surgeons must consider implant design, expected component size and positioning of the acetabular implant in order to reduce early failures when performing large-bearing metal-on-metal hip resurfacing and THR.

**Supplementary material**

Appendices giving further information on proliferation assay and hip joint contact are available with the electronic version of this article on our website at www.jbjs.org.uk

The author or one or more of the authors has received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article. In addition, benefits have been made or will be directed to a research fund, foundation, educational institute, or other non-profit organisation with which one or more of the authors are associated.

**References**


Tribological analysis of failed resurfacing hip prostheses and comparison with clinical data

T J Joyce1, D J Langton2, S S Jameson2, and A V F Nargo2
1Centre for Rehabilitation and Engineering Studies, School of Mechanical and Systems Engineering, Newcastle University, Newcastle upon Tyne, UK
2Joint Replacement Unit, University Hospital of North Tees, Hardwick, Stockton-on-Tees, UK

The manuscript was received on 21 July 2008 and was accepted after revision for publication on 8 January 2009.

DOI: 10.1243/13506501JET484

Abstract: Metal-on-metal resurfacing hip prostheses offer potential benefits over total hip replacement for younger and more active patients. Although some reported clinical results of resurfacing hip prostheses are excellent, other outcomes are less positive. To aid with understanding the balance of benefits related to these devices, analysis of failed resurfacing prostheses can contribute critical insights. However, because these implants are so new there are relatively few such prostheses available for independent ex vivo analysis. From a single-surgeon clinical cohort, a number of failed resurfacing hip prostheses were obtained and studied. It was found that roughness values of the articulating surfaces had increased so that the theoretical lubrication regime would shift from the fluid film to the boundary. In turn, this would likely result in increased wear from the articulating surfaces. High ion levels were seen in the patients from whom the explants were obtained, thus supporting the hypothesis that wear was linked with failure of the explanted hip resurfacing prostheses.

Keywords: metal-on-metal, cobalt chrome, hip resurfacing, hip prostheses, explant

1 BACKGROUND

Total hip replacement has been hailed as the outstanding achievement in orthopaedics of the twentieth century [1] and tribological expertise has contributed greatly to this success. Hip replacement is now a commonplace surgical procedure in many countries and the most recent data indicate that in England and Wales in the year 2006–2007 more than 55,000 primary hip replacement procedures were undertaken [2]. The average age of these patients was 68 years [2]. The majority of these people will enjoy the benefits of pain relief, increased independence, and enhanced mobility. Moreover, the success of the procedure has led to increased demand for surgical intervention and, coupled with an ageing population, it has been recognized that the prevalence of hip replacement will grow rapidly [3].

Unfortunately, not all of the hip prostheses that are implanted will show long-term success. The majority of total hip prostheses employ a hard metal or ceramic femoral component that articulates against a polyethylene acetabular socket. It is now recognized that excessive wear from the polyethylene component can result in a negative cascade of events within the body, which can eventually result in osteolysis, loosening of the implant, pain, and the need for revision surgery [4]. It is also acknowledged that younger patients tend to place greater demands on their hip prostheses, which in turn leads to greater wear and consequently reduced longevity and survivorship [5]. Much effort has therefore been directed at reducing wear volumes from hip prostheses, so patients of all ages can benefit from hip replacement. For example, the wear properties of polyethylene can be improved by irradiation cross-linking. From one clinical trial, a 40 per cent lower wear rate of cross-linked polyethylene acetabular cups has recently been reported [6].
One method of eliminating the problem of polyethylene wear debris is to employ a wear couple of cobalt chrome molybdenum (CoCr) articulating against itself. This material combination has led to a renaissance of metal-on-metal hip prostheses [1]. An exciting subsequent development has been to offer metal-on-metal resurfacing prostheses [7]. These have a relatively large diameter and are therefore less likely to dislocate [8]. They also offer the potential benefits inherent in removing less femoral bone stock compared with conventional total hip replacement. For such reasons, these resurfacing hip prostheses tend to be implanted in younger patients. Recent data from the UK have shown that for patients under the age of 55 who receive a hip replacement, 46 per cent will receive a resurfacing implant [9].

What then are the tribological benefits of metal-on-metal resurfacing hip prostheses? By minimizing surface roughness as far as possible, controlling sphericity and reducing the clearance between the head and the cup, and given the material properties of CoCr plus the relatively large size of the articulation, this theoretically allows fluid film lubrication to be achieved during part of the gait cycle [10]. Various designs of hip resurfacing prostheses are offered by several orthopaedic joint manufacturers [7]. The market leader is the earliest contemporary design, the Birmingham Hip Replacement (BHR). This device has shown some very good mid-term clinical results [8, 9]. More recent designs include the Durom™ from Zimmer and the Articular Surface Replacement (ASRTM) from De Puy. Both of these designs differ from the BHR in that they employ a sub-hemispheric acetabular cup to reduce acetabular bone removal prior to implantation and smaller diametral clearances to promote the fluid film lubrication.

As will be appreciated, the truest test of any total joint replacement occurs when it is implanted into the body. Although it may be of small consolation to the individual involved, an examination of a failed replacement joint can provide crucial evidence as to how these devices can be improved for the benefit of future patients suffering from crippling musculo-skeletal diseases. Such explant analysis has usefully been undertaken for knee [11], finger [12], and toe [13] prostheses as well as for total hip replacements [14]. Given their relatively recent introduction, there are comparatively few papers that have investigated explanted resurfacing hip prostheses [15].

Another issue with metal-on-metal hip joints, total replacement as well as resurfacing, is the exposure of the body to metal ions. There are concerns that high levels of systemic chromium (Cr) and cobalt (Co) ions may cause organ toxicity, carcinogenicity [16, 17], and teratogenicity [18]. There is currently no conclusive evidence of these adverse effects but, as yet, there is an absence of large epidemiological studies in the literature [17]. Clearly it is desirable to reduce the metal ion loads to which patients are exposed.

2 METHODS

The largest independent, single surgeon cohort of resurfacing hip prostheses in the UK has formed part of an ongoing clinical investigation that includes approximately 390 ASRTM prostheses [19]. Patients in this clinical study were monitored in a number of ways. Standing antero-posterior pelvic radiographs were taken and Einzel-Bild-Roentgen-Analyse (EBRA) software was used to measure the acetabular cup inclination and anteverision angles [20, 21]. Many of the patients underwent whole blood metal ion analysis (Co and Cr concentrations) at a minimum 1 year following their resurfacing hip replacement [19]. Whole blood samples were measured by inductively coupled plasma mass spectrometry, which is regarded as the most sensitive method for measuring systemic exposure to Cr and Co ions [18]. Ten patients (2.4 per cent) within this 390 ASRTM cohort presented with pain and a number were operated on and fitted with a ceramic-on-ceramic hip prosthesis. From two patients, ASRTM hip resurfacing femoral head and acetabular cup were available for explant analysis. From another three patients, only femoral components were available. All five patients were female. The duration of implantation varied between 8 and 35 months. It should be noted that none of the failed prostheses described in this article were associated with failure due to fracture of the femur. This is important as femoral neck fractures are said to be the main reason for early hip resurfacing failure [22].

The roughness values of the articulating surfaces of the explanted resurfacing hip prostheses were measured using a ZYGO NewView 5000 non-contacting profilometer. A similar device has been used previously to analyse resurfacing hip prostheses tested in vitro [23]. For each component ten roughness measurements were taken and then the average of these was calculated. The readings were taken at the pole and then at nine points obtained by intersecting 'longitude' lines spaced 40° apart, combined with a position 30° offset from the central axis of the component. A Mitutoyo Crysta 544 co-ordinate measuring machine was used to measure the spherical diameter of the head and the cup, following the guidance of the appropriate British Standard [24], and thus the diametral clearance could be calculated. Mitutoyo report an accuracy (E) of $E = (3.5 + 0.45L/100) \mu m$, where $L$ is the length of the component being measured, as well as a probe repeatability of $4 \mu m$, for the Crysta 544 co-ordinate measuring machine. The same roughness and dimensional measurements were also taken from a 'Do Not Implant' (DNI) ASRTM prosthesis (Fig. 1) and compared.
Using this dimensional data and modelling the resurfacing hip prosthesis as an equivalent ball-on-plane model and employing elastohydrodynamic theory [25] allowed the minimum effective film thickness \((h_{\text{min}})\) to be calculated from

\[
\frac{h_{\text{min}}}{R_x} = 2.80 \left( \frac{\eta u}{E^*R_x} \right)^{0.65} \left( \frac{w}{E^*R_x^2} \right)^{-0.21}
\]

Here, \(R_x\) is the equivalent radius (m), \(\eta\) is the viscosity of the lubricant (Pa s), \(u\) is the entraining velocity (m/s), \(E^*\) is the equivalent elastic modulus (Pa), and \(w\) is the load (N). In turn, given that \(R_x\) is the surface roughness and assigning subscript 1 to the head and subscript 2 to the acetabular socket of the resurfacing hip prosthesis under consideration, the lambda values were calculated from

\[
\lambda = \frac{h_{\text{min}}}{\left( (R_{a1})^2 + (R_{a2})^2 \right)^{1/2}}
\]

This allowed the lubrication regime to be identified, as \(\lambda < 1\) indicates boundary lubrication, \(\lambda > 3\) designates fluid film lubrication, and between these values mixed lubrication is indicated [26].

Before these calculations could be undertaken, the equivalent radius \((R_x)\) was calculated from

\[
\frac{1}{R_x} = \frac{1}{R_1} - \frac{1}{R_2}
\]

where \(R\) refers to the radius of the component and subscript 1 refers to the femoral head and subscript 2 to the socket of the resurfacing hip prosthesis under consideration. The equivalent modulus of elasticity was determined from the equation

\[
\frac{1}{E^*} = 0.5 \left( \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \right)
\]

Again, \(E\) refers to Young's modulus of the component and subscript 1 refers to the femoral head and subscript 2 to the acetabular socket of the resurfacing hip prosthesis, similarly for the two Poisson's ratios.

Hip joints can move at a range of speeds. Taking an average angular velocity \((\omega)\) of 1.5 rad/s [27] then allowed entraining velocity to be calculated for the particular head diameter \((d)\) under consideration using the equation

\[
u = \frac{\omega d}{4}
\]

A range of loads can be taken by the natural hip joint. From the literature, an average load of 2500 N was taken [27]. In addition, a viscosity of the synovial fluid lubricant of 0.0025 Pa s was assumed, together with a Young's modulus of 210 GPa and a Poisson's ratio of 0.3 for CoCr [27].

3 RESULTS

The results of the ex vivo analysis alongside the related clinical data are summarized in Table 1. Clinically, all five acetabular cups were implanted with inclination angles over 50° and anteverision over 29°. Each of these explanted resurfacing hip prostheses was associated with high ion levels in the patient from whom it was removed. All prostheses were towards the small size of the available range, with an articulating diameter of less than or equal to 50.5 mm. Data from the wider clinical ASR™ cohort showed that median whole blood metal ion levels were 1.99 µg/l for Co and 3.42 µg/l for Cr, whereas the mean cup inclination and anteverision angles were 48.5° and 19.8°, respectively [19]. For comparative purposes the values of Co and Cr in the blood

<table>
<thead>
<tr>
<th>Sample/patient</th>
<th>DNI</th>
<th>1♀</th>
<th>2♀</th>
<th>3♀</th>
<th>4♀</th>
<th>5♀</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roughness head (µ m Ra)</td>
<td>0.010</td>
<td>0.135</td>
<td>0.045</td>
<td>0.047</td>
<td>0.025</td>
<td>0.062</td>
</tr>
<tr>
<td>Roughness cup (µ m Ra)</td>
<td>0.012</td>
<td>0.058</td>
<td>0.044</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Diametral clearance (µ m)</td>
<td>87</td>
<td>110</td>
<td>96</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Articulating diameter (mm)</td>
<td>48.5</td>
<td>50.5</td>
<td>42.5</td>
<td>45.5</td>
<td>46.5</td>
<td>42.5</td>
</tr>
<tr>
<td>Lambda value</td>
<td>3.9</td>
<td>0.38</td>
<td>0.65</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Implantation time (months)</td>
<td>–</td>
<td>27</td>
<td>8</td>
<td>35</td>
<td>22</td>
<td>24</td>
</tr>
<tr>
<td>Ion levels Cr (µ g/l)</td>
<td>–</td>
<td>5.3</td>
<td>35.9</td>
<td>22.0</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ion levels Co (µ g/l)</td>
<td>–</td>
<td>7.8</td>
<td>87.5</td>
<td>32.2</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cup inclination (°)</td>
<td>–</td>
<td>50</td>
<td>65</td>
<td>60</td>
<td>–</td>
<td>57</td>
</tr>
<tr>
<td>Cup anteverision (°)</td>
<td>–</td>
<td>32</td>
<td>39</td>
<td>31</td>
<td>29</td>
<td>–</td>
</tr>
</tbody>
</table>
of pre-operative patients have been reported as 1.75 and 0.05 μg/l, respectively [28].

The data for the ‘DNI’ components can be contrasted with that from the five patients, identified by numbers 1–5 in Table 1. As can be seen, the measurements of the ‘DNI’ ASR<sup>TM</sup> components, with the head and cup surface roughness values of 0.010 and 0.012 μm, respectively, and a diametral clearance of 87 μm, give a lambda value of 3.9 and imply that such an implant would operate under fluid film lubrication.

In contrast, the two prostheses for which the explanted head and cup were available can be seen to each have a lambda value of less than 1 and therefore at the time of revision surgery would have operated under boundary lubrication. Where a head and a cup were not available (patients 3–5), a diametral clearance as well as a compound surface roughness could not be calculated and therefore the lubrication regime could not be determined.

It is clear from the roughness values of the articulating surfaces of the explanted resurfacing hip prostheses that they have roughened when in the body. This is clearest in the case of patient 1, with an average roughness value of 0.135 μm Ra for the head and 0.058 μm for the cup. Together with a diametral clearance of 110 μm these results indicated that, at the time of removal, the prosthesis would have operated in the boundary lubrication regime. Other explanted femoral components showed articulating surfaces with average roughness values of between 0.025 and 0.062 μm. The explanted acetabular cup from patient 1 also showed evidence of the rim damage.

Images from the ZYGO non-contacting profilometer are given in Figs 2 to 3. Figure 2 shows a typical pair of images of the surface of the ‘DNI’ cup. It can be seen that a roughness value of 0.006 μm Ra has been measured. This value indicates the smoothness of the surface as does the uniformity and lack of scratches seen on the oblique plot of Fig. 2.

In contrast, Fig. 3 shows a similar pair of images taken from the explanted femoral head of the prosthesis of patient 1. Here, it can be seen that a roughness value of 0.210 μm Ra has been measured. Together with the topography this data indicate that significant roughening has occurred to this articulating surface in vivo.

### 4 DISCUSSION

The values measured from the ‘DNI’ prosthesis, specifically a diametral clearance of 87 μm and head and cup surface roughness values of 0.010 and 0.012 μm Ra, respectively, were credible. A representative diametral clearance of ASR<sup>TM</sup> metal-on-metal resurfacing prostheses is said to be approximately 100 μm [29, 30].
Roughness values for the articulating surfaces of metal-on-metal hip prostheses have been offered as typically 0.005 and 0.010 μm [31] and between 0.005 and 0.025 μm Ra [27].

Perhaps the key question is what caused these resurfacing hip prostheses to fail? From the larger cohort of which these failed prostheses were a subset, it has been shown that higher blood ion levels were associated with three factors: acetabular cup inclination greater than 45°, acetabular cup anteversion greater than 20°, and components with articulating surfaces of diameter less than or equal to 51.5 mm [19].

Considering the latter point first, the tribological theory informs us that, all other things being equal, smaller implants will tend to operate further from the fluid film lubrication than larger implants. However, if the roughness and clearance for the 'DNI' prosthesis were to be applied under the conditions detailed above to the smallest size of failed ASR™, at 42.5 mm articulating diameter, then this results in a lambda value of 2.6. This is fairly close to the lambda value of 3 that implies fluid film lubrication. Values of lambda that approach 3 have been described as the 'mild'- mixed lubrication regime [1]. Therefore, size alone cannot explain the failures, although smaller implants may tend to work in the mixed lubrication regime rather than under fluid film lubrication.

Can the high angles of acetabular cup inclination and anteversion have been a contributing factor to the failure of these resurfacing hip prostheses? This is certainly possible as, at high angles of inclination, not only have higher metal ion levels been reported in patients [32], but so too has edge loading [15]. Such edge loading was also seen in the form of rim damage on one of the two explanted cups studied. This edge loading may have initiated the wear and in turn increase in surface roughness that was measured. In addition to noting the correlation between the acetabular cup inclination and ion levels in patients fitted with resurfacing hip prostheses, the importance of edge loading has also been reported recently [33]. These authors postulated that at such high angles the 'arc of cover' between the femoral head and the acetabular cup was reduced and they found that smaller arcs of the cover were associated with higher wear [33]. Such an explanation could also apply to the ASR™ prostheses discussed in the current article. Another very recent study has reported a 21–27-fold greater wear rate in rim-loaded hip resurfacing explants compared with non-rim-loaded samples [22]. It should also be recognized that in 'conventional' metal-on-polymer total hip prostheses failure rates have increased with inclination of the acetabular cup [34], while the importance of acetabular position to low wear has also been identified in ceramic-on-ceramic total hip prostheses [35].

A recent simulator study, mainly on 28 mm diameter metal-on-metal total hip replacements, but also involving some 39 mm diameter hip resurfacings showed a five-fold increase in wear when the inclination angle was changed from 45° to 55° [36]. A separate simulator study also showed greater wear as inclination angles increased [37].

Although a small number of studies of explanted resurfacing hip prostheses have been undertaken [15, 38], none of these have offered values of roughness for the articulating surfaces. Therefore, this article is the first to offer such data. In the boundary lubrication regime, with its preponderance of surface-to-surface interaction, increased wear would be expected. In turn, this would result in higher volumes of wear debris. In addition, the relatively large size of the resurfacing hip prostheses means that, under boundary lubrication conditions, wear is taking place over a large sliding distance so that volumes of wear debris are maximized. Recently, large amounts of particulate wear debris have been linked with a toxic reaction in localized cells and the attendant failure of resurfacing hip prostheses [39]. Clearly the high wear could also be directly linked with the higher ion levels, which were measured in the patients from whom these explanted resurfacing hip prostheses were obtained [19].

5 CONCLUSIONS

The results presented in this article suggest that explanted resurfacing hip prostheses, which have been implanted with high cup inclination and anteversion angles, have higher values of surface roughness compared with a new prosthesis of the same type. The explanted prostheses may therefore have operated under boundary rather than fluid film lubrication. In turn, this may have led to greater than expected wear, concomitant higher ion levels in the patients, and may also be linked with early failure of these prostheses.

REFERENCES


Joyce, T. J. Examination of failed ex vivo metal-on-metal metatarsophalangeal prosthesis and comparison with theoretically determined lubrication regimes. *Wear*, 2007, 263(7-12), 1050-1054.


A study of the wear of explanted metal-on-metal resurfacing hip prostheses

Thomas J. Joyce, David J. Langton, Antoni V.F. Nargol

Abstract

Due to their recent introduction there are few studies of retrieved resurfacing hip prostheses. Nine such components associated with groin pain in patients, and five associated with early fracture of the femur, were obtained and analysed using a roundness measuring machine. While the ‘fracture’ components showed no more than 3 µm out of roundness, components associated with groin pain showed between 15 and 92 µm out of roundness values. These latter results indicate wear and correlated with high metal ion levels in these patients, therefore the groin pain was likely associated with an adverse reaction to excessive metal wear debris.

1. Introduction

A recent article in the Lancet was entitled ‘the operation of the century: total hip replacement’ [1]. Tribological expertise has contributed greatly to the success of this surgical procedure which helps many patients suffering from common musculoskeletal diseases. Hip replacement is now commonplace in many countries and the most recent data indicates that in England and Wales in the year 2008 more than 64,000 primary hip replacement procedures were undertaken [2]. The average age of these patients was 67 years [2]. The majority of these people will enjoy the benefits of pain relief, increased independence and enhanced mobility. Moreover, the success of the procedure has led to increased demand for surgical intervention and, coupled with an ageing population, it has been recognised that the prevalence of hip replacement will grow rapidly [3].

Unfortunately, not all of the hip prostheses which are implanted will show long-term success. The majority of total hip prostheses employ a hard metal or ceramic femoral component, which articulates inside a polyethylene acetabular socket. It is now recognised that excessive wear from the polyethylene component can result in a negative cascade of events within the body which can eventually result in osteolysis, loosening of the implant, pain, and the need for revision surgery [4]. It is also acknowledged that younger patients tend to place greater demands on their hip prostheses, which in turn leads to greater wear and consequently reduced longevity and survivorship [5]. Much effort has therefore been directed at reducing wear volumes from hip prostheses so patients of all ages can benefit from hip replacement. For example, the wear properties of polyethylene can be improved by irradiation cross-linking. From one clinical trial a 40% lower wear rate of cross-linked polyethylene acetabular cups has recently been reported [6].

One method of eliminating problems of polyethylene wear debris is to employ a wear couple of cobalt chrome molybdenum (CoCr) articulating against itself. This material combination has led to a renaissance of metal-on-metal hip prostheses [7]. An exciting subsequent development has been to offer metal-on-metal resurfacing prostheses [8]. These have a relatively large articulating diameter and are therefore less likely to dislocate [9]. They also offer the potential benefits inherent in removing less femoral bone stock compared with conventional total hip replacement. For such reasons these resurfacing hip prostheses tend to be implanted in younger patients. Recent data from the UK has shown that for patients under the age of 55 who receive a hip replacement, 46% will receive a resurfacing implant [10].

What then are the tribological benefits of metal-on-metal resurfacing hip prostheses? By minimising surface roughness as far as possible, controlling sphericity and reducing the clearance between the head and cup, and given the material properties of CoCr plus the relatively large size of the articulation, this theoretically allows fluid film lubrication to be achieved during part of the gait cycle [11]. In turn this should mean that wear of the articulating surfaces is minimised.

Various designs of metal-on-metal hip resurfacing prostheses are offered by several orthopaedic joint manufacturers [8]. The market leader is the earliest contemporary design, the Birmingham Hip Replacement (BHR). This device has shown some very good mid-term clinical results [9,10]. More recent designs include the Durom™ from Zimmer and the Articular Surface Replacement (ASR™) from De Puy. Both of these designs differ from the BHR in that they employ a sub-hemispheric acetabular cup to reduce...
acetabular bone removal prior to implantation and smaller diametral clearances to promote fluid film lubrication.

Like many other hip resurfacing systems, the ASR™ is designed so that the femoral component is cemented in place whereas the acetabular component is a cementless press-fit.

Heads are as cast while cups are also subject to hot isostatic pressing [8]. Both components are manufactured from high carbon (≥ 0.15%) CoCr. An image of an ASR™ prosthesis is shown in Fig. 1. To the left is the sub-hemispheric acetabular cup component. The porous coating designed to encourage bone ingrowth can be seen on the outside of the cup. Note that the inner, articulating surface does not extend all the way to the rim of the cup. Instead there is a step which matches with an 'introducer'—a device used during surgery to aid fitting of the cup within the pelvis. To the right is the femoral component. In comparison with a traditional total hip replacement, note the short stem which can be seen protruding to the right.

As may be appreciated, the truest test of any replacement joint occurs when it is implanted into the body. Although it may be of small consolation to the individual involved, an examination of a failed replacement joint can provide crucial evidence as to how these devices can be improved for the benefit of future patients suffering from crippling musculo-skeletal diseases. Such explant analysis has usefully been undertaken for knee [12], finger [13] and toe [14] prostheses as well as for total hip replacements [15].

Given their relatively recent introduction there are comparatively few papers which have investigated explanted resurfacing hip prostheses [16–19].

An important issue with metal-on-metal hip joints, total replacement as well as resurfacing, is the exposure of the body to metal ions. There are concerns that high levels of systemic chromium (Cr) and cobalt (Co) ions may cause organ toxicity, carcinogenicity [20,21] and teratogenicity [22]. There is currently no conclusive evidence of these adverse effects but, as yet, there is an absence of large epidemiological studies in the literature [21].

Clearly it is desirable to reduce the metal ion loads to which patients are exposed.

One accepted concern related to metal-on-metal hip resurfacing prostheses is early fracture of the femur [23,24]. This is thought to be due to a number of reasons including the surgeon's learning curve, gender differences in patients, as well as their body mass [25]. Such fractures generally occur soon after implantation, usually within one year. It has therefore been stated that femoral neck fractures are the main reason for early hip resurfacing failure [17].

Very recently there has emerged a body of evidence which suggests that, despite the theoretical advantages of metal-on-metal hip resurfacing prostheses in terms of promoting fluid film lubrication and minimising surface-to-surface contact, wear may be a serious problem in some of these devices [18,26–28]. By examining a number of explanted metal-on-metal hip resurfacing devices, it was intended to offer some evidence to inform this nascent concern.

2. Material and methods

The largest independent, single surgeon cohort of resurfacing hip prostheses in the UK [26] has formed part of an ongoing clinical investigation which includes approximately 500 ASR™ prostheses. Patients in this clinical study were monitored in a number of ways. Standing antero-posterior pelvic radiographs were taken and Einzel-Bild-Roentgen-Analyse (EBRA) software was used to measure acetabular cup inclination and anteverision angles [29,30]. Many of the patients underwent whole blood metal ion analysis (Co and Cr concentrations) at a minimum one year following their resurfacing hip replacement [26]. After one year, it has been argued that the 'bedding-in' period will have ceased and steady state wear has been achieved [31,32]. Whole blood samples were measured by Inductively Coupled Plasma Mass Spectrometry (ICPMS), which is regarded as the most sensitive method for measuring systemic exposure to Cr and Co ions [22].

From ten patients within this ASR™ cohort, fourteen explanted components were obtained for analysis. Five of these patients had reported pain and a characteristic effusion was seen at revision operation. From four of these patients, ASR™ femoral head and acetabular cup were available for explant analysis. From the fifth patient, only the femoral component was available. All five patients were female. Duration of implantation varied between 8 and 28 months. The remaining five ASR™ explants were obtained from patients whose femurs had fractured. In these cases only the femoral component was available for analysis. Here, devices were obtained from males and females and within eight months of implantation.

Each explanted component was examined using a Zeiss TSK Rondcom60A roundness measuring machine. Essentially, out of roundness refers to the deviation in shape from a perfect circle.

Modern manufacturing of resurfacing hip prostheses allows acetabular and femoral components to be produced with an out of roundness of less than 5 μm. Exploated components with greater out of roundness values imply that either material has been removed locally (wear) or that the component has been deformed. Devices such as the Zeiss TSK Rondcom60A roundness measuring machine have a resolution of approximately 0.1 μm. Out of roundness measurements were taken on three planes for each acetabular and femoral component. For the acetabular cups these planes were at 3, 7 and 11 mm below the rim. For the femoral heads these were at 3, 7 and 11 mm from the ‘pole’ of the head. Three traces were used at set distances between 3 and 11 mm as this gave a consistent procedure. As each prosthesis was subject to a unique range of loading and motion in the individual patient, so the exact area of wear or deformation would be unique too. The aim of measurements between 3 and 11 mm was to take in a substantial region where any changes might have occurred.

3. Results

The maximum out of roundness values for the fourteen ASR™ components taken from the ten patients are given in Table 1. As can be seen, those components associated with early fracture of the femur all showed the lowest out of roundness values, with a maximum of 3.1 μm. Given such low out of roundness values, all
Table 1

<table>
<thead>
<tr>
<th>Patient</th>
<th>Head OOR (µm)</th>
<th>Cup OOR (µm)</th>
<th>Reason for revision</th>
<th>Duration (months)</th>
<th>Cr (µg/l)</th>
<th>Co (µg/l)</th>
<th>Cup inclination (°)</th>
<th>Cup anteverision (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.3</td>
<td>-</td>
<td>Pain, effusion</td>
<td>28</td>
<td>-</td>
<td>-</td>
<td>62</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>17.7</td>
<td>14.8</td>
<td>Pain, effusion</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>58</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>91.8</td>
<td>64.0</td>
<td>Pain, effusion</td>
<td>8</td>
<td>36</td>
<td>88</td>
<td>65</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>38.0</td>
<td>28.8</td>
<td>Pain, effusion</td>
<td>27</td>
<td>5</td>
<td>8</td>
<td>50</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>32.9</td>
<td>23.1</td>
<td>Pain, effusion</td>
<td>17</td>
<td>15</td>
<td>19</td>
<td>51</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>1.8</td>
<td>-</td>
<td>Fracture</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7 d</td>
<td>3.1</td>
<td>-</td>
<td>Fracture</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>46</td>
<td>31</td>
</tr>
<tr>
<td>8 d</td>
<td>2.5</td>
<td>-</td>
<td>Fracture</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9 d</td>
<td>1.7</td>
<td>-</td>
<td>Fracture</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>47</td>
<td>7</td>
</tr>
<tr>
<td>10 d</td>
<td>2.0</td>
<td>-</td>
<td>Fracture</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1 shows that, where pain and effusion were associated with the reason for revision, the out of roundness values were much greater, up to a maximum of 91.8 µm in the case of the femoral head from patient 3. The roundness traces from this component are shown in Fig. 3, and it can be seen that the maximum out of roundness is on the plane 3 mm from the pole of the femoral head. The contrast with the circular shape of Fig. 2 is notable.

The roundness traces from the retrieved femoral head of patient 5, another 'effusion' failure, are shown in Fig. 4. On the three planes, maximum out of roundness values of 18.3, 24.9 and 32.9 µm have been measured. This implies that wear or deformation was taking place over a substantial area of the femoral head. Such a result was seen with all of the ASR™ femoral heads associated with ‘effusion’ failures.

For the acetabular cups associated with pain and effusion, out of roundness traces for patients 3 and 4 are shown in Figs. 5 and 6, respectively. All of these signify that the greatest wear or deformation took place towards the edge of the cup, as this is indicated by the third measurement, that on a plane 3 mm below the rim. Fig. 5 shows that out of roundness values of 38.5, 48.1, and 64.0 µm were measured on the three planes, and therefore that wear or deformation took place deep inside the acetabular cup. With out of roundness values on the three planes of 15.1, 20.8 and 28.8 µm Fig. 6 similarly shows that with this component wear or deformation again occurred over a substantial area of the articulating surface of the cup.

The clinical data in Table 1 shows that patients who suffered from groin pain and an effusion also had high ion levels. In turn they had high angles of acetabular cup anteversion and/or inclination. It has been shown in the larger ASR™ cohort that high anteversion and inclination are directly related to high ion levels [26].

Please cite this article as: Joyce TJ, et al. A study of the wear of explanted metal-on-metal resurfacing hip prostheses. Tribol Int (2010), doi:10.1016/j.triboint.2010.04.006
Roundness
Data No.: 1, 2, 3

1 P-P=18.3µm
2 P-P=24.9µm
3 P-P=32.9µm

Fig. 4. Roundness traces taken of an ASR™ femoral head associated with an 'effusion' and groin pain failure (patient 5).

Roundness
Data No.: 1, 2, 3

1 P-P=15.1µm
2 P-P=20.8µm
3 P-P=20.8µm

Fig. 6. Roundness traces taken of an ASR™ acetabular cup associated with an 'effusion' and groin pain failure (patient 4).

4. Discussion

For a new ASR™ component, a typical out of roundness value would be of the order of 4 µm [33]. Therefore, from Table 1, it can be seen that all the five components taken from patients whose femurs had fractured (maximum out of roundness 3.1 µm) showed minimal deformation or wear after removal. As such they serve as controls against which the effusion failure components can be judged. Additionally, the out of roundness values of the 'fracture' components help to verify the accuracy of the measuring technique and match such measurements to independent data [33]. From a clinical failure analysis point of view, it could be concluded from the out of roundness results related to the early 'fracture' failures that these are not correlated to the wear of material from the articulating surfaces of the femoral components. However, the contrast with the 'effusion' failure components and their high out of roundness values is stark.

Could the 'effusion' components have shown deformation rather than wear? There are two key reasons why deformation is unlikely to have caused the high out of roundness values seen with the effusion failures. Firstly, all retrieved components were likely to have been subject to a similar range of loading within the various patients. Why then should only the components associated with effusion failures have deformed when all those associated with fracture of the femur have shown minimal deformation? Secondly, and perhaps most importantly, deformation cannot explain the high ion levels seen in the blood of patients with effusion failures but wear from the articulating surfaces of the metal-on-metal resurfacing hip prostheses can explain these high ion levels. Therefore the out of roundness measurements provided valuable information about the wear of the metal-on-metal resurfacing prostheses.

From the out of roundness measurements, the maximum out of roundness (defined as the deviation from a perfect circle) was taken as indicative of the maximum linear wear depth. On the 'effusion' failures these linear wear depths were significant and contrasted with the early fracture failures where it was not possible to separate out any wear from the manufacturer's tolerances for out of roundness.

To the authors' best knowledge, the dichotomy between low wear components associated with early femoral fracture and high wear explants associated with groin pain and effusion has not previously been reported. Moreover we are not aware of other researchers having offered out of roundness measurements from failed resurfacing hip prostheses. A roundness measuring machine
has recently been used to measure the wear of explanted metal-on-metal total hip prostheses [34]. The authors described these explanted prostheses as ‘bearing couples that had been in successful use for more than seven years (range, seven to thirty-four years)’ [34]. Therefore they differ from the cohort described in this paper, which concerns itself with resurfacing devices which are considered failures.

However, other researchers have offered wear assessments of retrieved metal-on-metal resurfacing hip prostheses, but using coordinate measuring machines (CMM). While CMM do have certain advantages, they lack the sensitivity found with a dedicated roundness measuring machine. For example in an early study it was noted that when a CMM with a claimed accuracy of ±2 μm was used, only a limited number of the explanted resurfacing hip components could be measured, which were outside of manufacturing tolerance [35]. Where wear could be measured, for paired head and cup components, wear depths per year of 4.4, 16 and 58 μm were offered. If the roundness values reported in the current paper are taken as the maximum wear depth then the values of Beaufle et al. would compare with values of 21.7, 233.7, 29.7 and 39.5 μm wear depth per year for patients 2, 3, 4 and 5, respectively.

Using another CMM to examine various designs of explanted resurfacing hip prostheses, values of maximum wear in the range of <2–164 μm have been reported [16]. Two microns was therefore offered as the resolution of the CMM. As with other studies, there is a considerable range of wear values, but the data of Campbell et al. offers agreement with the values of up to 98.1 μm found in this study.

Another recent study used a CMM to look at a large number of explanted resurfacing components of various designs [36]. An accuracy of 3 μm was claimed for the CMM. Wear was given in the form of mm³ per day. Again, a wide variation in wear rates was reported, between 22 and 27 times higher wear being measured in ‘edge-loaded’ implants rather than non edge-loaded. The authors also noted that two-thirds of their components were revised due to fracture of the femur.

Most recently, the wear of eight femoral and two acetabular explanted BHR components was reported [19]. Time in vivo of the components was between 7 and 24 months, similar to the 1-28 months reported in this paper. A CMM was used and a spatial resolution of 1 μm was claimed. Linear wear of the eight BHR heads was measured to be between 1.7 and 44.7 μm, which compares with a range of 1.7–98.1 μm reported in this paper for ten ASR™ femoral heads. For the two BHR acetabular components Witzleben et al. reported linear wear of 9.2 and 31.5 μm, which compares with 14.8–64.0 μm for the four ASR™ acetabular components measured in this paper. These results, based on a limited sample size of explanted components, suggest that the wear of ASR™ prostheses is greater than that of the BHR.

A conventional method of offering wear is to give the weight change. Unfortunately, however, for explanted components this is generally not possible, as the precise weight prior to implantation is not recorded by manufacturers. Therefore, given that gravimetric methods of wear measurement are inappropriate, wear measurements based on dimensional changes are employed, such as out of roundness and those obtained from a CMM. It should also be noted that the exact engineering specification relating to the prostheses prior to implantation cannot be known precisely. Manufacturers tend not to release figures related to their products such as out of roundness. Moreover, as would be expected from modern engineering components like hip prostheses, such values fall within a range of tolerances and again manufacturers do not disclose such figures which could be valuable to competitors. However some independent comparative measurements have been done. For example, and most applicable to this study, out of roundness values of the order of 4 μm have been offered for ASR™ components [33].

Can the high angles of acetabular cup inclination and anteversion have been a contributing factor to the failure of these ASR™ resurfacing hip prostheses? This is certainly possible as at high angles of inclination, not only higher metal ion levels have been reported in patients [37], but also edge loading [16]. Such edge loading was also seen in the form of rim damage on the explanted cups studied. In addition to noting the correlation between acetabular cup inclination and ion levels in patients fitted with resurfacing hip prostheses, the importance of edge loading has also been reported recently [38]. These authors postulated that at such high angles the ‘arc of cover’ between femoral head and acetabular cup was reduced and they found that smaller arcs of cover were associated with higher wear [38]. Such an explanation could also apply to the ASR™ prostheses discussed in the current paper. One study has reported a twenty-one to twenty-seven fold greater wear rate in rim loaded hip resurfacing explants compared with non-rim loaded samples [17]. It should also be recognised that in ‘conventional’ metal-on-polymer total hip prostheses failure rates have increased with inclination of the acetabular cup [39] while the importance of acetabular position to low wear has also been identified in ceramic-on-ceramic total hip prostheses [40].

In a recent paper a number of explanted ASR™ components, all associated with pain and effusion in patients, were measured to determine their surface roughness and thus the lubrication regime during gait [18]. It was found that all surfaces had roughened compared with an unused pair of head and cup ASR™ components, so much so that instead of operating under fluid film lubrication during gait, they would instead operate in the boundary lubrication regime [18]. In the boundary lubrication regime, with its preponderance of surface-to-surface interaction, increased wear would be expected. In turn this would result in higher volumes of metallic wear debris. In addition the relatively large size of the resurfacing hip prostheses means that, under boundary lubrication conditions, wear is taking place over a large sliding distance so that volumes of wear debris are maximised. Recently, large amounts of particulate wear debris have been linked with a toxic reaction in localised cells and the attendant failure of resurfacing hip prostheses [27]. Clearly the high wear could also be directly linked with the higher ion levels, which were measured in the patients from whom these explanted resurfacing hip prostheses were obtained [26].

5. Conclusion

From a series of explanted metal-on-metal resurfacing hip prostheses with a known clinical history it was found that failures associated with effusion and pain in patients were associated with values of out of roundness far higher than at manufacture. These changes were associated with wear of the devices and this assertion was supported by the high ion levels seen in the blood of patients. In addition, all of these failures were associated with cups which had been positioned at high angles of inclination and anteversion. The effusion failures contrasted markedly with those associated with early fracture of the femur. Here, all components had an out of roundness no greater than 3 μm, which strongly implied that little or no wear of these components had taken place prior to retrieval.

References


Please cite this article as: Joyce TJ, et al. A study of the wear of explanted metal-on-metal resurfacing hip prostheses. Tribol Int (2010). doi:10.1016/j.triboint.2010.04.006
Chapter Six
Volumetric wear assessment of failed metal-on-metal hip resurfacing prostheses

J.K. Lord a,*, D.J. Langton a, A.V.F. Nargol b, T.J. Joyce a

a School of Mechanical & Systems Engineering, Stephenson Building, Clarendon Road, Newcastle University, Newcastle upon Tyne NE1 7RU, England, UK
b Joint Replacement Unit, University Hospital of North Tees, Hardwick Road, Stockton TS19 8PE, UK

ARTICLE INFO

Article history:
Received 9 February 2011
Received in revised form 12 July 2011
Accepted 22 July 2011
Available online 30 July 2011

Keywords:
Joint replacement
Hip resurfacing prosthesis
Metal-on-metal
Wear determination
Surface analysis

ABSTRACT

Recent advancements in hip arthroplasty have allowed the operation to boast excellent results and high survivorship. However, failures do still occur and a major cause is complications arising from wear debris. It is essential therefore that debris is minimized by reducing wear at the bearing surface. One proposed method of achieving this wear reduction is through the use of metal-on-metal articulations. One of the latest manifestations of this biomaterial combination is in designs of hip resurfacing which are aimed at younger, more active patients who might wear out a conventional metal-on-polymer hip prosthesis. However, do these metal-on-metal hip resurfacings show less wear when implanted into patients?

Using a co-ordinate measuring machine and a bespoke computer program, volumetric wear measurements for retrieved Articular Surface Replacements (ASR®-DePuy) metal-on-metal hip resurfacings were undertaken. Thirty-two femoral heads and twenty-two acetabular cups were measured. Acetabular cups exhibited mean volumetric wear of 29.00 mm³ (range 1.35–109.72 mm³) and a wear rate of 11.02 mm³/year (range 0.30–63.59 mm³/year). Femoral heads exhibited mean wear of 22.41 mm³ (range 0.72–134.22 mm³) and a wear rate of 8.72 mm³/year (range 0.21–31.91 mm³/year). In the 22 cases where both head and cup from the same prosthesis were available, mean total wear rates of 21.66 mm³/year (range 0.51–95.50 mm³/year) were observed. Compared with in many vitro tests, these are significantly higher than those expected in a well functioning metal-on-metal hip resurfacing prosthesis and are of concern.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Hip prostheses are an important tool for reducing pain and restoring function to patients with musculoskeletal disorders such as arthritis. Recent advancements have made hip arthroplasty a common operation, with over 65,000 primary hip replacements being performed in the UK in 2009 [1]. One such advancement was the re-introduction of metal-on-metal (MoM) hip prostheses. These so called ‘second-generation’ MoM prostheses are typically made from cobalt–chromium–molybdenum alloy (CoCrMo) and have been shown to have improved wear properties over more traditional metal-on-polyethylene (MoP) articulations [2]. Hip simulator studies have shown MoM wear rates in the region of 0.5–1 mm³/million cycles [3,4], a 10- to 40-fold decrease compared with 10–20 mm³/million cycles [5,6] for MoP.

Another more recent advancement was the introduction of MoM resurfacing hips. Compared with total hip replacement (THR), these devices are intended to conserve bone and offer a more physiological load transfer [7]. For these reasons, they are commonly implanted in younger, more active patients [8]. Additionally, they are designed to operate under more favourable fluid film lubrication, for at least part of the gait cycle, which should reduce wear, compared with the boundary or mixed lubrication of total hip replacement [9]. Simulator studies of MoM resurfacings have demonstrated wear rates in the region of 0.03–3.59 mm³/million cycles [10–12], though it should be noted that the upper end of this region is during the ‘running-in’ process, whereby wear is higher in the first few months after implantation than in the remainder of the prosthesis life [10,11].

Although short-term survival studies of MoM hip resurfacing prostheses have been encouraging [8,13,14], there are still many reported cases of early failure, with a 6.3% revision rate at 5 years reported in the UK in 2010 [1]. There are numerous modes of failure [15], and an underlying cause in most cases is wear at the bearing surface and consequent creation of metallic debris [16,17]. An accurate calculation of the amount of wear occurring from the prosthesis is therefore an important step in identifying the causes of wear to improve the longevity of hip prostheses.

Several techniques have been described for wear estimation of both MoP and MoM hip prostheses. Attempts have been made to measure wear in vivo using radiographic data [18], and various formulae have been proposed [19]. Radiographic estimation has shown resolution in the region of 0.055–0.3 mm [20]. This is particularly inaccurate for metal-on-metal prostheses where wear rates
as low as 0.006 mm/year (6 μm/year) have been demonstrated [21]. Maximum linear wear depth has been used as a quantification of wear [21, 22]. Although this gives an indication, it does not account for variable wear across the component surface. A component with deep, isolated wear may have lost less material in volumetric terms than another component with shallow wear over a large area. Calculation of volumetric wear would therefore be a much more useful tool than wear depth [23]. In vitro, volumetric wear is commonly calculated gravimetrically [24, 25]. However, this is not practical with retrieved prostheses as original component masses are unknown.

There are a handful of studies offering ex vivo volumetric wear rates of MoM hip prostheses. In 1996, Kothari et al. used a coordinate measuring machine (CMM) to evaluate 22 retrieved McKee–Farrar total hip replacements [26]. Three hundred and twenty-five points were measured on each sample and the ‘accuracy’ of the CMM used was ±5 μm. Although accuracy in this context was not explicitly defined in this paper, it is reasonable to accept the definition offered in ISO 10360 “Geometrical product specification (GPS) – Acceptance and reverification tests for coordinate measuring machines (CMM)”. Here accuracy is defined as the maximum permitted form error when a reference sphere is measured with 25 evenly distributed points [27]. With accuracy of ±5 μm and possible wear rates as low as 6 μm/year as noted above, there is potential for large errors. Indeed, in 2006 Becker et al. evaluated the influence of measurement accuracy in CMM based approaches and recommended a minimum accuracy of ±2 μm [28]. However a later study by the same authors comparing two CMMs (a “standard precision” 2.9 μm and a “high precision” 0.8 μm) concluded that a high precision CMM is “essential for assessing wear in modern hard-on-hard bearings” [29]. Becker et al. examined retrieved 28 mm MoM THRs and these were all femoral heads with no acetabular cups examined [29]. Morlock et al. reported in 2006 on a CMM based volumetric wear measurement methodology [30]. This method was then used in 2008 to report on 267 retrieved hip resurfacing components (although wear data on only 58 components [including 26 pairs] was tabulated in the paper) [15]. The CMM used by Morlock et al. was said to be accurate to ±3 μm. Bills et al. published a CMM based volumetric wear measurement method in 2007 [31], as did Witzleb et al. in 2009 [32]. Bills et al. stated that most average CMMs have an accuracy of approximately 3 μm and as such would not be accurate enough for useful volumetric measurements of hard-on-hard orthopaedic bearings [31]. Both Bills et al. and Witzleb et al. used CMMs with accuracy of ±1 μm, but the methods were applied to small numbers of retrievals (4 and 10 components respectively). Neither set of authors gave the articulating diameters of the hip components they measured. This retrieval and measurement data is summarised in Table 1. Perhaps most importantly, of the above publications, only Morlock et al. [30] and Becker and Dirix [29] provided any data on the accuracy of their calculations. Morlock et al. claimed errors for volumetric calculations within 8% when applying their method to a simulated data set, though data was not offered to support this. Because the data set was simulated, this error value is only for the calculation of wear and does not indicate errors arising from their CMM measurements or from differentiating between manufacturing tolerance and wear.

Becker et al. showed percentage error for the high precision CMM decreasing from approximately 15% to 2% when linear wear depths were increased from 3 μm to 15 μm. The standard precision CMM varied from 55% to 10% errors across the same range. However, neither Morlock et al. nor Becker et al. offered their actual volumetric wear and so it is not possible to quantify these percentage errors in terms of mm³.

The present paper offers a method for calculating volumetric wear of retrieved MoM resurfacing hip prostheses using a combination of coordinate measuring machine data and Matlab (The Mathworks, Inc.). A validation study is included and the method is applied to 54 retrieved hip resurfacing components to assess in vivo wear.

2. Materials and methods

2.1. Materials

Research Ethics Committee approval was obtained for all work carried out (REC/09/H0905/41). Thirty-two femoral and twenty-two acetabular components were obtained from revision surgeries. All were MoM hip resurfacing prostheses of a single design, the Articular Surface Replacement (ASRTM. DePuy, Leeds, United Kingdom). All were implanted and removed by the same surgeon (AVFN). The ASRTM is a design cast from high carbon content CoCrMo alloy with heat treated acetabular components [7]. The nominal articulating diameters for the components were between 43 mm and 53 mm. Retrieval occurred after 2–58 months in vivo and was attributed to four main reasons: adverse reactions to metal debris (ARMD [33]) (19 heads, 18 cups); fracture of the femur (10 heads, 2 cups); avascular necrosis (AVN) (2 heads, 1 cup); and infection (1 head, 1 cup). Fractures of the femur were further subdivided into early fracture (up to 7 months in vivo) and late ARMD fracture (22–53 months in vivo) (Table 2).

2.2. Data collection

After retrieval, the explants were soaked in 10% formalin for one week before being rinsed thoroughly in water. The articulating surfaces were cleaned carefully using acetone and a lint-free cloth in order to remove loose material and minimise spurious measurements. The samples were then scanned using a Mitutoyo LEGEX322 co-ordinate measuring machine (CMM). The CMM was dedicated to measuring only the hip prostheses discussed in this study. Co-ordinate measuring machines rely on a contact stylus which is touched against the sample surface and measures a point in space.

Table 2

<table>
<thead>
<tr>
<th>Failure mode</th>
<th>Number of cups</th>
<th>Number of heads</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMD</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>AVN</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Early fracture</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>ARMD fracture</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Infection</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1
Summary of previous CMM based volumetric wear calculations for explanted MoM hip prostheses.

<table>
<thead>
<tr>
<th>Lead author [reference]</th>
<th>Year</th>
<th>Number of components measured</th>
<th>CMM accuracy (μm)</th>
<th>Number of points taken</th>
<th>Size of errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kothari et al. [26]</td>
<td>1996</td>
<td>22 pairs</td>
<td>±5</td>
<td>325</td>
<td>Not given</td>
</tr>
<tr>
<td>Bills et al. [31]</td>
<td>2007</td>
<td>2 pairs</td>
<td>±1</td>
<td>Not given</td>
<td>Up to 8%</td>
</tr>
<tr>
<td>Morlock et al. [15]</td>
<td>2008</td>
<td>58 (including 26 pairs)</td>
<td>±3</td>
<td>Not given</td>
<td>Up to 8%</td>
</tr>
<tr>
<td>Witzleb et al. [32]</td>
<td>2009</td>
<td>10 (including 2 pairs)</td>
<td>±1</td>
<td>1297</td>
<td>Not given</td>
</tr>
<tr>
<td>Becker and Dirix [29]</td>
<td>2009</td>
<td>44 femoral heads</td>
<td>±2.9 and ±0.8</td>
<td>15,960</td>
<td>Max. 15% and 55%</td>
</tr>
<tr>
<td>Present study</td>
<td>2009</td>
<td>54 (including 22 pairs)</td>
<td>±0.9</td>
<td>Up to 7128</td>
<td></td>
</tr>
</tbody>
</table>

as noted [32], the method in 2008 to report on 267 retrieved hip resurfacing components (although wear data on only 58 components [including 26 pairs] was tabulated in the paper) [15]. The CMM used by Morlock et al. was said to be accurate to ±3 μm. Bills et al. published a CMM based volumetric wear measurement method in 2007 [31], as did Witzleb et al. in 2009 [32]. Bills et al. stated that most average CMMs have an accuracy of approximately 3 μm and as such would not be accurate enough for useful volumetric measurements of hard-on-hard orthopaedic bearings [31]. Both Bills et al. and Witzleb et al. used CMMs with accuracy of ±1 μm, but the methods were applied to small numbers of retrievals (4 and 10 components respectively). Neither set of authors gave the articulating diameters of the hip components they measured. This retrieval and measurement data is summarised in Table 1. Perhaps most importantly, of the above publications, only Morlock et al. [30] and Becker and Dirix [29] provided any data on the accuracy of their calculations. Morlock et al. claimed errors for volumetric calculations within 8% when applying their method to a simulated data set, though data was not offered to support this. Because the data set was simulated, this error value is only for the calculation of wear and does not indicate errors arising from their CMM measurements or from differentiating between manufacturing tolerance and wear.
The LEGEX322 CMM was fitted with a SP-25 scanning head. Such a head allows measurements to be taken continuously, in contrast to most CMMs which take a measurement by 'pecking', i.e. gently colliding with an object to trigger a measurement point, then retreating, repositioning and re-colliding to take the next measurement. The SP-25 scanning head allowed continuous contact and data measurement; compared with a conventional head it reduced typical measurement times from over 8 h down to 20 min. A second ISO 10360 definition of accuracy exists for scanning heads: the maximum permissible error of measured radii when a referencesphere is scanned along 4 defined lines [34]. With a scanning head, accuracy decreases as scanning length increases. For the present set-up, the accuracy stated by Mitutoyo is ±0.008 mm, where L is the measurement length in mm. For the largest component in this study (53 mm diameter head), the largest value of L is 41.6 mm, corresponding to measurement accuracy of 0.88 µm. Femoral head samples were held in place by their stem using a self-centring three-jawed chuck (Fig. 1) to prevent movement during the scanning process. In order to prevent deformation from the same three-jawed chuck, acetabular samples were held in a clay mould.

Programs were written in the CMM software (MCOSMOS') to allow head and cup components to be scanned. Determination of the original spherical surface is critical to the accurate calculation of volumetric material loss. The program aimed to identify the origin of the spherical components from as wide an area of the articulating surface as possible. For femoral components, four points were taken at 90° intervals around the full 360° of the equator in the X-Y plane. Three points were taken in the Z-X plane at 25° intervals. From these seven points a sphere was calculated. If the sphericity of this initial sphere was found to be within the manufacturing tolerance of 4 µm (data supplied by manufacturer) then a Cartesian co-ordinate system was defined with the origin set according to the centre of the sphere (Fig. 1). If the sphericity was outside of the 4 µm tolerance, for example due to one of the measurements being taken within a worn area of the component, then, using the MCOSMOS software, the coordinate system was rotated by 10° about the z axis and the process repeated until a suitably unworn area was located. In the rare event of the method failing to find a satisfactory form after 36 passes, the area over which the points were taken was restricted to a 300° area around the equator with the aim being to minimise the probability of contacting a worn area. Even in highly worn components, wear is typically localised and this adjustment always proved to be successful in allowing a spherical origin to be found. For acetabular cups, the process was identical to that of the femoral program except that areas within 30° of the rim of the cup were not used in the calculation of the sphere. This decision was based on the principle that in most heavily worn cups, the wear is located primarily at the rim of the cup, a result which has also been reported elsewhere [15,35,36]. The procedure described above provided a rapid methodology to determine the approximate centre of the sphere.

To determine the definitive centre of the sphere, 100 points were taken in the YZ plane moving from equator to equator for the femoral heads but limited to a 120° scan about the pole in the case of the acetabular cups (for reasons described above). The coordinate system was then rotated 22.5° about the z axis and the process repeated seven times, so that a total of 800 points were taken. Any points which were calculated to be greater than or less than 4 µm deviation from the initial spherical form, as determined from the initial seven points, were discarded as they were unlikely to represent the original surface and so could not be used. All other points were retained and used in the calculation of the second sphere. The centre of the second sphere was then taken as the definitive origin. This method was developed on the principle that even heavily worn samples typically show a sharply demarcated transition between worn and unworn areas. Points taken over worn areas are highly likely to be much greater than 4 µm in deviation from the original calculated form and are not used to determine the definitive origin.

Finally, scans were taken every 5° around the circumference, starting 5 mm below the equator and converging on the pole. This was done for femoral heads and acetabular cups and allowed for between 6048 and 7128 data points to be collected for each head and 3024–4104 for each cup (depending on the articulating diameter of the component). At each point the 3-dimensional position was recorded in Cartesian co-ordinates, relative to the centre of the sample.

2.3. Volumetric wear calculation

Co-ordinate data collected from the CMM was read in to a Matlab program written by the first author of this paper. The data was split into three matrices representing the Cartesian co-ordinates at each measured point. Using these points, the surface of the component was reconstructed. Each point was connected to its adjacent points to form gridsquares (Fig. 3). The exact radius of the component at each measured point was calculated using \( \sqrt{X^2 + Y^2 + Z^2} = r \) where X, Y, and Z represent the Cartesian co-ordinates of the point and r is the radius of the sphere. These radii were presented in a histogram similar to that shown in Fig. 4. In the most worn samples wear occurred unevenly across the surface resulting in a wide range of measured radii values, with any given value occurring relatively infrequently. In contrast, the unworn areas exhibited high uniformity of radii across a large portion of the surface. The ‘worn’ radii are demonstrated in Fig. 4 by the long ‘tail’ towards the left showing
A wide range of infrequent values corresponding to the wear area, while the narrow range, towards the right, of frequently occurring radii values correspond to the unworn area. Thus, identification of the original 'unworn' radius was made by calculating the mode of the measured radii. Note that, despite the high calculated wear of 39.78 mm³, 59% of points (3936 of 6696) are within manufacturing tolerance of ±4 μm. With the unworn radius known, the linear wear depths were calculated by subtracting the measured radii data from the unworn radius. Any measured radii greater than the original radius were regarded as being unworn material and simply part of the manufacturing form within the manufacture's tolerance of sphericity or roundness. As these radii represented unworn areas, they were disregarded in further calculations of wear.

A mean wear depth for each gridsquare on the reconstructed surface was calculated by taking a mean of the depths at the four corners. The area of each gridsquare was then calculated and multiplied by the corresponding mean wear depth to give a wear volume. These individual volumes were then summed for the entire component to give an overall volumetric wear. A diagram showing the reconstructed surface and coloured according to linear wear depth was produced. Such wear diagrams were used to characterise the wear pattern and severity for all analysed components (Figs. 5 and 6). Again, there is a clear demarcated transition between the worn and unworn areas in the highly worn sample (Fig. 5).

2.4. Validation

2.4.1. Ceramic masterball
Validation occurred in two stages. First, a 20 mm nominal diameter (19.9881 mm actual diameter) ceramic masterball was scanned and processed using the method described above. Due to the ceramic material and the tight manufacturing tolerances used for creating a masterball (within 0.5 μm sphericity), this component was expected to show very little deviation in radii. It was unworn and therefore any volumetric "wear" measured was...
expected to be due to form error from manufacture rather than material removal.

2.4.2. Gravimetric comparison

Secondly, the method described in Sections 2.2 and 2.3 was validated against established gravimetric methodology using a sample femoral head component. The component was an un-implanted 36 mm nominal diameter total hip replacement head of CoCrMo. The sample was cleaned thoroughly in an acetone bath for 5 min. It was then left to dry for 1 h on a lint-free cloth and then weighed on a high precision scale (Denver Instrument, sensitivity 0.1 mg). The sample was weighed six times, and an average taken. The sample was also scanned using the CMM, so that the effect of form error on apparent "wear" could be evaluated.

The CoCrMo femoral head sample then had a quantity of material removed to simulate wear. As the intention here was simply to remove material, sandpaper was used. In this way it was possible to produce a wear pattern of variable depth across the surface. Following material removal, the sample was cleaned to remove any debris, weighed, measured and analysed again using the CMM. Three scans were taken as per the methodology in Section 2.2. The sample was removed and replaced between scans; this was done to assess repeatability of measurements for a given sample. More material was then removed from the femoral head and the process of gravimetric and dimensional (CMM) measurements repeated. In total there were three stages of material removal, with three scans taken at each stage. The volumetric wear calculated from each scan was compared to the volume of material lost determined by the gravimetric method. This comparison was done to test accuracy of the CMM measurement methodology as volumetric wear increased. It was assumed that volumes obtained gravimetrically represented the 'gold standard' (for the 36 mm CoCrMo femoral head, a change in weight of 0.1 mg was equivalent to a volumetric change of 0.012 mm$^3$) and the accuracy of the CMM method was assessed against the gravimetric method. Since the CMM measurement method is identical for femoral heads and acetabular cups, it is reasonable to assume that all validation carried out on a head is equally applicable to a cup.

3. Results

3.1. Validation

3.1.1. Masterball

The CMM method calculated a radius for the masterball of 9.9945 mm. This is 0.4 $\mu$m larger than the actual radius. This error is within the scanning limits of the CMM (accurate to within 0.9 $\mu$m). The calculated "wear" volume was 0.04 mm$^3$. This is not actual wear but form error inherent in manufacture. Moreover it is a trivial volume compared to those being measured on retrieved prostheses, presented in Section 3.2. The results from the masterball measurements are summarised in Table 3. The measured radial deviations are presented as a histogram in Fig. 7. The measurements were evaluated and provide a Gaussian distribution around the zero point, indicative of variations arising from a manufacturing process rather than from wear.

Each time the CMM was used, linear wear depths calculated by the method described in this paper were compared with those produced by the CMM software. In every case, both methods were in exact agreement.

3.1.2. Gravimetric

The initial scan of the as manufactured CoCrMo femoral component revealed form error leading to a 'wear' calculation of 0.4 mm$^3$. As will be seen this value is small compared to the volumes presented in the wear measurement methodology as volumetric wear increased. It was assumed that volumes obtained gravimetrically represented the 'gold standard' (for the 36 mm CoCrMo femoral head, a change in weight of 0.1 mg was equivalent to a volumetric change of 0.012 mm$^3$) and the accuracy of the CMM method was assessed against the gravimetric method. Since the CMM measurement method is identical for femoral heads and acetabular cups, it is reasonable to assume that all validation carried out on a head is equally applicable to a cup.

3. Results

3.1. Validation

3.1.1. Masterball

The CMM method calculated a radius for the masterball of 9.9945 mm. This is 0.4 $\mu$m larger than the actual radius. This error is within the scanning limits of the CMM (accurate to within 0.9 $\mu$m). The calculated "wear" volume was 0.04 mm$^3$. This is not actual wear but form error inherent in manufacture. Moreover it is a trivial volume compared to those being measured on retrieved prostheses, presented in Section 3.2. The results from the masterball measurements are summarised in Table 3. The measured radial deviations are presented as a histogram in Fig. 7. The measurements were evaluated and provide a Gaussian distribution around the zero point, indicative of variations arising from a manufacturing process rather than from wear.

Each time the CMM was used, linear wear depths calculated by the method described in this paper were compared with those produced by the CMM software. In every case, both methods were in exact agreement.

3.1.2. Gravimetric

The initial scan of the as manufactured CoCrMo femoral component revealed form error leading to a 'wear' calculation of 0.4 mm$^3$. As will be seen this value is small compared to the volumes presented in the wear measurement methodology as volumetric wear increased. It was assumed that volumes obtained gravimetrically represented the 'gold standard' (for the 36 mm CoCrMo femoral head, a change in weight of 0.1 mg was equivalent to a volumetric change of 0.012 mm$^3$) and the accuracy of the CMM method was assessed against the gravimetric method. Since the CMM measurement method is identical for femoral heads and acetabular cups, it is reasonable to assume that all validation carried out on a head is equally applicable to a cup.

Table 3

| Measurement of the ceramic masterball indicating the size of the errors in the presented wear measurement method. |
|-----------------------------------------------|-----------------------------------------------|
| Radius (mm) | Volume (mm$^3$) |
| Actual | 9.9941 | 0 |
| CMM | 9.9945 | 0.04 |
| Difference | 0.0004 | 0.04 |

Note: Actual size supplied with masterball; difference in radius of 0.4 $\mu$m is within claimed accuracy of LEGEX322 (0.9 $\mu$m).

Fig. 7. Histogram of masterball scan. Distribution was evaluated and is Gaussian. Minimum point $= -1.2 \mu$m. Maximum point $= +1.3 \mu$m. Positive 'linear wear depth' indicates manufacturing form and the ability of the CMM wear measurement methodology to identify this. Calculated wear volume $= 0.04$ mm$^3$. 
Table 4

Comparison of Matlab method to gravimetric method. Mean volume is the mean of the 3 measurements at each stage. Mean absolute error is the mean of the error of the 3 measurements. St. Dev. = Standard deviation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Material removal (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
</tr>
<tr>
<td>Gravimetric - weight (mg)</td>
<td></td>
</tr>
<tr>
<td>Gravimetric converted to volume (mm³)</td>
<td>42.0</td>
</tr>
<tr>
<td>Mean Matlab volume (mm³)±St. Dev.</td>
<td>5.1</td>
</tr>
<tr>
<td>Mean absolute Matlab error (mm³)</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Presented in Section 3.2. Results of the gravimetric validation procedure are shown in Table 4. The mean Matlab calculation of wear volume is presented from the three CMM scans at each level of material removal, along with the standard deviation of these values. The difference between the Matlab and gravimetric method is presented as a percentage. As can be seen, at the three stages the mean absolute error was 0.53, 0.50 and 0.24 mm³ respectively.

3.2. Retrieval analysis

 Retrieval analysis and wear data for all components is shown in Table 5. Acetabular cups exhibited mean volumetric wear at revision of 29.09 mm³ (range 1.35-109.72 mm³). The mean wear rate was 11.02 mm³/year (range 0.30-63.59 mm³/year). The femoral heads had mean volumetric wear of 22.41 mm³ (range 0.72-134.22 mm³). The mean wear rate was 8.72 mm³/year (range 0.21-31.91 mm³/year). Table 6 shows this data summarised by failure mode. As can be seen, combined wear rates for the AVN (0.51 mm³/year) failures appear to be within the region expected from simulator testing (0.03-3.59 mm³/million cycles). Wear rates for the ARMD components however are high (17.68 mm³/year).

4. Discussion

 The method presented in this paper provides a validated assessment of volumetric wear of retrieved metal-on-metal hip components. The absolute error was within 0.53 mm³ for all scans (Table 4, row 4). As noted in Section 1, only two previous studies were found that offered the error range for their CMM based wear measurement method. In one study, the error was quoted to be within 8% [30] although data was not offered to validate this. The second study offered errors from 2% to 10% [29]. The present method has been shown to be accurate to within 0.5 mm³ for a range of wear volumes. Clearly, as the volume of wear increases, the percentage errors decrease. It is not possible to discuss how this level of accuracy compares to other CMM based volumetric calculations as the other studies mentioned earlier [25,31,32] did not include validation. To the authors' best knowledge, this is the first time that a validation of a CMM based wear volume calculation for metal-on-metal hip prostheses has been offered in a peer-reviewed paper.

 It should be noted that in the current study, scans were taken every 5° around the circumference regardless of component size. On larger components there are therefore slightly larger gaps between data points, particularly towards the equator. For example, the widest gap on a 43 mm component (where 6048 points were taken) would be 1.88 mm. For a 53 mm component (7128 points), this increased to 2.31 mm. However, the size of each grid square is taken into account during volume calculation and so this difference in gap size will not significantly affect results, even in cases of high edge wear which was commonly seen on the acetabular components. Moreover, as can be seen from Table 1, the number of points taken in the present study is greater than all but one other study, where such information has been offered.

The method of wear measurement described in this paper allows for clear characterisation of the wear area, including coverage and depth (although, as this paper is primarily concerned with volumetric wear, these results are not included here). It is likely that the region where wear occurs on mated components is linked to their position in vivo. Acetabular cup position in vivo has been used to explain high concentrations of metal ions in the blood of patients who are fitted with cups at high inclination angles [38]. These high ion concentrations are likely to occur with edge-loaded cups; here the major contact forces (and thus the main region of wear) are concentrated towards the rim of the cup [39]. Knowing wear volumes of explanted components and ion concentrations in the patients they were removed from allows these two key issues to be linked.

Table 6 shows wear volumes matched against the five failure modes of ARMD, ARMD fracture, early fracture, AVN and infection. Although the wear rate appears high for early fracture components, this may be due to the 'running-in' process [10,11]. Early fracture of the femoral neck is a recognised complication, which is claimed to have been minimised by careful implantation and patient selection [15,40]. The three AVN failures are worthy of note. One AVN head was retrieved after 38 months in vivo and its wear rate of 0.41 mm³/year is comparable to those found in simulator studies of MoM hip resurfacings, which have been demonstrated in the range of 0.03-3.59 mm³/million cycles [10-12]. One million cycles is generally agreed to be equivalent to 1 year in vivo [41]. The paired head and cup AVN failure components were retrieved after 54 months and showed a combined wear rate of just 0.51 mm³/year. The infection retrieval was slightly above the values predicted by simulator results, with a combined wear rate of 3.98 mm³/year. However, combined wear rates of 17.64 and 68.50 mm³/year in the ARMD and ARMD fracture groups respectively are significantly higher than those expected in well functioning MoM hip resurfacings. From an examination of these combined wear rates it can be speculated that the clinical implication of the greatest wear rates is bone fracture. In patients the other effects of high wear manifested themselves as pain, effusions, tissue necrosis, difficulty in moving and early failure necessitating another major operation [33].

Typical in vivo wear rates for MoP hip prostheses are of the order of 35-62 mm³/year [42]. While the wear rates reported in this paper for the majority of MoM hip resurfacings are less than for MoP hip prostheses it should be recognised that another vital factor is size of the wear debris and the number of particles. Mode polyethylene wear particles have been offered in the size range 0.5-1 μm [43-45], though particles up to 10 μm in size are common [17,44]. Typical metal wear particles have been found to be of the order of 40 nm in size [46,47]. Taking these sizes and values of 62 mm³/year for MoP and 18 mm³/year for MoM hip resurfacings means that a failed hip resurfacing joint will subject the patient to a number of particles 4-70 times greater than seen in an equivalent MoM hip. This is in the range offered by a 1998 study which suggested that 13-500 times more metal particles were produced in a MoM THR than PE particles in an equivalent MoP prosthesis [48].

Turning to the few other ex vivo studies of MoM hip resurfacings, Morlock et al. also found high wear rates for certain sub-groups of components, with mean wear rates for 14 rim-loaded heads and
Table 5
Retrieval and wear data for all 54 components. Combined wear data is provided where both head and cup components were retrieved. Paired components are numbered consecutively. The ten single heads are numbered Head 51–Head 510.

<table>
<thead>
<tr>
<th>Device</th>
<th>Nominal diameter (mm)</th>
<th>Time to retrieval (months)</th>
<th>Reason for failure</th>
<th>Wear volume (mm³)</th>
<th>Wear rate (mm³/year)</th>
<th>Combined volume (mm³)</th>
<th>Combined wear rate (mm³/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head 1</td>
<td>43</td>
<td>8</td>
<td>ARMD</td>
<td>19.65</td>
<td>29.48</td>
<td>25.01</td>
<td>37.52</td>
</tr>
<tr>
<td>Cup 1</td>
<td>43</td>
<td>8</td>
<td>ARMD</td>
<td>5.36</td>
<td>8.04</td>
<td>25.01</td>
<td>37.52</td>
</tr>
<tr>
<td>Head 2</td>
<td>47</td>
<td>14</td>
<td>ARMD</td>
<td>6.08</td>
<td>5.21</td>
<td>10.17</td>
<td>8.72</td>
</tr>
<tr>
<td>Cup 2</td>
<td>47</td>
<td>14</td>
<td>ARMD</td>
<td>4.09</td>
<td>3.51</td>
<td>10.17</td>
<td>8.72</td>
</tr>
<tr>
<td>Head 3</td>
<td>47</td>
<td>17</td>
<td>ARMD</td>
<td>2.96</td>
<td>2.09</td>
<td>5.35</td>
<td>3.78</td>
</tr>
<tr>
<td>Cup 3</td>
<td>47</td>
<td>17</td>
<td>ARMD</td>
<td>2.39</td>
<td>1.69</td>
<td>5.35</td>
<td>3.78</td>
</tr>
<tr>
<td>Head 4</td>
<td>43</td>
<td>18</td>
<td>ARMD</td>
<td>7.93</td>
<td>5.29</td>
<td>10.38</td>
<td>6.92</td>
</tr>
<tr>
<td>Cup 4</td>
<td>43</td>
<td>18</td>
<td>ARMD</td>
<td>2.45</td>
<td>1.63</td>
<td>10.38</td>
<td>6.92</td>
</tr>
<tr>
<td>Head 5</td>
<td>49</td>
<td>19</td>
<td>ARMD</td>
<td>19.79</td>
<td>18.44</td>
<td>138.91</td>
<td>87.73</td>
</tr>
<tr>
<td>Cup 5</td>
<td>49</td>
<td>19</td>
<td>ARMD</td>
<td>19.79</td>
<td>18.44</td>
<td>138.91</td>
<td>87.73</td>
</tr>
<tr>
<td>Head 6</td>
<td>49</td>
<td>21</td>
<td>ARMD</td>
<td>11.17</td>
<td>6.38</td>
<td>49.68</td>
<td>28.39</td>
</tr>
<tr>
<td>Cup 6</td>
<td>49</td>
<td>21</td>
<td>ARMD</td>
<td>38.51</td>
<td>22.01</td>
<td>49.68</td>
<td>28.39</td>
</tr>
<tr>
<td>Head 7</td>
<td>51</td>
<td>22</td>
<td>ARMD</td>
<td>3.68</td>
<td>2.01</td>
<td>7.30</td>
<td>3.98</td>
</tr>
<tr>
<td>Cup 7</td>
<td>51</td>
<td>22</td>
<td>ARMD</td>
<td>3.68</td>
<td>2.01</td>
<td>7.30</td>
<td>3.98</td>
</tr>
<tr>
<td>Head 8</td>
<td>45</td>
<td>22</td>
<td>ARMD</td>
<td>3.78</td>
<td>2.06</td>
<td>5.71</td>
<td>3.11</td>
</tr>
<tr>
<td>Cup 8</td>
<td>45</td>
<td>22</td>
<td>ARMD</td>
<td>1.93</td>
<td>1.05</td>
<td>5.71</td>
<td>3.11</td>
</tr>
<tr>
<td>Head 9</td>
<td>51</td>
<td>27</td>
<td>ARMD</td>
<td>39.78</td>
<td>17.68</td>
<td>49.24</td>
<td>21.88</td>
</tr>
<tr>
<td>Cup 9</td>
<td>51</td>
<td>27</td>
<td>ARMD</td>
<td>39.78</td>
<td>17.68</td>
<td>49.24</td>
<td>21.88</td>
</tr>
<tr>
<td>Head 10</td>
<td>41</td>
<td>27</td>
<td>ARMD</td>
<td>13.31</td>
<td>5.92</td>
<td>46.26</td>
<td>20.56</td>
</tr>
<tr>
<td>Cup 10</td>
<td>41</td>
<td>27</td>
<td>ARMD</td>
<td>13.31</td>
<td>5.92</td>
<td>46.26</td>
<td>20.56</td>
</tr>
<tr>
<td>Head 11</td>
<td>51</td>
<td>30</td>
<td>ARMD</td>
<td>2.93</td>
<td>1.17</td>
<td>4.50</td>
<td>2.58</td>
</tr>
<tr>
<td>Cup 11</td>
<td>51</td>
<td>30</td>
<td>ARMD</td>
<td>2.93</td>
<td>1.17</td>
<td>4.50</td>
<td>2.58</td>
</tr>
<tr>
<td>Head 12</td>
<td>46</td>
<td>36</td>
<td>ARMD</td>
<td>5.36</td>
<td>1.79</td>
<td>15.81</td>
<td>5.27</td>
</tr>
<tr>
<td>Cup 12</td>
<td>46</td>
<td>36</td>
<td>ARMD</td>
<td>5.36</td>
<td>1.79</td>
<td>15.81</td>
<td>5.27</td>
</tr>
<tr>
<td>Head 13</td>
<td>49</td>
<td>39</td>
<td>ARMD</td>
<td>75.09</td>
<td>23.10</td>
<td>131.16</td>
<td>40.36</td>
</tr>
<tr>
<td>Cup 13</td>
<td>49</td>
<td>39</td>
<td>ARMD</td>
<td>75.09</td>
<td>23.10</td>
<td>131.16</td>
<td>40.36</td>
</tr>
<tr>
<td>Head 14</td>
<td>41</td>
<td>45</td>
<td>ARMD</td>
<td>4.38</td>
<td>1.17</td>
<td>7.30</td>
<td>3.98</td>
</tr>
<tr>
<td>Cup 14</td>
<td>41</td>
<td>45</td>
<td>ARMD</td>
<td>4.38</td>
<td>1.17</td>
<td>7.30</td>
<td>3.98</td>
</tr>
<tr>
<td>Head 15</td>
<td>41</td>
<td>45</td>
<td>ARMD</td>
<td>8.92</td>
<td>2.38</td>
<td>13.30</td>
<td>5.55</td>
</tr>
<tr>
<td>Cup 15</td>
<td>41</td>
<td>45</td>
<td>ARMD</td>
<td>8.92</td>
<td>2.38</td>
<td>13.30</td>
<td>5.55</td>
</tr>
<tr>
<td>Head 16</td>
<td>45</td>
<td>52</td>
<td>ARMD fracture</td>
<td>13.77</td>
<td>3.24</td>
<td>20.19</td>
<td>4.75</td>
</tr>
<tr>
<td>Cup 16</td>
<td>45</td>
<td>52</td>
<td>ARMD fracture</td>
<td>13.77</td>
<td>3.24</td>
<td>20.19</td>
<td>4.75</td>
</tr>
<tr>
<td>Head 17</td>
<td>46</td>
<td>52</td>
<td>ARMD fracture</td>
<td>13.77</td>
<td>3.24</td>
<td>20.19</td>
<td>4.75</td>
</tr>
<tr>
<td>Cup 17</td>
<td>46</td>
<td>52</td>
<td>ARMD fracture</td>
<td>13.77</td>
<td>3.24</td>
<td>20.19</td>
<td>4.75</td>
</tr>
<tr>
<td>Head 18</td>
<td>49</td>
<td>58</td>
<td>ARMD</td>
<td>20.71</td>
<td>4.28</td>
<td>54.28</td>
<td>11.23</td>
</tr>
<tr>
<td>Cup 18</td>
<td>49</td>
<td>58</td>
<td>ARMD</td>
<td>20.71</td>
<td>4.28</td>
<td>54.28</td>
<td>11.23</td>
</tr>
<tr>
<td>Head 19</td>
<td>47</td>
<td>53</td>
<td>ARMD fracture</td>
<td>28.19</td>
<td>6.51</td>
<td>44.02</td>
<td>10.16</td>
</tr>
<tr>
<td>Cup 19</td>
<td>47</td>
<td>53</td>
<td>ARMD fracture</td>
<td>28.19</td>
<td>6.51</td>
<td>44.02</td>
<td>10.16</td>
</tr>
<tr>
<td>Head 20</td>
<td>47</td>
<td>53</td>
<td>ARMD fracture</td>
<td>31.91</td>
<td>136.1</td>
<td>366.10</td>
<td>95.50</td>
</tr>
<tr>
<td>Cup 20</td>
<td>47</td>
<td>53</td>
<td>ARMD fracture</td>
<td>31.91</td>
<td>136.1</td>
<td>366.10</td>
<td>95.50</td>
</tr>
<tr>
<td>Head 21</td>
<td>51</td>
<td>51</td>
<td>ARMD fracture</td>
<td>43.22</td>
<td>71.78</td>
<td>176.32</td>
<td>41.49</td>
</tr>
<tr>
<td>Cup 21</td>
<td>51</td>
<td>51</td>
<td>ARMD fracture</td>
<td>43.22</td>
<td>71.78</td>
<td>176.32</td>
<td>41.49</td>
</tr>
<tr>
<td>Head 22</td>
<td>53</td>
<td>53</td>
<td>ARMD fracture</td>
<td>53.35</td>
<td>7.70</td>
<td>176.32</td>
<td>41.49</td>
</tr>
<tr>
<td>Cup 22</td>
<td>53</td>
<td>53</td>
<td>ARMD fracture</td>
<td>53.35</td>
<td>7.70</td>
<td>176.32</td>
<td>41.49</td>
</tr>
</tbody>
</table>

Table 6
Wear rates and volumes at retrieval for all components. Data is grouped by failure mode.

<table>
<thead>
<tr>
<th>Failure mode (number of cups, heads)</th>
<th>Mean wear rate (mm³/year) (range)</th>
<th>Mean wear volume (mm³) (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cups</td>
<td>Heads</td>
<td>Combined</td>
</tr>
<tr>
<td>ARM 1(19)</td>
<td>9.26 (1.05–69.30)</td>
<td>17.64 (3.11–87.73)</td>
</tr>
<tr>
<td>ARM fracture 2(26)</td>
<td>36.75 (9.91–63.59)</td>
<td>68.50 (41.49–95.50)</td>
</tr>
<tr>
<td>Early fracture 0(4)</td>
<td>8.26 (3.51–17.10)</td>
<td>2.42 (0.72–4.06)</td>
</tr>
<tr>
<td>AVN 1(12)</td>
<td>0.30</td>
<td>0.93</td>
</tr>
<tr>
<td>Infection 1(1.1)</td>
<td>1.89</td>
<td>3.98</td>
</tr>
</tbody>
</table>

Cups: 19.30 (1.93–109.72) 21.92 (2.96–75.09) 38.69 (5.35–138.91)
Heads: 142.94 (42.10–243.77) 56.49 (7.24–134.22)
Combined: 271.21 (176.32–366.10)
15 rim-loaded cups of 8.69 and 15.88 mm³/year respectively [15]. For non-rim loaded components the mean wear rate dropped to 0.40 mm³/year (12 heads) and 0.58 mm³/year (17 cups). Morlock et al. did not differentiate between implants from different manufacturers. However, hip resurfacing design does have an impact: recent survival data has shown higher 5-year failure rates for ASR™ (12.0%) than other resurfacing designs such as the Birmingham Hip Resurfacing (BHR™, Smith and Nephew, Warwick, United Kingdom) (4.3%) or the Adept (Finsbury Orthopaedics Ltd., Leatherhead, UK) (5.0%) [1]. Witzleb et al. have reported on ten BHR™ explanted components (8 heads, 2 cups) which exhibited wear rates as high as 22.08 mm³/year, though with a mean of 3.36 mm³/year [32]. This mean wear rate is significantly lower than the rate for ASR™ reported in this paper and this goes some way to explaining the differing survival rates offered by the National Joint Registry. Components in Witzleb’s study were retrieved for femoral neck fracture (4 heads), femoral subsidence (2 heads), aseptic cup loosening (1 head), and infection (1 head, 1 cup). The heaviest cup wear was seen in the case of infection (31.5 mm³) after 15 months in vivo while the heaviest head wear occurred on a femoral neck fracture implanted at 70° abduction (17.8 mm³ after 24 months in vivo). In comparison, ASR™ cups in the present study exhibited a mean wear volume at retrieval of 24.97 mm³ with maximum wear on an individual component of 229.00 mm³ after 46 months in vivo (retrieved after ARMD fracture). ASR™ heads in the present study exhibited a mean wear volume at retrieval of 25.44 mm³ with maximum wear on an individual component of 134.22 mm³ after 51 months in vivo (retrieved after ARMD fracture). This data is summarised in Table 7. The present study shows that the heaviest wear is seen on late fracture components, in agreement with Witzleb et al.

What might explain the difference in wear rates between many ex vivo results and in vitro hip simulator studies? It is possible that hip simulators do not currently apply a physiologically realistic test to MoM resurfacing joints. Kamali et al. have suggested that a stop/start motion, a change in frequency from 1 Hz to 0.5 Hz, and alternating kinetic and kinematic profiles would provide a more physiologically relevant test protocol [49]. Additionally, in the past it has been rare for components to be tested in vitro at ‘extreme’ angles which are seen in vivo. One simulator study of six 40 mm articulating diameter MoM devices [25] demonstrated a 47-fold increase in steady-state volumetric wear rate from 0.24 to 1.7 mm³/million cycles when cup inclination angle was increased from 35° to 60°. In another simulator study on 39 mm diameter hip resurfacing components, tests were conducted with the cup at 45° inclination, and then a second test was done with the cup at 55° inclination with the addition of microlateralization [35]. At 45° inclination an overall wear rate of 1.61 mm³/million cycles was reported, compared with 8.99 mm³/million cycles at 55° inclination plus microlateralization. If a hip resurfacing patient is expected to undertake 1.9 million steps per year [49] then a patient with a malpositioned cup could expect a wear rate of 17.1 mm³/year. This is uncannily close to the mean wear rate of 17.64 mm³/year reported for the ARMD failures in this paper. Thus, while hip simulators can give results which match ex vivo hip resurfacing wear rates, it may be that studies under ideal conditions will fail to reproduce the wear rates which have been reported in this paper.

It has been recognised that accurate component positioning, particularly of the acetabular cup, is key in minimising failure of MoM hip resurfacing devices [33]. A recent Medical Device Alert (MDA) reported higher than anticipated rates of revision for ASR™ acetabular cups and recommended that cups are implanted with inclination angles between 40° and 45° [50]. Whether this tight range can be achieved is open to question [51].

5. Conclusion

This paper has presented a novel method for calculating volumetric wear of ex vivo MoM hip prostheses. It was applied to the largest series of ex vivo ASR™ resurfacing prostheses reported in the scientific literature, though it is equally applicable to other resurfacing designs as well as all other MoM total hip replacements. The method has been shown to be accurate to approximately 0.5 mm³ of volumetric wear across a range of wear volumes and has proved repeatable across multiple scans of the same component.

Measurements of retrieved ASR™ components have shown wear rates significantly greater than is expected in a well functioning MoM resurfacing hip prosthesis. This is of real concern for patients implanted with the device. In late 2010 in the UK a MDA was issued preventing further implantation of ASR™ hip replacements [52]. On 26 August 2010 the ASR™ was withdrawn. However, some 93,000 ASR™ devices have been implanted worldwide between 2003 and 2010. Vital conclusions can be drawn from studying failed hip prostheses so that risks to patients can be minimised in future designs.

References

Chapter Seven
Reducing Metal Ion Release Following Hip Resurfacing Arthroplasty

David J. Langton, MRCS\textsuperscript{a,b,c,*}, Thomas J. Joyce, PhD\textsuperscript{c}, Navjeet Mangat, MRCS\textsuperscript{b,d}, James Lord, MEng\textsuperscript{a,c}, Maarten Van Orsouw, MD\textsuperscript{a,e}, Koen A. De Smet, MD\textsuperscript{a,e}, Antoni V.F. Nargol, FRCS (Tr & Orth)\textsuperscript{a,b}

KEYWORDS

\textbullet{} Metal ion release \textbullet{} Resurfacing arthroplasty \textbullet{} Implant \textbullet{} Ion concentration

Increased chromium (Cr) and cobalt (Co) concentrations following metal-on-metal (MoM) hip arthroplasty are associated with local and systemic pathologic changes.\textsuperscript{1-6} It is unquestionable that efforts should be made to identify modifiable variables leading to ion release to minimize the risk of adverse effects.

To date, the variables that have been shown to significantly affect metal ion concentrations in the absence of renal disease are femoral component diameter,\textsuperscript{7} acetabular cup angles of inclination and anteversion,\textsuperscript{7-9} time from surgery,\textsuperscript{10} and activity.\textsuperscript{11} Of these variables, only cup orientation can be regarded as realistically under the surgeon's control. However, it remains to be proven whether an optimal cup position exists and whether this is affected by implant design. Our previous work and those of others have suggested that the coverage angle provided by the acetabular cup is critical.\textsuperscript{7,12} There is evidence, concordant with laboratory data, to suggest that clearance may also be an important variable in determining in vivo wear rates.\textsuperscript{13}

There were 3 aims of this study:
1. To investigate the relationship between volumetric wear rate and serum metal ion concentrations
2. To establish the incidence of excessive metal ion release following resurfacing with commonly used devices
3. To identify cup orientations associated with lowest ion concentrations for each device and to propose mechanisms leading to increased wear.

PATIENTS AND METHODS

There were 723 patients in this series, 446 men and 277 women. All surgeries were performed by

Disclosure: David Langton has received payment for work as a consultant to DePuy and Wright Medical; Koen De Smet is a consultant for Wright Medical.
The remaining authors have nothing to disclose.

None of the authors' families have any commercial or financial interest in the material in the study.
\textsuperscript{a} Joint Replacement Unit, University Hospital of North Tees, Hardwick, Stockton-on-Tees, TS19 8PE, UK
\textsuperscript{b} North Tees and Hartlepool NHS Trust, University Hospital of North Tees, Hardwick, Stockton-on-Tees, TS19 8PE, UK
\textsuperscript{c} School of Mechanical and Systems Engineering, Newcastle University, Stephenson Building, Claremont Road, NE1 7RU, England, UK
\textsuperscript{d} Northern Deanery, North East Strategic Health Authority, Waterfront 4, Goldcrest Way, Newburn Riverside, Newcastle Upon Tyne, NE15 8NY, UK
\textsuperscript{e} ANCA Medical Centre, Krijgslaan 181, 9000, Ghent, Belgium
\textsuperscript{*} Corresponding author. 199 Rodsley Avenue, Gateshead, NE8 4LB, UK

E-mail address: djlangton22@doctors.org.uk

Orthop Clin N Am 42 (2011) 169-180
0030-5898/11 - see front matter © 2011 Elsevier Inc. All rights reserved.
2 experienced hip resurfacing surgeons who use metal ion analysis as part of routine follow-up. Both surgeons perform more than 100 hip resurfacings per year. Surgeon 1 is based in the United Kingdom (at site 1) and surgeon 2 is based in Belgium (at site 2). At both sites, blood samples were collected from the patients at a minimum time of 12 months after surgery to avoid the confounding factor of higher levels of wear during the run-in period. These patients did not have any other metal implants at the time of measurement. Three resurfacing devices were studied: the Articular Surface Replacement (ASR; DePuy International, Leeds, UK), the Birmingham Hip Resurfacing (BHR; Smith and Nephew, Warwick, UK), and the Conserve® Plus (C+; Wright Medical Technology, Memphis, TN, USA). Serum ion analysis was performed using inductively coupled plasma mass spectrometry. Volumetric wear rates of explanted components determined by coordinate measuring machine analysis were compared with corresponding serum ion concentrations in vivo. The relationships between ion levels and component size and cup orientations remaining in vivo were investigated. The mean (range) age of patients from site 1 was 56 years (25–83 years) and the mean (range) time from surgery to blood sampling was 29 months (12–72 months) for the patients with the ASR implant device and 66 months (48–81 months) for those with the BHR implant device. The mean age of patients from site 2 was 51 years and the mean (range) time from surgery to blood sampling was 19 months (12–29 months) for the patients with the ASR implant device, 51 months (12–106 months) for those with the BHR implant device, and 28 months (12–54 months) for those with the C+ implant device.

### Implants

From site 1, there were 223 patients with ASR and 72 with BHR implants. From site 2 there were 271 patients with BHR, 136 with C+, and 21 with ASR implants. **Table 1** shows the differences in design between the implants in this study.

### Sample Collection

The samples were obtained at site 2 using an intravenous catheter (insyte-WTM; Becton Dickinson, Franklin Lakes, NJ, USA). After the catheter had been introduced, the metal needle was withdrawn and the first 5 mL of blood was discarded to avoid possible contamination from the needle. A second 5 mL was collected using a vacuum tube (Venosafe VF-106SAHL; Terumo Europe NV, Leuven, Belgium). The samples were analyzed at the Laboratory of Clinical Biology at Ghent University Hospital, Ghent, Belgium. The laboratory quotes its quantification limit as 0.5 μg/L with a reproducibility of 5%. At site 1, serum samples were collected in a similar manner. Venous cannulation was performed with a 21-gauge stainless steel needle (Venflon, Becton Dickinson, Helsingborg, Sweden), with disposal of the first 5 mL of blood to avoid contamination. All samples were centrifuged to separate blood and serum fractions, frozen, and sent for blinded trace element analysis at the Trace Element Laboratory of the Royal Surrey County Hospital, Guildford, United Kingdom. This laboratory also quotes its quantification limit as 0.5 μg/L with a reproducibility of 5%.

### Radiographic Analysis

At the time of collection of the sample, the University of California, Los Angeles activity scores were

---

### Table 1
**Comparison of device characteristics**

<table>
<thead>
<tr>
<th>Device Characteristics</th>
<th>ASR</th>
<th>BHR</th>
<th>C+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtended articular surface angle (°)</td>
<td>144–160°</td>
<td>158–165°</td>
<td>162–165°</td>
</tr>
<tr>
<td>Mean nominal radial clearance (μm)</td>
<td>50</td>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>Manufacturing method of head</td>
<td>As cast</td>
<td>As cast</td>
<td>HIP/SA</td>
</tr>
<tr>
<td>Manufacturing method and treatment of cup</td>
<td>HIP/SA</td>
<td>As cast</td>
<td>HIP/SA</td>
</tr>
<tr>
<td>Carbon content</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

* Subtended articular surface angles increase with increasing cup diameter.

b HIP/SA, cast process and heat treatment by hot isostatic pressure/surface annealed.

c High carbon content defined as ≥0.20%.

recorded and weight-bearing pelvic radiographs were obtained. From these radiographs, the following parameters were measured using ImageJ software (National Institutes of Health, Bethesda, MD, USA): femoral stem angle relative to the femoral shaft (SSA), femoral stem to femoral neck angle (SNA), and femoral offset and femoral component to femoral neck ratios. Einzel-Bild-Roentgen-Analyse (EBRA, University of Innsbruck, Innsbruck, Austria) software was used to analyze all available radiographs to obtain angles of cup inclination and anteversion. The accuracy of this software has been discussed in the literature. In most cases, one well-centered radiograph was available for analysis. Theoretical contact patch to rim (CPR) distance was calculated for all patients.

To test the validity of comparing Cr and Co levels between the 2 populations, patients with BHR implants from site 2 were matched with patients with BHR implants from site 1 by the length of time to blood sampling from surgery. From these patients, we selected only those with femoral components of size 50 mm and larger as we have previously showed these sizes to be relatively resistant to cup position. To further control for the effects of cup position, cups with inclinations greater than 55° and less than 10° or anteversion greater than 30° were excluded. The Cr and Co levels for the matched populations were then compared using Mann-Whitney tests for nonparametric data (Table 2). P values less than .05 were deemed significant. Two sample t-tests comparing the relevant parameters can be seen in Table 2. Cr concentrations in patients from UK were found to be slightly but significantly increased when compared with patients from Belgium. There was no significant difference between the groups with regard to Co. For this reason, we used only serum Co in the analysis.

For the 3 resurfacing devices, the relationship between each variable and serum Co was examined using the Spearman rank correlation. Rank correlation was preferred because of the nonparametric nature of the ion data. Scatter plots were used to identify nonlinear relationships. To examine the relative sensitivity of each device to the effects of cup orientation, the patients from each implant group were divided into subgroups of inclination (<25°, 25°–27°, 27°–29°, 29°–31°, 31°–33°, 33°–35°) and anteversion (<0°, 0°–5°, 5°–7°, 7°–9°, 9°–11°). From each subgroup the median Co level was calculated and plotted. To further examine this relationship, patients were placed into zones of cup orientation based on the EBRA measurements (Fig. 1). We then calculated the percentage of patients in each zone (for each device) with serum Co levels more than 7 µg/L. This figure has recently been quoted by the Medicines and Healthcare Products Regulatory Agency/Metal on Metal Working Committee of the United Kingdom, indicating a poorly performing bearing surface with recommendations for close patient follow-up. Windows SPSS version 15.0 (SPSS Inc, Chicago, IL, USA) was used for statistical analysis throughout.

**Survival Analysis Using Metal Ion Concentrations**

In our previous work we showed that there was a suggestion that there is a temporary immunity to increases in the wear of malpositioned cups/cups. With this principle in mind, we collected all the metal ion results from all patients in the study at all periods, including repeat samples from individual patients at different times postoperatively. Of the 723 patients, 124 had undergone repeat testing (as part of routine screening), giving a total of 847 samples. Using these data, we conducted a Kaplan-Meier survival analysis comparison between the patients in each device group, with the latest follow-up being the time in months.

**Table 2**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Site 1</th>
<th>Site 2</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>40</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Time to Sample (mo)</td>
<td>66 (48–81)</td>
<td>64 (50–81)</td>
<td>.190</td>
</tr>
<tr>
<td>Femoral Size (mm)</td>
<td>51.6</td>
<td>52.4</td>
<td>.107</td>
</tr>
<tr>
<td>Cup Inclination (°)</td>
<td>48.5 (33–55)</td>
<td>48.8 (34–55)</td>
<td>.890</td>
</tr>
<tr>
<td>Cup Anteversion (°)</td>
<td>17.04 (10–39)</td>
<td>20.5 (10–36)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Serum Cr (µg/L)*</td>
<td>3.49</td>
<td>1.80</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Serum Co (µg/L)*</td>
<td>1.31</td>
<td>1.20</td>
<td>.812</td>
</tr>
</tbody>
</table>

* Median values, all other values are means (range).
Langton et

to the description byWaiter and colleagues. Described in the Appendix. Univariate metric wear was categorized as anteverted or retroverted according to the radius of the explant. As each component, dependent on the radius of the acetabular component, which showed larger joints to be less vulnerable to the effects of cup position. The patients were then also categorized by cup position and a Cox proportional hazard model constructed.

EXPLANT ANALYSIS

The wear of retrieved ASR matched femoral and acetabular components from site 1 (n = 35 pairs) was measured by a coordinate measuring machine (CMM) using a scanning head (Legex 322, Mitutoyo Halifax, UK) with a spatial resolution of less than 1 µm in the area of measurement. Measurements were made for every 5° on 18 concentric circles as well as at the pole of the component, which resulted in a total of 4500 to 6000 measurements for each component, dependent on the radius of the explant. The wear scars of each component were categorized as anteverted or retroverted according to the description by Walter and colleagues. Volumetric wear was calculated using the method described in the Appendix. Univariate linear regression was used to examine the relationship between volumetric wear rates and serum Co and Cr levels. Using the Medicines and Healthcare products Regulatory Agency (MHRA) level of 7 µg/L as a positive test and abnormal bearing surface wear as an annual wear rate greater than 3 mm³/yr, (In our recent work, 3 mm³/yr was the lowest bearing surface wear rate of a resurfacing device associated with adverse reaction to metal debris.) the sensitivity and specificity of serum Cr and Co testing was evaluated.

RESULTS A: EX VIVO

Explant Analysis and Serum Ion Concentrations

Volumetric wear correlated well with serum Cr (r² = 0.66, P<.01) and serum Co concentrations (r² = 0.78, P<.01). Specificity and sensitivity (95% CI) of serum Co levels greater than 7 µg/L were found to be 1.00 (0.67–1.00) and 0.93 (0.83–0.98), respectively. Specificity and sensitivity (95% CI) of serum Cr levels greater than 7 µg/L were found to be 0.80 (0.48–0.97) and 0.93 (0.83–0.97), respectively. Wear maps of retrieved components were generated and can be seen adjacent to the corresponding radiographs (Fig. 2). In Fig. 2, for simplification, red areas represent wear of greater than 20 µm deviation from a perfect spherical form. All explants that were found to have extremely high wear rates were found to have edge loading of the acetabular components. We define edge loading here as a progressive increase in the measured wear depths of the cups so that maximal wear depths occur at the edge of the articular surface.

RESULTS B: IN VIVO

Femoral Size

In all 3 devices, we identified an inverse relationship between femoral component size and serum Co concentrations (Table 3). This relationship was not significant in the C+ device. However, the largest values were found in the patients in whom larger-diameter bearings had been suboptimally positioned.

Cup Inclination

Increasing serum Co concentrations were associated with increasing inclination angle of the ASR cup (Fig. 3; see Table 3). In the patients with BHR and C+ implant devices, lowest ion levels were associated with cups placed with inclinations of 45° to 50°. These results were consistent through the size ranges.
Reducing Metal Ion Release

Fig. 2. (A) Retrieved components of a 45-year-old female patient with bilateral ASR resurfacings. Right ASR failed at 19 months, the left failed at 58 months. Both failures were secondary to pain, large joint effusions, and areas of tissue necrosis. The right cup was placed with inclination/anteversion angles of 57°/27°, and the corresponding femoral component shows signs of anterior rim loading with possible anterior subluxation. The left cup was placed with inclination/anteversion angles of 41°/0°. The wear scar is directed in an opposite direction to the right, indicating posterior rim loading and subluxation. The time to onset of pain in the left hip was much longer compared with the right and this potentially reflects the less steep increase in metal ion levels associated with low cup angles as opposed to high cup angles. (B) Femoral component of a 49-year-old male patient (ASR) with femoral fracture at 4 years associated with metallosis/aseptic lymphocyte dominated vasculitis associated lesion. The main wear scar is retroverted suggesting that the mechanism of wear was posterior edge loading/subluxation during deep flexion and internal rotation. The cup is orientated in 40° of inclination and 8° of anteversion (radiological). (C) Retrieved components from a 50 year old male who suffered a femoral neck fracture at 7 months. The cup was placed in 41° of inclination and 20° of anteversion. Patterns of wear suggest possible equatorial bearing secondary to low clearance.
Table 3
The Spearman rank correlations between each examined variable and serum Co concentration for each device

<table>
<thead>
<tr>
<th>Variables</th>
<th>ASR</th>
<th>BHR</th>
<th>C+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femoral Size</td>
<td>-0.22 (.01)</td>
<td>-0.21 (&lt;.01)</td>
<td>-0.12 (.16)</td>
</tr>
<tr>
<td>Cup Inclination</td>
<td>0.41 (&lt;.01)</td>
<td>0.09 (.07)</td>
<td>-0.11 (.20)</td>
</tr>
<tr>
<td>Cup Anteversion</td>
<td>0.15 (.130)</td>
<td>0.13 (.03)</td>
<td>0.20 (.021)</td>
</tr>
<tr>
<td>Time from Surgery</td>
<td>0.17 (.05)</td>
<td>0.07 (.50)</td>
<td>-0.07 (.432)</td>
</tr>
<tr>
<td>SSA</td>
<td>0.06 (.53)</td>
<td>0.18 (.05)</td>
<td>0.18 (.03)</td>
</tr>
<tr>
<td>SNA</td>
<td>-0.07 (.57)</td>
<td>0.18 (.190)</td>
<td>0.01 (.90)</td>
</tr>
<tr>
<td>Stem Neck</td>
<td>-0.14 (.17)</td>
<td>0.04 (.457)</td>
<td>0.25 (.05)</td>
</tr>
<tr>
<td>Offset</td>
<td>0.01 (.98)</td>
<td>-0.23 (.004)</td>
<td>-0.01 (.93)</td>
</tr>
<tr>
<td>UCLA Score</td>
<td>-0.29 (.09)</td>
<td>0.010 (.85)</td>
<td>-0.05 (.54)</td>
</tr>
<tr>
<td>Head Neck Ratio</td>
<td>0.47 (.01)</td>
<td>0.22 (&lt;.01)</td>
<td>0.12 (.17)</td>
</tr>
<tr>
<td>CPR</td>
<td>-0.49 (&lt;.01)</td>
<td>-0.255 (&lt;.01)</td>
<td>-0.09 (.28)</td>
</tr>
<tr>
<td>Age</td>
<td>0.02 (.82)</td>
<td>0.06 (.65)</td>
<td>0.10 (.25)</td>
</tr>
</tbody>
</table>

All patients in the study are included. Note that there is a strong positive correlation among the patient group as a whole between femoral size and offset/SSA/SNA. When femoral size is controlled for these variables nonsignificant relationships to Co levels is shown. P values are given in parentheses.

Abbreviations: SSA, femoral stem angle relative to the femoral shaft, SNA, femoral stem to femoral neck angle.

**Cup Anteversion**

The ASR proved to be extremely sensitive to both increased and decreased anteversion (Fig. 4). The median Co concentration of patients with BHR devices with cups placed with anteversion greater than 30° was greatly increased, although this relationship was absent in patients with BHR devices with femoral implant sizes larger than 50 mm. The C+ device proved to be relatively resistant to the effects of anteversion; however, there were only a limited number of patients (n = 4) with a cup anteversion greater than 30°.

**The Interaction of Bearing Diameter and Cup Orientation**

Large joints of each device were compared with small joints of each device in cup zone 3, 4, and 5. These zones were selected because cups placed in these positions are not generally viewed as optimally placed, but they are commonly encountered.

![Fig. 3. Median serum Co values associated with each subgroup of inclination. All patients in the study are included. Red, ASR patients; blue, BHR patients; green, C+ patients. Median Co level associated with ASR cup inclinations greater than 60° is removed for graphical representation (21.5 μg/L).](image)

![Fig. 4. Median serum Co values associated with each subgroup of anteversion. All patients in the study are included. Red, ASR patients; blue, BHR patients; green, C+ patients.](image)
in surgical practice. In all devices, the median Co values were found to be lower in the large joint groups, 3.71 versus 2.38 μg/L ($P = .08$) for the ASR, 2.50 versus 1.60 μg/L ($P < .01$) for the BHR, and 3.00 versus 1.35 μg/L ($P = .04$) for the C+ groups.

**Design Factors Other than Bearing Diameter and Cup Orientation**

Site 2 patients with the BHR device were matched with those with the C+ device according to the duration of time postoperation to blood sampling. These 2 devices were selected for comparison because they have the largest arcs of cover, opposing heat treatments and differing clearances (see Table 1). From these patients, we selected only those with femoral components of size 50 mm and above for reasons outlined earlier. To further control for the effects of cup position, cups with extreme inclinations (>55°) and/or anteversions (<10° or >30°) were excluded. The Cr and Co levels for the matched populations were then compared using Mann-Whitney tests for nonparametric data. Two sample $t$-tests comparing the relevant parameters of the 2 groups can be seen in Table 4.

**Incidence of Serum Co Levels Greater than 7 μg/L** Between Devices by Cup Placement Zone

Fig. 5 shows the percentage of patients with serum Co concentrations greater than 7 μg/L related to the zone of cup placement. In zone 1, serum Co concentrations in 10% of the patients with the ASR device were above this level. The C+ device was most resistant to variations in cup placement, consistent with the results described earlier. However, blood samples in patients with the BHR device were taken at a significantly greater period postoperatively compared with the patients with the C+ device, hence our reason to conduct the survival analysis described later.

**RESULTS C: SURVIVAL ANALYSIS**

Direct comparison of the 3 devices with all patients and all results included showed a significant increase in the failure rate of the ASR device compared with the BHR and C+ devices (Fig. 6). There was no significant difference in survival between the groups with C+ and BHR devices. When patients with the large-joint implants were examined separately, there were no significant differences in the survival of each device and cup orientation was not found to be a significant variable. When only patients with the small-joint implants were included in the Cox hazards model, the ASR device was found to have a hazard ratio of 7.58 compared with the BHR device ($P < .01$). There was a sequential increase in the hazard ratio with cups placed further from zone 1 (all $P$ values < .05).

**DISCUSSION**

Serum metal ion concentrations have been shown to correlate well with maximum wear depths of explanted femoral resurfacing prostheses. However, linear wear depths can be unrepresentative of total volumetric loss of material. In this study, we have demonstrated for the first time a strong correlation between serum ion concentrations and total volumetric wear rates. Recently, there has been a report of multiple end organ damage secondary to Co intoxication after a conventional total hip replacement. It has also been shown that revision surgery secondary to adverse reactions to metal debris frequently lead to unsatisfactory outcomes due to widespread tissue destruction. It could be argued that reducing articulating surface wear is the most

<table>
<thead>
<tr>
<th>Joint Characteristics</th>
<th>BHR</th>
<th>C+</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>21</td>
<td>35</td>
<td>—</td>
</tr>
<tr>
<td>Time to Sample (mo)</td>
<td>34 (12–57)</td>
<td>28 (12–54)</td>
<td>.207</td>
</tr>
<tr>
<td>Size (mm)</td>
<td>52.2</td>
<td>51.5</td>
<td>.240</td>
</tr>
<tr>
<td>Cup Inclination (°)</td>
<td>50.7 (41–55)</td>
<td>50.3 (33–55)</td>
<td>.771</td>
</tr>
<tr>
<td>Cup Anteversion (°)</td>
<td>16.8 (10–19)</td>
<td>16.0 (10–19)</td>
<td>.326</td>
</tr>
<tr>
<td>Serum Cr (μg/L)*</td>
<td>2.00</td>
<td>1.60</td>
<td>.13</td>
</tr>
<tr>
<td>Serum Co (μg/L)*</td>
<td>1.80</td>
<td>1.10</td>
<td>.02</td>
</tr>
</tbody>
</table>

Fourthes are means (and range) unless denoted by "*" which indicate median values.
important consideration when a resurfacing device is implanted, given the proven and theoretical risks of local and systemic exposure to metal debris. The Medicines and Healthcare Products Regulatory Agency of the United Kingdom have recommended close follow-up of patients with Co or Cr concentrations greater than 7 μg/L. The aim of this study was to gain further understanding of the risk of excessive metallic debris exposure after implantation of commonly used resurfacing devices by studying the largest independent metal ion collection in existence. We also hoped to provide evidence of the mechanisms leading to accelerated wear and suggest explanations as to why some devices and bearing sizes may be more susceptible to these processes.

Large diameter MoM joints are designed to harness a beneficial lubricating film to reduce wear. Components are manufactured with such precision that, in theory, the moving joint should develop a fluid film of sufficient thickness for the joint to function in fluid film lubrication (λ > 3). It is recognized that harnessing of this lubricating film can be encouraged by increasing the femoral component size and decreasing the clearance between the femoral and acetabular components. Factors thought to impair fluid film generation include rim contact and component subluxation. If the prosthesis is starved of a lubricating film, wear will increase dramatically over the large bearing surfaces. Evidence of this tendency to boundary lubrication in poorly performing prostheses can be found in the fact that the ion values considered statistical outliers in this data are found in the patients implanted with larger-diameter resurfacings.

Leslie and colleagues showed in a hip simulator study that increased cup inclination angles (>55°) increased the rate of wear of resurfacing components secondary to edge loading. There is an expanding body of clinical data that is consistent with these results. Leslie and colleagues also found that wear rates can be dramatically increased when edge loading is combined with microseparation of the femoral and acetabular components. This can best be achieved in clinical practice by combining high anteversion angles with high inclination angles. The joint is then vulnerable to posterior impingement of the femoral neck on the cup rim during extension, in some extreme cases during the gait cycle. This mechanism in itself can disrupt the fluid film but can also leverage the head anteriorly creating stripe wear on the femoral component and rim damage/edge wear on the cup (see Fig. 2A). In this way, the mechanical situation Leslie and colleagues showed to increase metallic wear 17-fold in vitro is potentially replicated in vivo.

Fig. 5. Cup orientation zone and risk of serum Co concentrations greater than 7 μg/L. Red, ASR patients; blue, BHR patients; green, C+ patients.

Fig. 6. Kaplan Meier survival analysis conducted using serum Co >7 μg/L as a failed joint.
Reducing Metal Ion Release

The ASR device has a smaller arc of coverage relative to the other devices (see Table 1), which has certain implications. When matched for size and orientation, component contact always occurs closer to the rim of an ASR cup than in the other devices. The 3 cups, ASR, BHR, and C+, also have different rim designs. The ASR introducer inserts adjacent to the articular surface leaving a sharp edge, which may mean that subluxation of the ASR femoral component is particularly detrimental. The recessed nature of the articular rim in comparison with the outer rim also means that the advantage of a subhemispherical cup design in terms of avoiding impingement has been largely negated. It is to be noted that impingement is related to not only the cup design but also the angle of function of the femoral component. In practical terms, the ASR device does not provide a greater ROM than the other devices.

Healthy subjects require an average of 104° of hip flexion for sitting, 112° to rise from sitting, and 125° to stoop to pick up an object from the floor. Hip extension reaches a mean peak of 20° during the normal adult gait cycle. Given this information, range of motion studies involving hip resurfacing models suggest that the necks of resurfaced femurs frequently impinge on the acetabular cup during activities of daily living. However, if impingement alone was the dominant mechanism of ion release, then wider variation is expected in the ion results associated with larger-diameter resurfacings because they have a poor head to neck ratio. This is clearly not the case. That is not to say that impingement is not an important factor, rather that the smaller variation in ion concentrations of larger diameter resurfacings is likely to reflect impingement processes without the compounding effect of severe edge loading. This theory may well explain why inclination angles of the C+ and BHR devices at the higher limit of what is presently considered acceptable (45°–50°) are associated with the lowest median ion levels in this large series. Inclination angles in this range, in these devices, when coupled with reasonable angles of anteversion (10°–20°) confers a greater range of impingement-free flexion and extension as well as provides sufficient acetabular coverage. The most recent anatomic studies of the human pelvis show the acetabular inclination angle to be approximately 55° (relative to the anterior pelvic plane). It seems that once edge loading is eliminated as a factor, the more "physiologically" the acetabulum is reconstructed, the lower the associated metal ion concentrations.

Although there is clear evidence that the ASR cup favors lower inclination angles, there also seems to be a difference between devices in terms of the optimal angle of anteversion (see Fig. 4). The lowest median Co concentrations occur in...
patients with the ASR device with cups antevered between 15° and 20°. In the group with the BHR device, lowest Co levels are found in patients with cups positioned between 10° and 15° of anteverision. The C+ seems to tolerate a wider range of anteverision. This tolerance leads us to speculate that the negative effects of near-neutral anteverision observed with the ASR and the smallest BHR devices are primarily because of the effects of posterior edge loading (± anterior impingement) during deep flexion/stair climbing. The retroverted wear scars identified on retrieved femoral components, which were coupled with cups with low angles of anteverision, provide direct evidence of this effect (see Fig. 2A, B).

We found a highly significant positive correlation between Co and Cr levels and head/neck ratio in the patients with ASR and BHR devices, that is, the greater the head to neck ratio, the higher the observed ion concentrations are likely to be. The head/neck ratio of the largest implants is approximately 1.30 steadily increasing up to 1.50 for the smallest implants (all devices). We consider this ratio to be further evidence that the radius and coverage arc of the bearing surface is more important than the head/neck ratio (ie, tendency to impinge) in the process of metal ion generation. We do not know if this is primarily because of an improved lubrication regime or because of the protection from edge loading. We must also not discount the possibility that a highly localized concentration of metal debris per se causes thinning of the femoral neck and consequently increases the head/neck ratio. However, analysis of serial radiographs does not seem to prove a clear association between neck thinning and metal ion levels at this stage.

Patients who have a well-positioned C+ device have a slightly but significantly lower Co concentration than those with a BHR device when the devices are comparable in terms of time to blood sampling, femoral size, and cup orientation. This may be because of the beneficial effect of or a combination of the reduced clearance of the C+, the larger acetabular coverage, and a difference in the material properties of the components. It does seem to suggest, however, that the heat treatment of the C+ device does not have a significant effect on ion release at this length of follow-up. This information must be taken with the caveat that the designer surgeon of the BHR (Derek McMinn) has shown evidence that a previous double heat-treated device first began to fail at 5 years postsurgery. It was for this reason that we conducted our survival analysis.

The survival analysis presented in this article must be interpreted with caution for obvious reasons. This study was not prospective and the time points at which blood samples were taken/repeated in each patient group were not standardized. Also, although blood tests at each center are performed routinely, significant selection bias cannot be ruled out. For example, patients experiencing pain may have been more likely to attend clinics to give samples. Despite this limitation, we believe that the large number of patients involved in the study go some way to reducing the effects of the factors described earlier. More importantly we do believe that the time from implantation to venesection plays a significant role in the identification of accelerated wear, particularly in poorly positioned or poorly designed implants. In the patient group described in this article, we identified 8 patients with ASR device with theoretical CPR distances of less than 5 mm who all had Co concentrations less than 7 μg/L in blood samples taken within 2 years of surgery. We have previously shown a CPR distance of less than 5 mm to be strongly associated with high wear states. When these blood tests were repeated 2 years later, all but 2 of these patients were found to have Co levels greater than 7 μg/L. It is likely that this time delay is related to:

1. The duration of time that has to elapse before the edge of the developing wear area reaching the edge of the articular surface of the cup
2. The time needed for the bearing surfaces to roughen beyond a critical level.

Either of these 2 situations could lead to the inability of the joint to develop a sufficient lubricating film and therefore an increased tendency toward boundary lubrication.
Of the patients in this study, 13.6% were found to have serum Co concentrations greater than 7 μg/L. Increased concentrations were associated with reduced bearing diameter, suboptimal cup placement, and devices with reduced coverage arcs. In patients who received a BHR device or C+ resurfacing device with a femoral component larger than or equal to 49 mm mated with a cup placed between 30° and 60° of inclination and 0° and 30° of anteversion this percentage decreased to 3.35%.

Resurfacing cups positioned in 45° ± 5° of inclination and 15° ± 5° of anteversion seemed to be least vulnerable to the processes leading to metal ion release (see Fig. 6). Larger-diameter near-hemispherical components provide greater coverage and can tolerate greater variability in component orientation.

SUMMARY

The results of this study were consistent with the idea that the coverage arc provided by the resurfacing cup is critical in terms of metal ion generation. Smaller diameter bearings with insufficient coverage arcs are vulnerable to unacceptable increases in metal ion concentrations even when relatively well positioned. Of all the patients with the ASR device in this study, 23% were found to have Co concentrations greater than the MHRA guidance of 7 μg/L compared with 6% of the patients with the BHR and C+ devices. This incidence was further reduced to 3.4% when only patients with bearing diameters greater than or equal to 50 mm were included. Serum ion concentrations can reliably be used as surrogate markers of volumetric bearing surface wear.

APPENDIX

Cartesian coordinate data from the CMM were read into Matlab (The Mathworks, Inc, Natick, MA, USA). These data were used to reconstruct the surface of the component. At each point, the radius was calculated using the formula \( \sqrt{X^2+Y^2+Z^2} = r \), where X, Y, and Z represent the Cartesian coordinates of the point and r is the radius of the sphere. By considering only the unworn regions of the samples, it was possible to identify the original component radius in this manner. It was important to disregard the worn areas in this stage because they would adversely affect the calculation. Using a histogram such as that in Fig. 7, it is obvious that the worn radii are the low frequency values and the unworn surface is represented by the high frequency values of small range. With the original radius known, the linear wear depths were calculated by subtracting each calculated radius from the original radius. In the case of femoral heads, wear leads to a smaller radius being calculated. For acetabular cups, the calculated radii are larger.

Once the wear depths were calculated, a grid was interpolated between adjacent points (Fig. 8). This grid was used to calculate wear volume. The size of each grid square was calculated, and this size was multiplied by the mean wear depth at that square to give a volume. These small volumes were summed across the entire surface. A heat map was produced to show this information, allowing for quick visual analysis of wear severity and patterns (Fig. 9).

This method was validated against a gravimetric technique. A sample head was cleaned thoroughly in an acetone bath for 5 minutes. The sample head was left to dry for 1 hour on a lint-free cloth and then weighed on a high-precision scale (Denver Instrument, NY, USA, accuracy 0.1 mg). The sample was weighed 6 times and an average taken. After this, the sample was placed on the CMM using the method described earlier and the resulting data file was analyzed in Matlab. The sample then had a quantity of material removed to simulate wear and was cleaned, weighed, measured, and analyzed again in the same way. This procedure was repeated for a total of 4 times. The results are shown in Table 5.

REFERENCES


Chapter Eight
HIP

Adverse reaction to metal debris following hip resurfacing

THE INFLUENCE OF COMPONENT TYPE, ORIENTATION AND VOLUMETRIC WEAR

We sought to establish the incidence of joint failure secondary to adverse reaction to metal debris (ARMD) following metal-on-metal hip resurfacing in a large, three surgeon, multicentre study involving 4226 hips with a follow-up of 10 to 142 months. Three implants were studied: the Articular Surface Replacement; the Birmingham Hip Resurfacing; and the Conserve Plus. Retrieved implants underwent analysis using a co-ordinate measuring machine to determine volumetric wear. There were 58 failures associated with ARMD. The median chromium and cobalt concentrations in the failed group were significantly higher than in the control group (p < 0.001). Survival analysis showed a failure rate in the patients with Articular Surface Replacement of 9.8% at five years, compared with < 1% at five years for the Conserve Plus and 1.5% at ten years for the Birmingham Hip Resurfacing. Two ARMD patients had relatively low wear of the retrieved components. Increased wear from the metal-on-metal bearing surface was associated with an increased rate of failure secondary to ARMD. However, the extent of tissue destruction at revision surgery did not appear to be dose-related to the volumetric wear.

Metal-on-metal (MoM) hip resurfacing devices were re-introduced following a number of design changes.1,2 The encouraging early results of the Birmingham Hip Resurfacing (BHR; Smith and Nephew, Warwick, United Kingdom) designers’ series3 led to a rapid increase in the number of surgeons carrying out the procedure. A number of manufacturers subsequently developed their own implants, resulting in the large number of resurfacing systems in regular use in Europe and the United States.4 Each commercially available device has a different combination of modifications of the original designs.5 The central design features which are thought to influence wear are: the use of ‘as cast’ versus forged material; varying heat treatments of the femoral and acetabular components; the difference in diameter between mated components (the diametral clearance); the arc of acetabular cover, and the angle of function of the femoral component.6-10

With the technology currently available, the MoM bearing surface remains integral to the design of hip resurfacing. The main drawback of a metal articulation is the production of metal debris due to the combined effect of mechanical and corrosive wear. In fact, the popularity of the procedure has waned in the last two years following a number of reports of adverse reactions in the peri-prosthetic tissues of resurfaced hips.11-15

We have previously highlighted the disparity in blood chromium (Cr) and cobalt (Co) levels between patients receiving different hip resurfacings16 and have identified a relationship between increased wear of the articulat surface and the incidence of soft-tissue lesions.17 The aim of this study was to involve other centres and thereby increase the number of hip resurfacing procedures available for analysis in order to gain more understanding of the incidence of adverse reactions to metal debris (ARMD) in three commonly used hip resurfacing arthroplasty designs, further investigate the clinical effects of increased wear debris from MoM hip resurfacings, and determine whether the extent of tissue destruction in ARMD is related to volumetric wear from the bearing surfaces. ARMD is an umbrella term.17 It is used to describe joint failure secondary to surface wear of the bearing surface or corrosion debris, in the absence of any other obvious explanation. It encompasses metallosis, pseudotumour and aseptic lymphocytic-dominated vasculitis associated lesion (ALVAL).18

Patients and Methods
Three implant designs were used in the study: A: the Articular Surface Replacement (ASR;
ADVERSE REACTION TO METAL DEBRIS FOLLOWING HIP RESURFACING

Table I. Details of the three devices

<table>
<thead>
<tr>
<th>Component*</th>
<th>ASR</th>
<th>BHR</th>
<th>C+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtended articular surface angle (°)</td>
<td>144 to 160</td>
<td>158 to 168</td>
<td>162 to 165</td>
</tr>
<tr>
<td>Mean radial clearance (µm)</td>
<td>50</td>
<td>100</td>
<td>89</td>
</tr>
<tr>
<td>Impingement-free range of movement(°)</td>
<td>37.0 to 28.0</td>
<td>40.2 to 30.0</td>
<td>38.5 to 32.0</td>
</tr>
<tr>
<td>Wall thickness at rim (mm)</td>
<td>3.1</td>
<td>3.6/4.6</td>
<td>3.8</td>
</tr>
<tr>
<td>Manufacturing method of head</td>
<td>As cast</td>
<td>As cast</td>
<td>HIP/SA</td>
</tr>
<tr>
<td>Manufacturing method and treatment of component</td>
<td>HIP/SA</td>
<td>As cast</td>
<td>HIP/SA</td>
</tr>
<tr>
<td>Surface roughness (µm)</td>
<td>0.025</td>
<td>0.029</td>
<td>0.020</td>
</tr>
<tr>
<td>Deviation of roundness head (µm)</td>
<td>3.4</td>
<td>0.9</td>
<td>3.2</td>
</tr>
<tr>
<td>Deviation of roundness component (µm)</td>
<td>3.8</td>
<td>0.9</td>
<td>1.8</td>
</tr>
<tr>
<td>Carbon content</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

* ASR, Articular Surface Replacement; BHR, Birmingham Hip Resurfacing; C+, Conserve Plus
† subtended articular surface angles increase with increasing component diameter
‡ this refers to the maximum arc of movement in a single plane prior to impingement of the femoral neck on the rim of the acetabular component, assuming that the neck is flush to the base of the femoral component and that impingement will take place on the articular rim of the BHR and the C+ components and on the outer rim of the ASR component. The range of motion decreases with increasing component size for all devices
§ HIP/SA, cast process and heat treatment by hot isostatic pressure/surface annealed
¶ High carbon content is defined as ≥ 0.2%
Source: manufacturers' details and independent testing

Table II. Patient demographics, joint orientations and metal ion concentrations (centre 1). All values are mean values (range) unless marked with * in which case they are median values

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Design A</th>
<th>Design B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hips</td>
<td>430</td>
<td>180</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>56 (28 to 77)</td>
<td>51 (32 to 67)</td>
</tr>
<tr>
<td>% Female</td>
<td>44</td>
<td>43</td>
</tr>
<tr>
<td>Follow-up (mths)</td>
<td>37 (10 to 67)</td>
<td>65 (59 to 87)</td>
</tr>
<tr>
<td>Femoral size (mm)</td>
<td>48.6 (39 to 59)</td>
<td>47.6 (38 to 58)</td>
</tr>
<tr>
<td>Inclination angle (°)</td>
<td>48.5 (31 to 70)</td>
<td>48.3 (32 to 70)</td>
</tr>
<tr>
<td>Anteversion angle (°)</td>
<td>20.4 (3 to 39)</td>
<td>19.9 (-5 to 39)</td>
</tr>
<tr>
<td>Harris Hip Score</td>
<td>93 (35 to 100)</td>
<td>97 (51 to 100)</td>
</tr>
<tr>
<td>Total ion results</td>
<td>156</td>
<td>95</td>
</tr>
<tr>
<td>Serum metal ion levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr (µg/L)*</td>
<td>4.00 (0.6 to 115)</td>
<td>4.42 (1.8 to 77)</td>
</tr>
<tr>
<td>Co (µg/L)*</td>
<td>2.60 (0.4 to 228)</td>
<td>1.56 (0.7 to 190)</td>
</tr>
<tr>
<td>Whole blood metal ion levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cr (µg/L)*</td>
<td>3.89 (1.5 to 69.8)</td>
<td>3.92 (2.37 to 40)</td>
</tr>
<tr>
<td>Co (µg/L)*</td>
<td>2.39 (0.4 to 271)</td>
<td>1.44 (0.63 to 147)</td>
</tr>
</tbody>
</table>

Adverse reaction to metal debris failure at a minimum of ten months (%) | 27 (6.3) | 1 (0.56) |

DePuy, Leeds, United Kingdom); B: BHR; C: Conserve Plus (C+; Wright Medical, Memphis, Tennessee). The important differences between the three devices can be seen in Table I. All the patients of three experienced hip resurfacing surgeons (AVFN, KDS, JPH) who received a design A, B or C resurfacing prosthesis between January 1998 and January 2009 were involved in this study. A total of 4226 hips in 3888 patients were studied. Details of the patients are shown in Tables II, III and IV. Surgeon 1 (AVFN) is based in the United Kingdom (University Hospital of North Tees, Stockton, United Kingdom). Between 2002 and 2004 he used design B for all resurfacing procedures. From 2004 he used design A exclusively. This cohort of patients has been described previously. From 2007, patients at this centre have undergone routine whole blood and serum metal ion testing. Surgeon 2 (KDS) is based in Belgium (ANCA...

Table III. Patient demographics, joint orientations and metal ion concentrations (centre 2). All values are mean values (range) unless marked with * in which case they are median values

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Design A</th>
<th>Design B</th>
<th>Design C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hips</td>
<td>59</td>
<td>1922</td>
<td>961</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>54 (27 to 69)</td>
<td>52 (20 to 79)</td>
<td>55 (20 to 80)</td>
</tr>
<tr>
<td>% Female</td>
<td>39</td>
<td>39</td>
<td>41</td>
</tr>
<tr>
<td>Follow-up (mths)</td>
<td>31 (11 to 58)</td>
<td>68 (10 to 142)</td>
<td>34 (10 to 82)</td>
</tr>
<tr>
<td>UCLA¹ score</td>
<td>6.2 (5 to 9)</td>
<td>74 (3 to 10)</td>
<td>7.0 (4 to 10)</td>
</tr>
<tr>
<td>Femoral size (mm)</td>
<td>50.7 (45 to 59)</td>
<td>49.8 (38 to 58)</td>
<td>48.4 (38 to 62)</td>
</tr>
<tr>
<td>Inclination angle (*)</td>
<td>49.8 (32 to 60)</td>
<td>51.8 (32 to 70)</td>
<td>49.0 (31 to 65)</td>
</tr>
<tr>
<td>Anteversion angle (*)</td>
<td>20.8 (0 to 42)</td>
<td>21.8 (0 to 41)</td>
<td>22.3 (3 to 44)</td>
</tr>
</tbody>
</table>

Outcome scores
- Harris Hip Score: 97 (27 to 100) 98 (27 to 100) 98 (29 to 100)
- Total ion results: 26 368 144
- Serum metal ion levels
  - Cr (µg/L)*: 1.6 (0.6 to 184) 2.6 (0.8 to 93) 1.8 (0.3 to 40)
  - Co (µg/L)*: 1.9 (0.6 to 235) 1.9 (0.5 to 119) 1.4 (0.4 to 55)
- Adverse reaction to metal debris failure at a minimum of ten months (%): 2 (3.4) 23 (1.5) 4 (0.42)

Table IV. Patient demographics, and joint orientations (centre 3). Values are mean (range)

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Design B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hips</td>
<td>674</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>49 (22 to 70)</td>
</tr>
<tr>
<td>% Female</td>
<td>32</td>
</tr>
<tr>
<td>Follow-up (mths)</td>
<td>65 (10 to 130)</td>
</tr>
<tr>
<td>Femoral size (mm)</td>
<td>51.0 (38 to 62)</td>
</tr>
<tr>
<td>Inclination angle (*)</td>
<td>44.9 (30 to 64)</td>
</tr>
<tr>
<td>Anteversion angle (*)</td>
<td>14.8 (0 to 28)</td>
</tr>
</tbody>
</table>

Outcome scores
- Harris Hip Score: 97
- Adverse reaction to metal debris (%) 3 (0.45)

Clinic, Ghent, Belgium). He initially used design B but has subsequently also used designs A and C. Routine serum metal ion analysis is undertaken post-operatively. The early results of the patients receiving design B at this centre have been published previously. Surgeon 3 (JPH) is based in the United Kingdom (Freeman Hospital, Newcastle-upon-Tyne, United Kingdom) and uses design B exclusively. All surgeons used the posterior surgical approach.

At all centres, outcomes were assessed at six months, at one year and annually thereafter using the Harris hip score and the UCLA activity score. Any patient whose hip was revised or was listed for revision secondary to ARMD at the time of writing was recorded. At centres 1 and 2, standing radiographs were obtained at the time of blood sampling and analysed using EBRA software (University of Innsbruck, Innsbruck, Austria) to measure the inclination and anteversion of the acetabular component. This was carried out by one of the authors (DJL). At centre 3, the first 100 well-centred digitised standing radiographs were analysed in order to provide an assessment of the orientation of the acetabular component, whose position was allocated a ‘zone’ as shown in Figure 1. Zone 1 corresponds to a safe zone derived from our previous work based upon an analysis of metal ion results and explants.

The diagnosis of ARMD was made by the consultant in charge and was based on clinical presentation, findings at revision and the histological appearance of capsular tissue taken at revision surgery. The appearance of the peri-prosthetic tissues at revision was graded as follows: 0, no soft-tissue necrosis; 1, small localised areas of tissue necrosis; 2, widespread tissue necrosis, stability of the implant not obviously compromised and 3, widespread tissue necrosis with compromised stability of the implant.

Serum metal ion analysis. Blood samples for Cr and Co analysis were taken more than 12 months post-operatively to avoid the confounding effects of the running-in period. The methods of sampling and ion analysis at both centres have been described previously. Histopathological examination of tissues from revision procedures at centre 1. The processing of tissue specimens has also been described previously. In our experience, the vast majority of tissues retrieved from ARMD patients exhibit two dominant, and often co-existent, cellular responses: histiocytic and lymphocytic. For the purpose of this paper, tissues were described as having a dominant 'histiocytic' response if there was a band of histiocytes > 1.5 mm in width or a dominant lymphocytic (ALVAL) response if there were multiple perivascular lymphocytic cuffs > 1.5 mm in size. Metal particulate load in the tissues was graded using a scale
similar to Mirra's classification in order to allow correlation with the rate of volumetric wear. We used a more reproducible grading system derived from the assessment of tissue iron in liver biopsies. It is based upon the ease of identification of particles and the magnification power used to identify them: 0, granules absent or barely discernable at × 400; 1, easily confirmed at × 400, barely discernable at × 250; 2, discrete granules resolved at × 100; 3, discrete granules resolved at × 25; 4, masses visible to the naked eye.

The average diameter of the lymphocytic cuff, the width of the histiocytic band and the integrity of the surface membrane were recorded for each patient and correlated to the rate of volumetric wear.

**Explant analysis.** The wear of all available femoral and acetabular explants from centres 1 and 3 was measured by a co-ordinate measuring machine using a scanning head (Legex 322; Mitutoyo, Halifax, United Kingdom) with a spatial resolution of < 1 μm in the area of measurement. Volumetric wear was calculated using a validated method.

Volumetric wear rates were correlated with serum metal ion results using Spearman rank univariate analysis.

**Survival analysis.** A Cox proportional hazards model was constructed using surgeon and implant design as qualitative variables and bearing diameter as a quantitative variable. Given the numbers involved in the study, a p-value of < 0.01 was considered to be statistically significant.

**Results.**

At the time of writing, there were 60 failures related to ARMD. The incidence of ARMD by centre was as follows. Centre 1. There was a clear difference in the failure rates between the design A and B patient groups. At a mean follow-up of 65 months (59 to 87), one design B patient had been revised. This was in contrast to the design A group, in which 24 hips had been revised and three more were awaiting revision, amounting to a failure rate of 6.3% at a mean follow-up of 37 months (10 to 67).

Centre 2. Two patients with unilateral design A resurfacings had undergone revision, amounting to a rate of revision of 3.4% at a mean follow-up of 31 months (11 to 58). The failure rate secondary to ARMD was 1.2% (23 cases) in the design B group, at a mean follow-up of 68 months (10 to 142) and 0.42% (two cases) in the design C group at a mean follow-up of 37 months (10 to 82).

Centre 3. Only design B was used here, and three patients had been revised. This amounted to a failure rate of 0.45% at a mean follow-up of 65 months (10 to 130).

**Survival analysis.** Cox proportional hazards model showed that patients receiving design A implants were at a significantly greater risk of ARMD failure than patients with designs B and C. Smaller implants were at greater risk of early failure but the surgeon performing the procedure did not significantly affect survival (Table V).

The most common presenting symptom was pain, located predominantly in the groin and occasionally radiating to the greater trochanter and down the thigh, and frequently associated with clicking and clunking sensations. One hip was asymptomatic but as the patient had experienced a severe ARMD with an ASR (DePuy) total hip replacement on one side she asked for the contralateral ASR to be revised. At revision, there were variable degrees and combinations of soft-tissue necrosis, joint effusions in 30, macroscopic metallosis in 18, component loosening in 1, disk granules absent or barely discernable at × 250; 2, discrete granules resolved at × 100; 3, discrete granules resolved at × 25; 4, masses visible to the naked eye.

The average diameter of the lymphocytic cuff, the width of the histiocytic band and the integrity of the surface membrane were recorded for each patient and correlated to the rate of volumetric wear.

**Explant analysis.** The wear of all available femoral and acetabular explants from centres 1 and 3 was measured by a co-ordinate measuring machine using a scanning head (Legex 322; Mitutoyo, Halifax, United Kingdom) with a spatial resolution of < 1 μm in the area of measurement. Volumetric wear was calculated using a validated method.

Volumetric wear rates were correlated with serum metal ion results using Spearman rank univariate analysis.

**Survival analysis.** A Cox proportional hazards model was constructed using surgeon and implant design as qualitative variables and bearing diameter as a quantitative variable. Given the numbers involved in the study, a p-value of < 0.01 was considered to be statistically significant.

**Results.**

At the time of writing, there were 60 failures related to ARMD. The incidence of ARMD by centre was as follows. Centre 1. There was a clear difference in the failure rates between the design A and B patient groups. At a mean follow-up of 65 months (59 to 87), one design B patient had been revised. This was in contrast to the design A group, in which 24 hips had been revised and three more were awaiting revision, amounting to a failure rate of 6.3% at a mean follow-up of 37 months (10 to 67).

Centre 2. Two patients with unilateral design A resurfacings had undergone revision, amounting to a rate of revision of 3.4% at a mean follow-up of 31 months (11 to 58). The failure rate secondary to ARMD was 1.2% (23 cases) in the design B group, at a mean follow-up of 68 months (10 to 142) and 0.42% (two cases) in the design C group at a mean follow-up of 37 months (10 to 82).

Centre 3. Only design B was used here, and three patients had been revised. This amounted to a failure rate of 0.45% at a mean follow-up of 65 months (10 to 130).

**Survival analysis.** Cox proportional hazards model showed that patients receiving design A implants were at a significantly greater risk of ARMD failure than patients with designs B and C. Smaller implants were at greater risk of early failure but the surgeon performing the procedure did not significantly affect survival (Table V).

The most common presenting symptom was pain, located predominantly in the groin and occasionally radiating to the greater trochanter and down the thigh, and frequently associated with clicking and clunking sensations. One hip was asymptomatic but as the patient had experienced a severe ARMD with an ASR (DePuy) total hip replacement on one side she asked for the contralateral ASR to be revised. At revision, there were variable degrees and combinations of soft-tissue necrosis, joint effusions in 30, macroscopic metallosis in 18, component loosening in 8 and masses in 17. In some cases the effusion was massive (> 200 ml) and had extended through the abductor.
Table VI. Comparison of size, component orientations and metal ion levels between the asymptomatic and adverse reaction to metal debris (ARMD) cohorts. Median (range) values are given with Mann-Whitney U test for non-parametric data used to establish significance of results.

<table>
<thead>
<tr>
<th></th>
<th>ARMD</th>
<th>Asymptomatic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femoral size</td>
<td>46.00 (38 to 57)</td>
<td>50.00 (38 to 62)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Component inclination (°)</td>
<td>50.95 (31 to 70)</td>
<td>49.05 (23 to 78)</td>
<td>0.110</td>
</tr>
<tr>
<td>Component anteversion (°)</td>
<td>25.00 (-9 to 44)</td>
<td>16.28 (-9 to 42)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Serum Cr (µg/L)</td>
<td>26.05 (1.6 to 146)</td>
<td>2.9 (0.5 to 110)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Serum Co (µg/L)</td>
<td>39.45 (1.1 to 228)</td>
<td>1.9 (0.3 to 155)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Graph showing failure rates for each device, with patients split into groups by femoral component size. All patients in the study are included, although the design group A patients with femoral size of 39 mm are removed for graphical representation. There were two patients in this group, with one failure, amounting to a 50% failure rate.  

![Graph showing failure rates for each device](image)

**Fig. 2**

Box plot showing metal ion data for all implant designs. The inferior, middle and superior horizontal lines of the boxes represent the first quartile, median and third quartile. The ends of the whiskers correspond to the limits of the data, beyond which values are considered anomalous. The mean is displayed with a +, outliers with a *, extreme outliers with a ** and upper and upper and lower values with a *.

![Box plot showing metal ion data for all implant designs](image)

**Fig. 3**

musculature. The asymptomatic patient described above had severe soft-tissue destruction. Four patients presented with painless swellings in the lateral thigh and groin (two design B, two design A), one associated with femoral nerve symptoms. Revision surgery in these cases revealed psoas bursae containing caseous material. At centre 1, five male patients with design A implants were found to have increased levels of Cr and Co on routine screening. As they were asymptomatic they were simply observed. All five patients became symptomatic. Fractures of the femoral neck in association with gross macroscopic metallosis were found at revision. A psoas mass was also identified in one of these cases. Soft-tissue destruction was not extensive in these cases. Patients who were found to have a large effusion at revision surgery were revised significantly earlier than those found to have masses at revision (median time to revision (effusion) 21 months versus 52 months (masses), p < 0.001).

Table VI shows the significant differences in bearing size, acetabular component orientation and metal ion levels between the asymptomatic and ARMD cohorts. There was a clear trend towards increased risk of failure with decreasing size of femoral component in the three centres (Fig. 2). Figure 1 shows the relationship between the angle of inclination and anteversion of the acetabular component and failure. In two ARMD acetabular components we found they were in the safe zone for metal ion reduction (45° ± 5° inclination and 15° ± 5° for anteversion). An overview of the metal ion results for each implant...
group can be seen in Figure 3. Median levels of Cr and Co in ARMD patients (n = 37 with pre-revision ion levels) were significantly higher than in the asymptomatic cohort, with the median Co concentration × 20 greater in the failed group. The two patients described above, with acetabular components in the safe zone, did have ion levels comparable with the asymptomatic cohort (Co levels of 2.1 μg/l and 1.9 μg/l). Histology of the tissue specimens showed no macroscopic metallosis but a dominant ALVAL reaction. Both of these patients had grade 2 soft-tissue destruction at revision surgery.

Total rates of volumetric wear correlated well with serum Cr (r = 0.847, p < 0.001) and Co levels (r = 0.732, p < 0.001) and with particulate tissue load (r = 0.60, p < 0.001), but not with perivascular lymphocytic cuff thickness, histiocyte band width, surface membrane necrosis or macroscopic tissue necrosis at revision surgery. A relationship was identified between lymphocytic cuff diameter and surface membrane necrosis (r = 0.55, p = 0.011).

Discussion
This paper represents the largest collection of clinical and biochemical results from hip resurfacing patients in the current literature. We believe that the three-centre nature of the study, the experience of the surgeons involved and the use of three resurfacing designs provides a fair representation of the performance of modern hip resurfacing in the wider orthopaedic community.

Currently there is increasing concern over the potential adverse effects of metal debris. It remains to be shown whether these adverse reactions are dose-dependent and whether they are mediated primarily by an immune response to, or a direct toxic effect of the metal debris. Reported rates of ALVAL,21 pseudotumours9 and metallosis12 vary throughout the literature. There appears to be no consensus as to the boundaries between the described conditions. Liu et al9 define pseudotumour as “a soft-tissue mass associated with the implant which is neither malignant nor infective in nature”. It is not clear whether this term includes joint effusion. ALVAL is a histological diagnosis21 which has also been used to describe the clinical appearance of tissue necrosis and abnormal joint fluid at revision surgery.30 Metallosis is defined as aseptic fibrosis, local necrosis, or loosening of a device secondary to metallic corrosion and release of wear debris.31,32 At centre 1, metallic debris was identified, at least on a microscopic scale, in every tissue sample whether this term includes joint effusion. ALVAL is a pseudotumour, but a dominant ALVAL reaction. Both of these patients had grade 2 soft-tissue destruction at revision surgery.

In a smaller, single surgeon, single-site study, we previously highlighted clear differences in the release of metal ions between two resurfacing devices (designs A and B in the present study).16 De Smet et al27 demonstrated that serum metal ion concentrations correlate well with wear of retrieved femoral components. By using Cr and Co analysis as a surrogate indicator of wear, we therefore concluded that increased wear was associated with an increased probability of early failure. These results have been substantiated here. In this series, design A has the highest rate of failure secondary to ARMD (Fig. 4). It is also the most vulnerable to the effects of variation in acetabular component position in terms of ion generation. Design B is associated with a smaller range of Cr and Co values than design A and this is reflected in the lower failure rates secondary to ARMD. Patients receiving design C have the lowest serum ion levels and also the lowest failure rates secondary to ARMD. An obvious confounding factor here, however, is that the design C patients have a shorter mean follow-up than those with design B resurfacings. The Cox proportional hazards model should account for this factor in our analysis; however, we interpret these results with caution.

The proportional hazards model we used will not account for unknown specific design features which may have an unexpected effect once a certain period of time has elapsed.

There is overwhelming evidence to show that surgeons cannot consistently position the acetabular components precisely. Without exception, studies show wide variations in the angles of inclination of the acetabular component and, to an even greater extent, its anteversion (Fig. 5).34 Surgeons must accept that some variables may be beyond their control, for example changes in pelvic tilt35 during
pre-operative positioning, intra-operative pelvic rotation and patient size. Consequently they should choose implants which are compatible with the inevitable variability of acetabular component orientation.

While the three resurfacing systems have differences of design including varying clearances, heat treatment or no heat treatment of both the femoral and acetabular components, we believe that the design feature most likely to explain the disparity in performance is the arc of acetabular cover. The acetabular component of design C has a smaller arc of cover than the other devices (Table I). Despite this decreased cover, the range of motion prior to impingement is not compromised firstly, due to the recessed nature of the articular surface and also due to the smaller functional angle of the femoral component relative to the other two designs. These features increase the vulnerability of the device to the two mechanisms which appear to be critical in the acceleration of wear rates: edge-loading and microseparation/subluxation. The smaller arc of the cover, i.e. the more shallow the acetabular component, the greater is the tendency to rim-loading at equivalent angles of inclination and anteversion, when matched for size.

Line graph showing acetabular component placement zones as described in Figure 2. This chart shows the incidence of adverse reaction to metal debris (ARMD) for each design with relation to acetabular component orientation.

The relationship between the amount of wear debris and the extent of tissue destruction is not straightforward. Using serum ion levels as a surrogate measure, our results suggest that, when exposed to low levels of wear, most patients with a resurfaced hip do extremely well. Whether or not there is a causal relationship, patients exposed to more debris are more likely to have complications. In this series, the patients with extremely high levels of metal ions/ wear from the bearing surfaces did not, however, present exclusively with worsening pain, but with other symptoms including delayed fracture of the femoral neck and/or masses. At revision surgery, macroscopic metallosis was the dominant feature. This was in contrast to the striking appearance of joint fluid and tissue necrosis, which was seen in patients with moderately increased wear/metal ion levels. Neither volumetric wear rates nor joint/serum ion concentrations correlated with microscopic necrosis or the extent of tissue destruction observed at revision surgery. The relationship between the thickness of perivascular lymphocytic cuffs and surface layer necrosis suggests that tissue destruction is not a result of toxic concentrations of metal debris, but is more likely to be the result of an immune response provoked by the debris. Patients found to have pseudotumours were revised at a significantly longer period from primary surgery than those with no masses. We speculate that pseudotumours are part of the same pathological spectrum of disease which causes joint effusions and tissue death, but is a more advanced stage of the disease process. Metallic debris produced by different devices may have different sizes and shapes and wear particles produced by design A could conceivably have greater immunogenicity.
The central message, however, is that in the short- to midterm at least, the vast majority of patients with MoM resurfaced hips will not experience severe soft-tissue reactions when exposed to the levels of wear associated with well-functioning bearing surfaces. We conclude that increased wear from MoM hip resurfacings is associated with an increased probability of adverse clinical outcomes. These adverse outcomes include severe destruction of soft tissues and bony necrosis. It is likely that loosening of metal-on-metal bearings and hyper-sensitivity in patients with artificial joint hips: a clinical and histomorphological study. J Bone Joint Surg [Am]2005;87-A:26-36.


Adverse reactions to metal debris: histopathological features of periprosthetic soft tissue reactions seen in association with failed metal on metal hip arthroplasties

Sonali Natu, Raghavendra Prasad Sidaginamale, Jamshid Gandhi, David J Langton, Antony V F Nargol

ABSTRACT
Aim To describe the histopathology of localised adverse reactions to metal debris (ARMD) seen in association with failed metal on metal (MoM) hip arthroplasties. The nature of aseptic lymphocytic vasculitis associated lesion (ALVAL) is investigated.

Methods Periprosthetic soft tissues biopsied at time of revision from failed MoM hip arthroplasties from January 2007 to March 2011 were analysed. The inflammatory cell response was categorised into perivascular lymphocytic cuffing (ALVAL), lymphoid aggregate formation with or without germinal centres, metallosis characterised by sheets of macrophages with intracytoplasmic metallic debris and well-defined granulomas.

Results 123 patient samples were analysed, 36 males (29.2%) and 87 females (70.8%). Three cases showing complete tissue necrosis were excluded. Patients were reviewed between 3.27 to 69.6 months postarthroplasty, with an average of 30.92 months. 103 cases (85.8%) showed ALVAL, 18 cases also showed well-defined granulomas. Of the 103 cases with ALVAL, 60 cases also showed a diffuse chronic lymphocytic synovitis, and 40 cases showed lymphoid aggregates with germinal centres. 17 cases (14.2%) showed pure metallosis. Small vessels showing ALVAL expressed peripheral node addressin.

Conclusions ARMD is a spectrum of changes comprising of pure metallosis, ALVAL and granulomatous inflammation. ALVAL, a distinctive inflammatory response seen in ARMD, is a precursor of lymphoid neogenesis. Lymphoid neogenesis documented in a variety of chronic inflammatory conditions most probably contributes to tissue necrosis and prosthetic failure seen in MoM hip arthroplasties. The role of vascular changes in contributing to necrosis is unclear at this stage.

INTRODUCTION
Metal on metal total hip arthroplasties were used in the 1960s and 1970s. Modern hip resurfacing became a popular procedure in the late 1990s. It was introduced in younger patients as only the degenerate cartilage was removed along with a thin slice of bone, bone stock was preserved and hip replacement could be used in the future, in the event of failure. Over the years, metal hip articulations gained popularity and were preferred over polyethylene devices, in cases where osteolysis and aseptic loosening of the prosthesis were a concern. A variety of resurfacing devices are present in the market, which consist of ceramic or metal together or on their own. Metal on metal (where the acetabular cup and femoral head are resurfaced with metal) is a popular option and the National Institute for Health and Clinical Excellence guidelines (UK) for its use were published in 2002. The metal used is usually an alloy of cobalt and chromium. When the triad of patient characteristics, type of device and surgical technique are optimised, the chances of failure of the device are minimised. On implanting metal on metal devices, there is a usual run in period of approximately 6-12 months when blood concentrations of cobalt/chromium rise due to wear, and reach a steady state thereafter. There is a simultaneous rise in local tissue levels of metal ions. Failure of metal on metal hip devices is usually associated with raised tissue levels and blood levels of metal ions. The Medicines and Healthcare products Regulatory Agency, UK, has issued a warning for all metal on metal hip implants. This involves monitoring of blood metal ion levels and looking for symptoms of pain and large hip joint effusions.

Willert et al introduced the concept of aseptic lymphocytic vasculitis associated lesion (ALVAL) in 2005. A lymphocyte dominated immunological response within the periprosthetic tissues from metal on metal hip articulations obtained at the time of revision was described. The inflammatory response comprised a diffuse and perivascular chronic lymphocytic infiltrate with development of high endothelial venules (HEVs). Davies et al further detailed this unusual perivascular lymphocytic inflammation within the synovium. It was considered to be unusual because the usual response seen in polyethylene devices consisting of a macrophagic response with engulfed polyethylene debris leading to osteolysis was not seen. A semiquantitative analysis of the chronic inflammation was performed in various hip prostheses including metal on metal and metal on polyethylene. The perivascular lymphocytic cuffing was typically seen within the deeper subsynovium in metal on metal articulations. Both studies noted surface necrosis and metal debris, but did not clearly state the significance of the perivascular lymphocytic cuffing with regard to vascular damage. Subsequently, there have been articles describing the adverse reactions clinically
as pseudotumours, which are cystic/solid masses developing in relation to failed metal prostheses. These studies noted the presence of ALVAL or perivascular lymphocytic cuffing, but did not elaborate any further on whether there was lymphocyte-mediated vascular damage, that is, fibrinoid necrosis, on skinning, hyalinisation and/or vascular occlusion. Evans et al. in 1974 noted vascular obliterator changes in the form of fibrous intimal proliferation and fibrinoid necrosis in the periprosthetic tissues and concluded the vascular changes caused tissue and bone necrosis. Similar vascular changes were also seen by Brown et al., Jones et al. and Winter et al. However, Brown et al. could not conclude that the vascular wall changes seen led to the necrosis as such changes were commonly seen in the synovium affected by a chronic disease process.

Lymphoid neogenesis, also known as formation of tertiary lymphoid organs, has been described in various chronic inflammatory conditions including Helicobacter-induced active chronic gastritis, autoimmune diseases, neoplasia and graft rejection. The reorganisation of a chronic inflammatory cell response into lymphoid aggregates and follicles with germinal centres has been the subject of great interest. This is a dynamic process in which lymphocytic infiltrate evolves into lymphoid aggregates with germinal centres. Development of HEVs and dendritic cell networks play an important role in lymphoid neogenesis. In the face of a sustained immune response to a persistent antigen, the local tissue chronic inflammatory cell response is organised into tertiary lymphoid organs with ectopic germinal centres, capable of locally generating B cells and T cells. Anatomically, tertiary lymphoid organs are similar to secondary lymphoid organs also known as lymph nodes.

Lymphoid neogenesis while optimising the local immune response causes considerable tissue destruction and loss of function. There has been a vast amount of interest in its pathophysiology in the hope that it could be a potential target for treatment of chronic inflammatory conditions. Within the synovium, lymphoid neogenesis has been documented in inflammatory arthropathies but more significantly in rheumatoid arthritis (RA). In this article, we will describe the histopathology seen in association with metal on metal hip arthroplasties including ALVAL, lymphoid neogenesis, granulomatous inflammation and mettabolism. The pathology has been described by a generic all encompassing term, adverse reactions to metal debris (ARMD). The histopathology is briefly compared with other common inflammatory arthropathies.

**MATERIALS AND METHODS**

**Patients**

As a routine procedure, all patients considered for revision of metal prostheses undergo a preoperative protocol, according to the Medicines and Healthcare products Regulatory Agency, UK, guidance. This includes the measurement of Harris hip scores, updated pelvic radiographs, blood metal ion testing, ultrasound scan or aspiration of the hip joint effusion. During revision surgery, intraoperative macroscopic appearances were meticulously recorded by the senior surgeon, and tissue samples from periprosthetic soft tissue were sent to histopathology for analysis. Tissue samples were provided from between two and four sites surrounding the metal on metal implant. There was no attempt made to distinguish if the tissue was sent from the hip capsule, acetabulum or bursa. No femoral or bony tissue was included in this study. Up to 10 paraffin blocks/cassettes were processed per site, to avoid sampling error due to necrosis and to ensure the viable tissue was well represented. Tissue samples were sent to microbiology to rule out sepsis. All explants underwent wear analysis using the coordinate measuring machine at an independent centre. In this study, we have included all the samples received from January 2007 to March 2011.

**Histology**

Routine H&E stained slides were examined by a histopathologist independently of the clinical findings. Up to six tissue levels were examined where a chronic lymphocytic infiltrate was noted within and around large blood vessels (a larger vessel is defined as any vessel larger than an arteriole or venule). T lymphocyte markers used were CD4, CD8 and CD43 (SP57, L60 Ventana, Roche. Prediluted), and B lymphocyte markers used were CD 20 (L26 Ventana, Roche. Prediluted). HEVs were stained with MECA-79, also known as peripheral node addressin (PDNA) (Novus Biologicals, Littleton USA Dilution, 10 µg/ml).

The surface tissue necrosis was typed according to Davies et al. Type 1 surfaces showed intact synovial surface epithelium. In Type 2 surfaces there was a loss of the synoviocyte layer without fibrin deposition. Type 3 surfaces were associated with fibrin deposition, and in Type 4 surfaces, there was extensive necrosis with loss of architecture. The extent of Type 4 surface necrosis was used to grade the overall tissue necrosis within a given sample. In Grade 4 necrosis, more than 75% of the tissue sample showed Type 4 necrosis. In Grade 3 necrosis, between 25% and 75% showed Type 4 necrosis. Necrosis was considered to be Grade 2 when either less than 25% of the tissue showed Type 4 necrosis or the tissue showed Type 3 surface. Grade 1 necrosis consisted of Type 2 surface.

**The inflammatory response**

On H&E stained sections, the chronic lymphocytic infiltrate present was categorised into: (a) diffuse synovitis, (b) lymphoid aggregate synovitis and (c) germinal centre containing synovitis.

Diffuse synovitis was defined as a chronic lymphocytic infiltrate not organised into follicles or aggregates. Lymphoid aggregates were defined as perivascular collections of lymphocytes. The blood vessels in the centre of the aggregates were similar to HEVs seen within lymph nodes. Lymphoid aggregates contained two or more such vessels. Lymphocytic cuff thickness was measured using a graticule (eyepiece micrometre disc 21 mm Nikon Eclipse 80i). The radial thickness was measured from a centre within a cuff along the longest axis of the cuff. An average of approximately 10 cuffs was taken. Lymphocytic cuff thickness was graded 1–4. Grade 1 cuffs were <0.25 mm, Grade 2 were from 0.25 to 0.5 mm, Grade 3 were from 0.5 to 0.75 mm and Grade 4 were >0.75 mm. Lymphoid aggregates with germinal centres consisting of immature lymphoblasts, with or without plasma cells, were classified as germinal centre containing synovitis. T cell and B cell lymphocytic markers showed organisation into T cell and B cell areas. The presence of lymphoid aggregates within skeletal muscle if present was noted.

**Vascular changes within blood vessels surrounded by the lymphoid infiltrate**

Vascular changes such as hyalinisation and onion skinning due to pericyte proliferation were noted. Hyalinisation extending
into the surrounding extracellular tissues was stained with periodic acid Schiff (PAS) and Congo red.

Other features
Granulomatous inflammation whether organised into well-defined granulomas or sheets of histiocytes containing metallic debris was noted. The thickness of the sheets of histiocytes present was measured with the graticule. Grade 1 was <1 mm thick, Grade 2 between 1 and 2 mm and Grade 3 >2 mm. Metal particle load within the macrophages was assessed by the method used to assess tissue iron overload in liver biopsies. Energy dispersive x-ray spectrometry (EDX) was used to confirm the presence of metal particles and the components on deparaffinised tissue. The microscope used was environmental scanning electron microscope field emission gun (FEI XL30 ESEM-FEG). The samples were coated with carbon in order to make them electrically conducting. High vacuum was used to look at the samples for both imaging and analysis.

RESULTS
There were 123 patient samples in total; three of these showed complete tissue necrosis with no viable tissue for assessment and were excluded from the study. There were 36 males and 87 females, with an age range from 32 to 82 years. There were 87 (70.7%) patients who were under the age of 65 years. Indications for the arthroplasty were primary, secondary osteoarthritis and congenital hip dysplasia. None of these patients had RA. The patients were reviewed from between 5.27 and 69.6 months with an average of 30.92 months. Intraoperatively, there was variable amount of fluid collection within the periprosthetic tissue with foci of metallic tingeing. Necrotic tissue was often cheesy in appearance. Grossly, the necrotic tissue showed areas of metallic tingeing. The tissue was thickened, fibrotic and partly necrotic. With large fluid collections, the tissue had the appearance of a thinned out cyst wall. Rarely, there were preserved synovial villi. None of the cases showed features of sepsis on histology and microbiology. EDX was used on two cases, which confirmed the presence of cobalt and chromium. The type of metal on metal implants used is shown in table 2.

Typically, all cases showed varying grades of surface necrosis with complete loss of surface synoviocytes and architecture (figure 1A–H). The pathology seen is summarised in tables 3 and 4. The surface synovium was not intact in any of our cases, and there was no synoviocyte hyperplasia or villous hyperplasia. Multinucleated synoviocytes were not seen. Ghost outlines of necrotic synovial villi could be seen. Beneath the necrosis, aggregates and sheets of histiocytes containing fine black metallic debris were present. Often the histiocytes were arranged in aggregates adjacent to small blood vessels. The histiocytes had slate grey discolouration. In two cases, the histiocytes had a xanthomatous appearance. The metal particles were either needle shaped (figure 2A) or dot-like (figure 2B). Foreign body giant cells with larger metal debris were interspersed. In 17 cases, there was a pure metallosis type reaction without lymphocytic infiltration (figure 2C). Well defined naked sardoid-like granulomas were seen in 18 cases (figure 2D) in addition to the chronic lymphocytic infiltrate. The majority of the granulomas consisted of walling off epithelioid histiocytes, and multinucleated giant cells when present were of Touton type. Occasionally, necrotic bony fragments or calcific debris could be seen beneath the areas of necrosis. The deeper tissues showed the characteristic perivascular lymphocytic cuffing (103 cases) (figure 3A) with or without germinal centres. Germinal centres were seen in 40 cases. The germinal centres had a paler centre with immunoblasts with occasional plasma cells (figure 3B). In places with florid lymphoid aggregate formation with germinal centres, the inflammatory response bears a resemblance to small lymph nodes but without a capsule or peripheral sinuses. Russel bodies were not seen. Plasma cells in small numbers were seen in 45 cases with ALVAL. T and B lymphocytic markers showed organisation into distinct central B cell areas with surrounding T cell areas (figure 3C,D). A background of diffuse chronic lymphocytic inflammation was seen in 60 cases. This lymphocytic infiltrate consisted predominantly of admixture of CD4 and CD8 T lymphocytes, and plasma cells. Plasma cell infiltrate is usually mild in comparison with the lymphocytic infiltrate. Eosinophils and neutrophil polymorphs were seen in only one case in our study. Overall, the inflammatory cell response can be patchily distributed and can be of varying intensity in different sections. HEVs were seen centrally within the lymphocytic aggregates. HEVs showed plump cuboidal endothelium, and lymphocytes could be seen transiting from the lumen (figure 4A,B). The vessels were PNA positive (figure 4C,D), while those vessels not surrounded by the infiltrate were PNA negative. Structural changes in the vessel walls could be seen in the form of hyalisation, onion skinning and luminal obliteration (figure 5A–D) in 47 cases. The material was PAS positive and Congo red negative. Fibrinoid necrosis of the vessel wall was not seen. Leucocytoclasis was not seen, but in one case nuclear dust was noted (figure 6C). Eosinophils and neutrophil polymorphs were seen in small numbers in a predominantly lymphocytic infiltrate in this case (figure 6A,B,D). In five cases, vessels larger than the calibre of arterioles or venules showed lymphocytic infiltrate in the walls. In one case, the lymphocytes were identified to be CD8 T lymphocytes. The infiltrate was segmental and cut out in further levels (figure 7A–D).

DISCUSSION
The spectrum of changes seen in ARMD is distinctive. The conglomerate of surface necrosis, with macrophagic response containing fine metallic debris, sometimes forming granulomas, along with an evolving ALVAL/perivascular lymphocytic

<table>
<thead>
<tr>
<th>Grade</th>
<th>Ease of observation and magnification (eyepiece × objective lens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Granules absent or barely discernable ×400</td>
</tr>
<tr>
<td>1+</td>
<td>Barely discernable ×250, easily confirmed ×400</td>
</tr>
<tr>
<td>2+</td>
<td>Discrete granules resolved ×100</td>
</tr>
<tr>
<td>3+</td>
<td>Discrete granules resolved ×25</td>
</tr>
<tr>
<td>4+</td>
<td>Masses visible ×10, naked eye</td>
</tr>
</tbody>
</table>

**Table 2** Distribution of the type of metal implants

<table>
<thead>
<tr>
<th>Implant type</th>
<th>Resurfacing, n = 69</th>
<th>Stemmed, n = 54</th>
</tr>
</thead>
<tbody>
<tr>
<td>36 MOM</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Adopt</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ASR</td>
<td>59</td>
<td>0</td>
</tr>
<tr>
<td>ASR XL</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>BHR</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>BHR Hybrid</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Conserve</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 1** Assessment of metal particle load within the macrophages

<table>
<thead>
<tr>
<th>Grade</th>
<th>Ease of observation and magnification (eyepiece × objective lens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Granules absent or barely discernable ×400</td>
</tr>
<tr>
<td>1+</td>
<td>Barely discernable ×250, easily confirmed ×400</td>
</tr>
<tr>
<td>2+</td>
<td>Discrete granules resolved ×100</td>
</tr>
<tr>
<td>3+</td>
<td>Discrete granules resolved ×25</td>
</tr>
<tr>
<td>4+</td>
<td>Masses visible ×10, naked eye</td>
</tr>
</tbody>
</table>
infiltrate with or without germinal centres combined with a history of metal on metal implants is diagnostic of ARMD. The surface necrosis is variable and extensive. When there is minimal necrosis and the ALVAL response is not well developed, and because the synovium shows a limited repertoire of chronic inflammatory response, the tissue response in ARMD may be difficult to distinguish from other more common chronic inflammatory arthropathies. Lymphoid aggregate synovitis with or without germinal centres is seen in chronic inflammatory arthropathies including osteoarthritis, psoriatic arthritis and more commonly in RA. The majority of rheumatoid synovia show a diffuse chronic synovitis-like picture. A third of the rheumatoid patients show follicular synovitis with or without germinal centre formation. There is no associated surface necrosis, and synovial villous hyperplasia with synoviocyte proliferation is typically seen. Synovial multinucleated giant cells are seen on the surface, with the underlying tissue showing variable chronic inflammation. Plasma cells with Russel bodies are commonly seen. The inflammation in psoriatic arthritis does not vary significantly from RA, and the synovial inflammation seen in osteoarthritis is of a milder degree and intensity as compared with RA. What distinguishes ARMD seen in metal on metal tissues from others is the feature of extensive but variable surface necrosis. The inflammation is less often diffuse and is

Figure 1 (A—D, E—H) These photomicrographs (magnification ×100) show surface necrosis, beneath which are seen sheets and aggregates of histiocytes containing metallic debris. Aseptic lymphocytic vasculitis associated lesion response is seen in the deeper tissues, for example, adipose tissue (G) and skeletal muscle (H).
### Table 3: Spectrum of changes in ARMD seen in this study

<table>
<thead>
<tr>
<th></th>
<th>Stemed hip arthroplasty (54)</th>
<th>Resurfacing hip arthroplasty (59)</th>
<th>Total (123)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
<td>20</td>
<td>36</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>49</td>
<td>87</td>
</tr>
<tr>
<td><strong>Necrosis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 1</td>
<td>02</td>
<td>07</td>
<td>9 (7.3%)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>19</td>
<td>21</td>
<td>40 (32.5%)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>06</td>
<td>15</td>
<td>21 (17.1%)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>27</td>
<td>26</td>
<td>53 (43.1%)</td>
</tr>
<tr>
<td><strong>Histiocytes sheets</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm</td>
<td>23</td>
<td>12</td>
<td>35 (29.2%)</td>
</tr>
<tr>
<td>1–2 mm</td>
<td>15</td>
<td>24</td>
<td>39 (32.5%)</td>
</tr>
<tr>
<td>&gt;2 mm</td>
<td>10</td>
<td>27</td>
<td>37 (30.8%)</td>
</tr>
<tr>
<td>Absent</td>
<td>04</td>
<td>05</td>
<td>09 (7.5%)</td>
</tr>
<tr>
<td><strong>Granuloma</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>37</td>
<td>65</td>
<td>102 (85%)</td>
</tr>
<tr>
<td>Present</td>
<td>15</td>
<td>03</td>
<td>18 (15%)</td>
</tr>
<tr>
<td><strong>Particle load</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0+</td>
<td>10</td>
<td>07</td>
<td>17 (14.2%)</td>
</tr>
<tr>
<td>1+</td>
<td>11</td>
<td>18</td>
<td>29 (24.2%)</td>
</tr>
<tr>
<td>2+</td>
<td>22</td>
<td>25</td>
<td>47 (39.2%)</td>
</tr>
<tr>
<td>3+</td>
<td>05</td>
<td>14</td>
<td>19 (15.8%)</td>
</tr>
<tr>
<td>4+</td>
<td>04</td>
<td>04</td>
<td>8 (6.6%)</td>
</tr>
<tr>
<td><strong>Diffuse synovitis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>23</td>
<td>37</td>
<td>60 (50%)</td>
</tr>
<tr>
<td>Present</td>
<td>29</td>
<td>31</td>
<td>60 (50%)</td>
</tr>
<tr>
<td><strong>Plasma cells</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>33</td>
<td>42</td>
<td>75 (62.5%)</td>
</tr>
<tr>
<td>Present</td>
<td>19</td>
<td>26</td>
<td>45 (37.5%)</td>
</tr>
<tr>
<td><strong>Lymphoid aggregates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>07</td>
<td>10</td>
<td>17 (14.2%)</td>
</tr>
<tr>
<td>Present</td>
<td>45</td>
<td>58</td>
<td>103 (85.8%)</td>
</tr>
<tr>
<td><strong>Germinal centre</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>32</td>
<td>48</td>
<td>80 (66.7%)</td>
</tr>
<tr>
<td>Present</td>
<td>20</td>
<td>20</td>
<td>40 (33.3%)</td>
</tr>
<tr>
<td><strong>Lymphocyte cuff thickness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;0.25 mm</td>
<td>25</td>
<td>34</td>
<td>59 (49.2%)</td>
</tr>
<tr>
<td>0.25–0.5 mm</td>
<td>14</td>
<td>13</td>
<td>27 (22.5%)</td>
</tr>
<tr>
<td>0.5–0.75 mm</td>
<td>03</td>
<td>05</td>
<td>8 (6.7%)</td>
</tr>
<tr>
<td>&gt;0.75 mm</td>
<td>03</td>
<td>06</td>
<td>9 (7.5%)</td>
</tr>
<tr>
<td>Absent</td>
<td>07</td>
<td>10</td>
<td>17 (14.1%)</td>
</tr>
<tr>
<td><strong>ALVAL in skeletal muscle</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>46</td>
<td>58</td>
<td>104 (86.7%)</td>
</tr>
<tr>
<td>Present</td>
<td>06</td>
<td>10</td>
<td>16 (13.3%)</td>
</tr>
<tr>
<td><strong>Vascular wall changes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>33</td>
<td>40</td>
<td>73 (60.8%)</td>
</tr>
<tr>
<td>Present</td>
<td>20</td>
<td>27</td>
<td>47 (39.2%)</td>
</tr>
</tbody>
</table>

ALVAL, aseptic lymphocytic vasculitis associated lesion; ARMD, adverse reactions to metal debris.

### Table 4: Distribution of cases with vascular wall changes

<table>
<thead>
<tr>
<th>Vascular wall changes</th>
<th>Distribution of cases with vascular wall changes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>47/120 (39.2%)</td>
</tr>
<tr>
<td><strong>Necrosis</strong></td>
<td></td>
</tr>
<tr>
<td>Grade 1</td>
<td>0/9</td>
</tr>
<tr>
<td>Grade 2</td>
<td>13/40 (32.5%)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>50/21 (42.8%)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>25/53 (47.1%)</td>
</tr>
<tr>
<td><strong>Histiocytes sheets</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;1 mm</td>
<td>19/35 (54.2%)</td>
</tr>
<tr>
<td>1–2 mm</td>
<td>17/39 (43.5%)</td>
</tr>
<tr>
<td>&gt;2 mm</td>
<td>11/37 (29.7%)</td>
</tr>
<tr>
<td><strong>Granuloma</strong></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>34/102 (33.3%)</td>
</tr>
<tr>
<td>Present</td>
<td>13/18 (72.2%)</td>
</tr>
<tr>
<td><strong>Particle load</strong></td>
<td></td>
</tr>
<tr>
<td>0+</td>
<td>09/17 (52.9%)</td>
</tr>
<tr>
<td>1+</td>
<td>11/29 (37.9%)</td>
</tr>
<tr>
<td>2+</td>
<td>18/47 (38.2%)</td>
</tr>
<tr>
<td>3+</td>
<td>07/19 (36.8%)</td>
</tr>
<tr>
<td>4+</td>
<td>02/08 (25%)</td>
</tr>
<tr>
<td><strong>Plasma cells</strong></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>23/75 (30.6%)</td>
</tr>
<tr>
<td>Present</td>
<td>24/45 (53.3%)</td>
</tr>
<tr>
<td><strong>Germinal centre</strong></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>20/80 (25%)</td>
</tr>
<tr>
<td>Present</td>
<td>27/40 (67.5%)</td>
</tr>
<tr>
<td><strong>Lymphocyte cuff thickness</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;0.25 mm</td>
<td>19/59 (32.2%)</td>
</tr>
<tr>
<td>0.25–0.5 mm</td>
<td>15/77 (55.5%)</td>
</tr>
<tr>
<td>0.5–0.75 mm</td>
<td>15/66 (75%)</td>
</tr>
<tr>
<td>&gt;0.75 mm</td>
<td>07/99 (77.7%)</td>
</tr>
<tr>
<td><strong>ALVAL in skeletal muscle</strong></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>37/104 (35.5%)</td>
</tr>
<tr>
<td>Present</td>
<td>10/16 (62.5%)</td>
</tr>
</tbody>
</table>

ALVAL, aseptic lymphocytic vasculitis associated lesion.

Lymphoid neogenesis is a dynamic process in which perivascular lymphocytic cuffs expand and evolve into lymphoid aggregates with germinal centres. During this process, there is significant tissue remodelling of the inflamed tissues. Blood vessels acquire the morphological and immune-phenotype of HEVs as seen in lymph nodes. HEVs are efficient in homing and recruiting lymphocytes by expression of relevant ligands like PNA. Follicular dendritic cell network and a fibroblastic reticular framework are the essential ingredients for lymphoid neogenesis. It has been shown that synovial blood vessels readily acquire the phenotype of HEVs in RA. In all 103 cases of this case series, the synovial vessels surrounded by the lymphocytes histomorphologically showed tall and plump endothelial cells, similar to HEVs in lymph nodes. This was noted in all the papers describing ALVAL. In our study, PNA positivity was seen in those vessels surrounded by lymphocytes, whereas those vessels without the lymphocytic cuffing failed to express PNA. Recently, the morphology of HEVs has been described in detail. The endothelial cells are tall and plump, on Electron Microscopy (EM) have increased mitochondria, free ribosomes and well-developed Golgi apparatus, suggesting an increased metabolic activity. HEVs are surrounded by multiple layers of pericyte-like cells called fibroblastic reticular cells (FRCs). A narrow space between the endothelial basal lamina and the FRC allows the lymphocytes to move easily into the extravascular space. The FRCs also produce and lay down extracellular matrix components including fibronectin, collagen...
Figure 2 (A–C) These photomicrographs (magnification ×200) show metallic debris within the histiocytes. The debris is either needle shaped as seen in A or fine dot-like as in B. Larger deposits of metallic debris are seen without lymphocytic response in C. (D) This photomicrograph (magnification ×200) shows well-defined granulomas. Lymphoid aggregates can also be seen.

IV and laminins forming the thick basal lamina of HEVs. Histologically, HEVs show cuboidal plump endothelium, thick basal lamina and sheath of pericytes. The hyalinisation and onion skinning seen in the smaller vessels in 47/103 cases showing ALVAL can be explained as part of the process of development of HEVs. The transformation to HEVs appears to be progressive, with early development of high endothelial morphology, followed by laying down of thick basal lamina.

When ALVAL was first described, a T lymphocyte cell mediated immunologic response was proposed, but the authors did not elaborate on whether the perivascular lymphocytic response was causing actual vascular damage, although vascular occlusion and fibrinoid necrosis were noted in earlier studies with metal on metal implants. Evans et al noted vascular obliterative changes in periprosthetic tissue around the areas of necrosis and concluded that necrosis was a direct consequence of the vascular changes. Brown et al argued that vascular obliterative changes were commonly seen in the synovium of chronic diseases, and these changes were unlikely to have led to the necrosis.

Figure 3 (A) This photomicrograph (magnification ×200) shows a typical lymphoid aggregate centred on a blood vessel. The vessel shows lymphocytes within its wall and high endothelial venule morphology. There is no hyalinisation of the wall. (B–D) These photomicrographs (magnification ×200) show germinal centre formation. Immunohistochemistry showing organisation into central B cell areas (D) with surrounding T cell areas (C) is seen.
When the term vasculitis is used, one expects to see structural damage to the involved blood vessels with associated functional changes. This is easily seen histologically in immunologic leukocytoclastic vasculitides. However, cutaneous lymphocytic vasculitis is identified by the characteristic perivascular lymphocytic cuffing and changes within the vessel wall are more subtle. Structural and functional changes within the vessel walls are less readily identifiable, and hence there is uncertainty with regard to lymphocytic vasculitis being a true vasculitis. Are the lymphocytes just in transit through the vessel wall or do they cause actual damage? This is an argument that is often discussed in various dermatopathology review articles, giving a range of definitions that define lymphocytic vasculitis. When lymphocytes are seen in the muscular walls of the vessels it is de facto evidence of vasculitis. Diapedesis or transit of lymphocytes does not occur in arteries or veins.

In five cases, we have seen lymphocytes within larger vessel walls in the deeper tissues. The lymphocytic inflammation present was segmental, and did cut out in further tissue levels. In one case, we were able to identify this population to be CD8

---

Figure 4 (A-D) These photomicrographs (magnification ×200) show development of high endothelial venules with plump cuboidal endothelium. Lymphocytes can be seen transiting from the lumen into the wall. Concentric lymphocytic cuffing can be appreciated in B. Occasional plasma cells are also seen. (C, D) High endothelium highlighted by peripheral node addressin (PNAd) immunostain (MECA 79). In C, negative expression of PNAd in vessels not surrounded by the lymphocytic infiltrate is seen.

Figure 5 (A-D) These photomicrographs (magnification ×200) show hyalisation of the vessels surrounded by the lymphocytic inflammation. Note the laminated hyalisation particularly prominent in (A).
Figure 6  (A–D) These photomicrographs (magnification ×200) show lymphocytic infiltrate in vessels larger than arterioles or venules. There is no fibrinoid necrosis, but nuclear dust can be seen in (C). The infiltrate is predominantly lymphocytic with a few eosinophils seen in A and D. Note the lack of perivascular lymphocytic cuffing, indicating this may be the earliest stage in the immune response.

Positive T cells. Evidence of vascular wall destruction was seen in one case, and a further case showed nuclear dust (figure 6A,B,D). In 47 cases, we have seen vascular wall changes in the form of hyalinisation, onion skinning and occlusion in the smaller vessels. These changes could be seen due to one of two reasons: a direct consequence of lymphocyte mediated damage or as a result of transformation to HEVs. In our opinion, the sensitised lymphocytes play a pivotal role in causing necrosis. The vascular hyalinisation could be a reactive response secondary to ‘innocent bystander’ damage caused by the transiting sensitised lymphocytes, the main target of the lymphocytes being the metal debris. This could further contribute to ischaemic pain and necrosis. The vascular changes may not be the only reason for the necrosis, as is evident from our study. All 47 cases with small vessel wall changes showed variable amounts of necrosis. Twenty-five of these showed Grade 4 necrosis. However, the remaining 28/53 cases that showed Grade 4 necrosis did not show any vascular wall changes. Our study, like the earlier studies, does not demonstrate a clear association between the small vessel wall changes and the extent of necrosis. The role

Figure 7  (A–D) These photomicrographs (magnification ×200) show evolving lymphocytic cuffing around larger blood vessels. Again note lack of fibrinoid necrosis of the vessel walls, although luminal obliteration with destruction of the vessel wall can be seen in C. In B, the lymphocytes present in the wall expressed CD8 T lymphocyte markers.
of large vessel involvement is not completely understood at this stage.

Local tissue destruction and loss of function are the hallmark of chronic inflammation.\textsuperscript{25} The term metallosis is used to describe the condition where macrophages containing cytoplasmic metal debris predominate without an ALVAL type chronic lymphocytic response.\textsuperscript{35} Interestingly, none of these 13 cases with pure metallosis without ALVAL showed well-defined granulomas. The presence of cobalt and chromium was confirmed by EDX. Deparaflinised tissue was used, as unstained tissue sections were unsuitable. The metal debris was seen in similar, all cases, and the high cost of EDX precluded us from using it for all our cases. Distinct sarcoid-like naked granulomas were seen in 18 cases, indicating that cell mediated immunity (Type IV hypersensitivity) is seen in response to the metal debris. These granulomas were seen in addition to the perivascular ALVAL response. The granulomas were more commonly seen with total hip stemmed implants, where the only difference is the presence of titanium within the femoral stems as opposed to the resurfacing implants that contain cobalt chromium alloy. The presence of granulomatous inflammation in extensive necrosis can be mistaken for other necrotising granulomatous inflammation, in particular mycobacterial infection. A negative ZN stain and the history of metal implants would help make a diagnosis of ARMD. Mahendra et al observed a predominant macrophagic and T cell response in 45/52 cases and proposed a cytotoxic and cell-mediated response to cobalt and chromium particles.\textsuperscript{12} The presence of granulomas in our study is indicative of Type IV delayed type hypersensitivity response. Host factors may determine why some cases show a purely metallosis type response.

The necrosis seen in ARMD can be caused by the inflammatory mediators generated by macrophages, lymphoid neogenesis and the granulomas present. The joint fluid collection seen in association could be due to defective lymphatic drainage. Defective lymphatic drainage has been proposed to be a trigger for lymphoid neogenesis.\textsuperscript{27} Fluid collection can initiate and perpetuate a cyclical response leading to further pressure necrosis.

The findings described by us are only the beginning of an understanding of the pathogenesis of ARMD seen in metal on metal prostheses. The study has a few shortcomings. The data have not been standardised with regard to age, sex, duration and type of metal on metal implant. We have not been able to explain why some cases showed a pure metallosis type reaction without the development of ALVAL. Patient or host characteristics have an important role to play in determining the nature of the immune response. Earlier revisions (screening cases rather than symptomatic cases) will throw further light on this.

In summary, periprosthetic tissues from metal on metal arthroplasties show necrosis with associated chronic lymphocytic immune response. This response consists of lymphoid neogenesis which is a dynamic process, beginning as perivascular lymphocytic aggregates (ALVAL) leading to lymphoid follicles with germinal centres. These are responsible for perpetuating the immune response to the metal debris, most probably leading to tissue necrosis. Vascular hyalinisation seen within vessels surrounded by the immune response is either secondary to lymphocyte mediated vasculitic damage or could be due to transformation to HEVs. It remains to be seen if these changes contribute further to ischaemic pain or necrosis. Macrophages are an important component of this inflammatory response and in a small number of cases can be the only cell type without any lymphocytic infiltration.

Take-home messages

- Adverse reactions to metal debris consist of a spectrum of changes ranging from a pure metallosis type reaction to aseptic lymphocytic vasculitis associated lesion (ALVAL) and granulomatous inflammation.
- ALVAL is a distinctive response seen in failed metal on metal hip arthroplasties and can be distinguished from chronic inflammatory arthropathies by the presence of extensive surface tissue necrosis and loss of architecture, and metallic debris within macrophages.
- ALVAL begins as perivascular lymphocytic cuffing evolving into lymphoid aggregates with or without germinal centres.
- There does not appear to be a clear association between small vessel wall changes and the extent of necrosis. A small number of cases also showed lymphocytes within larger vessels. It is unclear at this stage if these changes contribute to the degree of necrosis.
- The inflammatory cell response contributes to tissue destruction, but further work needs to be done in elucidating the pathogenesis of necrosis.

Acknowledgements We thank Louise Hall, Jim Frances and Gloria Moorhead, laboratory technicians at the University Hospital of North Tees and Hartlepool NHS Foundation Trust, and the staff in the Chemical and Materials Analysis Unit, University of Newcastle, Newcastle upon Tyne.

Competing interests None.

Contributors SN: corresponding author and pathologist; RPS, JG, DUL: orthopaedic research registrars; AVFN: orthopaedic surgeon.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

Adverse reactions to metal debris: histopathological features of periprosthetic soft tissue reactions seen in association with failed metal on metal hip arthroplasties

Sonali Natu, Raghavendra Prasad Sidaginamale, Jamshid Gandhi, et al.

J Clin Pathol published online March 15, 2012
doi: 10.1136/jclinpath-2011-200398

Updated information and services can be found at:
http://jcp.bmj.com/content/early/2012/03/14/jclinpath-2011-200398.full.html

These include:

References
This article cites 31 articles, 5 of which can be accessed free at:
http://jcp.bmj.com/content/early/2012/03/14/jclinpath-2011-200398.full.html#ref-list-1

PnP
Published online March 15, 2012 in advance of the print journal.

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
### Topic Collections

Articles on similar topics can be found in the following collections

- Immunology (including allergy) (1282 articles)
- Inflammation (131 articles)
- Degenerative joint disease (28 articles)
- Histopathology (88 articles)
- Musculoskeletal syndromes (79 articles)
- Vascularitis (32 articles)

### Notes

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
Chapter Nine
There is widespread concern regarding the incidence of adverse soft-tissue reactions after metal-on-metal (MoM) hip replacement. Recent National Joint Registry data have shown clear differences in the rates of failure of different designs of hip resurfacing. Our aim was to update the failure rates related to metal debris for the Articular Surface Replacement (ASR). A total of 505 of these were implanted.

Kaplan-Meier analysis showed a failure rate of 25% at six years for the ASR resurfacing and of 48.8% for the ASR total hip replacement (THR). Of 257 patients with a minimum follow-up of two years, 67 (26.1%) had a serum cobalt concentration which was greater than 7 µg/l. Co-ordinate measuring machine analysis of revised components showed that all patients suffering adverse tissue reactions in the resurfacing group had abnormal wear of the bearing surfaces. Six THR patients had relatively low rates of articular wear, but were found to have considerable damage at the trunion-taper interface. Our results suggest that wear at the modular junction is an important factor in the development of adverse tissue reactions after implantation of a large-diameter MoM THR.

The results suggested that soft-tissue reactions were more likely to develop in association with accelerated wear of metal prostheses, and that the ASR device was more susceptible to accelerated wear compared with the BHR. These conclusions were in agreement with those from other centres. 

Factors linked to accelerated wear in resurfacing components include a smaller diameter of the bearing, sub-optimal orientation of the acetabular component and reduced acetabular cover. We initially attributed the increased rate of failure observed in patients with the ASR THR to the greater proportion of women in the ASR THR group (male:female, 40:60), believing that the smaller mean size of the bearings led to increased surface wear of the bearing and an increased likelihood of ARMD.

Since the above study was published, the overall failure rate of the ASR bearing has increased significantly. We have observed a disproportionate increase in the failure rate of the ASR THR in men. This phenomenon did not appear to be fully explained by the orientation of the acetabular component, the size of the bearing or the volumetric wear of the bearing surfaces.

Our aims therefore were: 1) to update data on the metal ions in ASR patients; 2) to update the failure rates of the ASR secondary to ARMD; and 3) to investigate the disparity in the failure rate between the THR and resurfacing groups despite the identical bearing surfaces.

**Patients and Methods**

**Implants.** Both the ASR hip resurfacing and THR systems use an identical cobalt (Co) chromium (Cr) acetabular component. The bearing surfaces of the femoral components are also of CoCr and are produced to the same geometrical and material specifications. The THR
head (ASR XL head) attaches to a titanium alloy Corail or SROM stem (both DePuy) through a CoCr taper junction. The tapers are produced in two sizes: 11/13, compatible with the SROM stem, and 12/14, compatible with the Corail stem. Each is available with a varying offset. The tapers are machined to a tolerance of ± 13 μm out-of-roundness (manufacturer's data). They remain fixed to the femoral components during explantation.

Our patients were taken from a single-surgeon prospective study of the ASR bearing surface which was begun in 2004. There were 418 ASR surface replacements and 87 ASR THRs, of which 30 bearings were on Corail stems and 57 on SROM stems. These patients have been described in detail previously.13 They were followed up at six weeks, three months and annually thereafter unless complications developed. The outcome was evaluated using the Harris hip score16 and the University of California, Los Angeles (UCLA) activity score.17

Since the introduction of the Medicines and Healthcare products Regulatory Agency guidelines,18 which recommend that all symptomatic patients with a MoM joint should undergo analysis of blood metal ions, a mass screening programme for all patients with MoM bearings was initiated at our centre. At the time of writing, 409 ASR patients have given samples, of whom 149 have given repeat samples with the aim of carrying out annual tests on all patients and more frequent testing for patients with increased levels.

Failure of the joint secondary to ARMD was recorded. The diagnosis was based on the clinical history, the findings at revision and the histological analysis of excised tissue. Any evidence of sepsis such as positive cultures, grossly increased inflammatory markers in the blood or histological evidence of infection, precluded the diagnosis of ARMD. A raised concentration of Co or Cr in the whole blood or serum was not a prerequisite for the diagnosis since there have been reports of tissue destruction in association with normal surface wear.19

Analysis of explants. All the ASR components which were revised underwent volumetric wear analysis of the bearing surfaces and internal surfaces of the tapers using a coordinate measuring machine (Legex 322; Mitutoyo, Hampshire, United Kingdom) with an accuracy of 0.8 μm. Volumetric and maximum linear wear rates were calculated using Matlab software (MathWorks, Natick, Massachusetts) based on a previously validated programme.20 In order to provide graphical representation of the wear at the taper junctions, the coordinate measuring machine was used to perform several out-of-roundness traces at height intervals of 0.5 μm on the internal surface of the tapers. The latter was also analysed by SEM (FEI XL30 ESEM-FEG; Philips, Eindhoven, The Netherlands) with microanalysis capability (EDX) (QuantaX; Rontec, Carlise, Massachusetts) working in a high vacuum at 25kV.

Statistical analysis. Due to the non-parametric nature of the data, Spearman's rank analysis was used to analyse relationships between variables and Mann-Whitney U tests were used to determine significant differences between groups. Significance was set at $p < 0.05$.

Results

Mid-term CoCr levels. There were 257 patients with unilateral ASR resurfacings and THRs who had a minimum follow-up of two years. Of these, 67 (26.1%) were found to have blood levels of Co or Cr which were greater than 7 μg/l (the figure quoted in the Medicines and Healthcare products Regulatory Agency guidance to guide clinicians in the identification of a poorly performing bearing surface) and 30 (11.7%) had concentrations greater than 20 μg/l. This is a level above which patients were found to have gross macroscopic metallosis at revision in the study of De Smet et al.21 findings which agree with our observations.

Of these 257 patients, the ASR THR patients ($n = 51$) had a significantly higher median concentration of Co in serum and whole blood than the ASR resurfacing patients (n = 206) (serum Co 3.78 μg/l versus 2.55 μg/l, $p = 0.018$, and whole blood Co 3.20 μg/l versus 2.10 μg/l, $p = 0.011$; Mann-Whitney test for non-parametric data).

Bearing diameter and metal ion concentrations. In the resurfacing group, there was a significant trend for ion concentrations to decrease as bearing diameter increased (Spearman's rank correlation = 0.135, $p < 0.001$). This was in contrast to the ASR THR group in which ion levels showed a non-significant increase as bearing diameter increased (Spearman's rank correlation = 0.105, $p = 0.414$). Revisions. The most common finding at revision was a joint effusion in association with varying degrees of soft-tissue necrosis. Gross macroscopic metallosis was often encountered as well as a few solid masses of tissue described as pseudotumours.

ASR resurfacing ARMD failures. At the time of writing there were 37 failures in the ASR resurfacing group. Volumetric wear rates ranged from 2.30 mm³ to 95.5 mm³ per year (Fig. 1). Thus all ASR resurfacing patients who developed ARMD had implants with wear greater than would be expected in pain-free patients with well-functioning prostheses.22 Kaplan-Meier analysis showed survival of 75% at six years for the resurfacing group as a whole (Fig. 1).

ASR THR ARMD failures. There were 25 ASR THR ARMD failures with 14 in the SROM group and seven in the primary ASR THR Corail group. A further four patients developed ARMD after conversion of a primary ASR resurfacings to an ASR THR (Corail) following early fracture. Kaplan-Meier analysis showed a survival of 51.2% at six years for the ASR THR group as a whole (Fig. 2) with 21 of the failures in patients in whom the acetabular components were placed in zones 1, 2 or 321 (Table I).

Bearing diameter and failure rates. As bearing diameter increased, ARMD revision rates decreased in the resurfacing group (Fig. 3). The same pattern was seen in the ASR THR patient group, however, of the four THRs with...
On a random group, the bearing sizes ranged from 1.27 mm³ to 24.08 mm³ per year (Fig. 3). In six patients wear was found to be less than 3 mm³/year. Maximum wear depths measured in the tapers of these six patients were found to be greater than 15 µm (15 to 78) in each case. Therefore in these cases the loss of material was not from the articulating surface but from the taper (Fig. 4). Of the ASR THR ARMD patients, nine were found to have blood and serum concentrations of Co and Cr lower than the threshold level of 7 µg/l suggested by the Medicines and Healthcare products Regulatory Agency.16

The maximum linear wear depths from the internal taper junctions of the femoral components from ARMD patients are shown in Figure 5. A single unused (contaminated during operative procedure) femoral component was available for examination. It was found to have a maximum out-of-roundness of 5 µm and therefore matched the specification provided by the manufacturer. Volumetric loss from the tapers retrieved from ARMD patients varied from 0.07 mm³ to 3.0 mm³.
Figure 4a – an out-of-roundness trace of a retrieved 11/13 taper sleeve from an Articular Surface Replacement XL femoral component. There are two largely undamaged areas which correspond to the anterior and posterior scalloped areas of the SROM stem with which it was coupled. Adjacent to these areas there is marked abnormality of the taper surface with wear depths reaching 40 μm. Figure 4b – Scanning electron microscopy (SEM) image of the areas with no surface change. The manufacturing form has been retained and the elemental composition is as expected. Figure 4c – SEM image of the worn areas. The manufactured form has been lost and there are chromium, phosphate and titanium rich deposits (points 0 and 1) implying corrosion.

**Discussion**

In the light of an increasing number of reports of soft-tissue reactions in peri-prosthetic tissues, there are concerns about the continued use of MoM bearings. Our previous study implied that excessive metal wear was the basic cause of these adverse reactions rather than an idiopathic response to a well-functioning prosthesis. These findings were consistent with those from other centres. While excess wear of the articular surface appeared to explain all of the resurfacing failures at our centre, in this study we have observed a number of patients with relatively well-positioned resurfacing head THR who have experienced ARMD with apparently well-functioning bearing surfaces. In each of these considerable damage was identified at the taper junctions.

Few patients appear to develop ARMD with a well-functioning prosthesis and this is consistent with the idea that ‘sensitive’ or ‘allergic’ patients are the exception rather than the rule. Patients who have undergone THR may be labelled allergic if the bearing surfaces of retrieved explants are examined, found to be normal, and the taper damage is not investigated. Often little or no metal debris is found in the capsular tissue in these cases and a heavy lymphocytic infiltrate with lymphoid neogenesis is present. In these instances an incorrect diagnosis of ‘metal allergy’ can easily be made. At our centre we have analysed a variety of MoM devices from over 100 patients from different centres and we have yet to encounter a patient who has developed ARMD in the absence of abnormal wear of the articular surface or taper junction. It is our belief that the concept of ‘allergy’ in this field of orthopaedics remains unproven and is not a unique condition to be looked upon differently in terms of diagnosis or treatment. Patients who quickly develop a lymphocyte-dominated soft-tissue reaction to a
relatively small concentration of metal debris have similar macroscopic and microscopic tissue appearances to those who develop pain a number of years after surgery and are found to have been exposed to massive concentrations of metal debris. All ARMD patients at our centre who underwent in vitro tests of metal allergy including lymphocyte transformation studies did not have excessive lymphocytic reactivity to Co or Cr ions. Our results agree with those of Kwon et al.

Garbuz et al. recently found that the serum levels of Co and Cr in patients with a large-diameter Durom (Zimmer, Warsaw, Indiana) THR were increased to a much greater extent than in those who received a Durom resurfacing. As in our study, the bearing surfaces of the resurfacing and THR systems were identical. They found a disproportionate increase in the concentrations of Co relative to Cr. In our study, we observed a similar phenomenon, with a significantly higher Co concentration in the ASR THR patients compared with the ASR resurfacing patients. In our ASR resurfacing patients, the median level of Co was found to decrease as the femoral diameter increased. The reverse was true in the THR patients. We speculate that this may have been due to increased mechanical stress on the tapers as the bearing diameter increased.

The catastrophic failure rates of the ASR bearing surface can largely be explained by the design of the acetabular component and its predisposition to edge wear. ASR acetabular components of smaller size are particularly vulnerable to this process because of their reduced arcs of cover. However, smaller acetabular components, when used in a THR system, have an even greater rate of failure secondary to ARMD than resurfacing devices of the same size. The contrast in performance between large ASR resurfacing acetabular components and their stemmed equivalents is even sharper. Seven large acetabular components were implanted in the THR group. They had a diameter of ≥ 60 mm which is the threshold size that we have previously shown to be more resistant to the effects of the position of the acetabular component in terms of bearing surface wear. Three have failed secondary to ARMD within five years. In two of the cases the acetabular components were optimally positioned and the volumetric articular wear rates were relatively low, as were the corresponding ion levels. In both of these cases, however, there was marked taper damage. By contrast, there were 35 equivalent-sized acetabular components used as pure resurfacing devices. All remain in situ except for one which failed because of avascular necrosis. We believe that the generation of metal debris from taper junctions explains the poor performance of the larger sized THR joints and also the increased failure rates of the smaller sizes relative to the pure resurfacings. The out-of-roundness traces show consistent patterns of localized taper damage adjacent to taper areas which have retained their manufacturing form. The patterns of material loss suggest that the tapers have been splayed open by mechanical forces. We speculate that the trend for the use of larger diameter, harder-wearing bearing surfaces without a compensatory change in taper morphology has culminated in the clinical outcomes described in our report. In our total explant collection we have observed severe taper damage with a number of commercially available MoM devices from around the world, including the 36 mm Pinnacle system (DePuy), the Adept (Finsbury/DePuy, Leatherhead, United Kingdom) and the Birmingham (Smith and Nephew) femoral components. We therefore do not believe that this is a problem specific to the ASR. The latest National Joint Registry Annual Report suggests that we are correct in this assumption.

In conclusion, no resurfacing patients in our study who experienced ARMD were found to have a well-functioning bearing surface or metal ion levels which were lower than the median levels for the group as a whole. Six ASR THR patients had failed joints which had relatively little measurable wear from the articulating surfaces. These prostheses were found to have significantly worn taper junctions. We believe that abnormal wear at the head-neck junction may be a major contributing factor to the development of ARMD in MoM THRs using bearing diameters of 36 mm and greater. We advise surgeons to have a high index of suspicion of ARMD in well-positioned MoM THRs even in the absence of elevated levels of metal ions. Asymptomatic patients may have severe soft-tissue destruction, a fact which has been reported in other centres. Surgeons need to be aware of potential taper damage when revising failed MoM joints and the trunion and internal surface of the tapers should be carefully inspected. Consideration should be given to the use of ceramic revision heads which have an internal titanium sleeve to protect the ceramic material.
The author or one or more of the authors have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article.

References
High failure rates with a large-diameter hybrid metal-on-metal total hip replacement

This study reports the mid-term results of a large-bearing hybrid metal-on-metal total hip replacement in 199 hips (185 patients) with a mean follow-up of 62 months (32 to 83).

Two patients died of unrelated causes and 13 were lost to follow-up. In all, 17 hips (8.5%) have undergone revision, and a further 14 are awaiting surgery. All revisions were symptomatic. Of the revision cases, 14 hips showed evidence of adverse reactions to metal debris. The patients revised or awaiting revision had significantly higher whole blood cobalt ion levels (p = 0.001), but no significant difference in acetabular component size or position compared with the unrevised patients. Wear analysis (n = 5) showed increased wear at the trunnion-head interface, normal levels of wear at the articulating surfaces and evidence of corrosion on the surface of the stem.

The cumulative survival rate, with revision for any reason, was 92.4% (95% confidence interval 87.4 to 95.4) at five years. Including those awaiting surgery, the revision rate would be 15.1% with a cumulative survival at five years of 89.6% (95% confidence interval 83.9 to 93.4).

This hybrid metal-on-metal total hip replacement series has shown an unacceptably high rate of failure, with evidence of high wear at the trunnion-head interface and passive corrosion of the stem surface. This raises concerns about the use of large heads on conventional 12/14 tapers.

Modular cobalt-chrome large-diameter femoral heads were introduced in 2003 to treat failure of the femoral component of metal-on-metal hip resurfacing when the acetabular component was well fixed. These had a 12/14 cone which allowed assembly on to an appropriate stem. Midland Medical Technologies (Birmingham, United Kingdom) who made the large diameter modular head did not manufacture a stem at that time, and surgeons used a variety of components from different manufacturers. Large modular heads on a femoral stem were increasingly used as a primary hip replacement instead of hip resurfacing for patients with poor quality of the femoral bone. The combination of a large durable bearing with a clinically proven femoral stem had perceived advantages, including a low rate of dislocation and a greater range of movement and the potential for increased longevity compared with conventional metal-on-polyethylene total hip replacement (THR). These features made it an attractive option for young active patients with degenerative disease of the hip.

The reported survival of metal-on-metal THRs has been favourable, with few failures and complications, although most studies have involved 28 mm diameter heads. The results from smaller series on larger diameter bearings have also been encouraging.

Recent reports of unexpected high failure rates for some types of hip resurfacing and a high incidence of adverse reaction to metal debris (ARMD) have led to an alert from the Medicines and Healthcare products Regulatory Agency prompting an urgent review of all metal-on-metal articulations.

This study reports the mid-term clinical and radiological results of a large-bearing metal-on-metal hybrid THR. Secondary aims were to identify potential sites of failure from retrieval wear analysis and to identify factors predictive of revision.

Patients and Methods

Between 2002 and 2007, 199 metal-on-metal THRs were implanted by the senior author (JML). The author's experience of hip resurfacing in post-menopausal women had been disappointing, owing to early failure of the femoral component, and the metal-on-metal hybrid THR was chosen because it offered the potential benefits of a large durable bearing combined with a femoral stem which had...
previously given excellent results. Other indications for the metal-on-metal hybrid THR included poor-quality proximal femoral bone resulting from conditions such as avascular necrosis of the femoral head, severe cystic degeneration or previous surgery. The cohort contained 110 women and 75 men (14 bilateral cases) with a mean age at surgery of 58.1 years (29 to 77) and a mean follow-up of 62 months (32 to 83).

**Implants.** All patients had a collarless polished tapered cobalt chrome stem (CPT; Zimmer, Warsaw, Indiana) implanted with cement (Palacos R+G; Heraeus Medical GmbH, Hanau, Germany). On the acetabular side, the Birmingham Hip Resurfacing (BHR) acetabular component (Smith & Nephew, Warwick, United Kingdom) coupled with a large modular metal femoral head (Midland Medical Technologies (MMT) Ltd, Birmingham, United Kingdom) was used until 2006, when they were replaced by the Adept resurfacing acetabular component and modular metal heads (Finsbury Orthopaedics, Leatherhead, United Kingdom). The MMT modular head was introduced in isolation without a matching stem. The choice of stem was left to the surgeon’s judgement, as was the practice at the time.15 The senior author initially used the same stem (MS30; Sulzer Orthopaedics, Alton, United Kingdom) as one of the inventors of the BHR. Sulzer Orthopaedics was then acquired by Zimmer. The senior author was already familiar with the CPT stem (Zimmer) and began to use it instead of the MS30 as it was made of chrome-cobalt, had the same 12/14 taper as the MS30 stem and had excellent clinical results.16 The Adept system was used from 2006, after which time the senior author used the uncemented Adept stem with the large diameter modular head. Currently there are no universally accepted standards, guidelines or testing methodologies for trunnions used in hip replacement surgery, and there are no firm guidelines for the accepted tolerances between different taper geometries.

All operations were undertaken through a posterolateral approach in a laminar flow operating theatre. All metal-on-metal hybrid THRs were scheduled to have radiological and clinical review at one, two and five years post-operatively. However, after the Medicines and Healthcare products Regulatory Agency alert in 201014 and concerns over the number of patients presenting with clinical and radiological failure, all patients were recalled for review. The following assessments were performed:

**Clinical analysis.** The Oxford hip score17 was recorded (0 = worst score, 48 = best score). Patients were specifically asked if they had experienced groin, start-up or trochanteric pain, new-onset clunking or clicking, or fatigue (limp after exertion). Those describing new-onset symptoms were categorised as ‘painful hips’.

**Radiological analysis.** Standardised digital anteroposterior (AP) pelvic and lateral hip radiographs were obtained. Assessment of the position of the acetabular component was carried out using Einzel-Bild-Roentgen-Analyse software18,19 (EBRA; University of Innsbruck, Innsbruck, Austria) on all radiographs from each patient. Measurements were performed by an independent assessor (DJL), and mean values were used for analysis. The EBRA version of the acetabular component was calculated as the angle between the true version and the proposed mid-range value (15°; proposed normal range of version 5° to 25°). This established a range of versions from the proposed optimum value. Radiographs were assessed by two authors (BJRFB, JML) for progressive radiolucencies around the acetabular (De Lee and Charnley)20 and femoral (Gruen, McNeice and Amstutz)21 components, osteolysis, bone resorption and component migration.

**Metal ion analysis.** Blood was sampled with a 21-gauge needle (Becton-Dickinson, Oxford, United Kingdom) and collected in trace element tubes containing sodium ethylenediaminetetraacetic acid (EDTA). Samples were measured by inductively coupled plasma mass spectrometry for whole blood cobalt (Co) and chrome (Cr) levels (expressed in nmol/l). Normal ranges were given as 0 nmol/l to 120 nmol/l for Co, and 0 nmol/l to 135 nmol/l for Cr (equivalent to 0 ppb to 7 ppb). Patients with bilateral hip replacements were analysed separately to avoid confounding bias.

**Joint fluid.** In cases of revision joint fluid was aspirated (either before or around the time of revision surgery) and analysed for Co and Cr levels using the same technique and equipment as for blood metal ion sampling.

**Explant analysis.** Revised components underwent volumetric wear analysis of the bearing surfaces using a coordinate measuring machine (Legex 322; Mitutoyo (UK) Ltd, Andover, United Kingdom) with an accuracy of 0.8 μm. Measurements were made every 5° on 18 concentric circles as well as at the pole of the component, thereby completing a total of 4500 to 6000 measurements for each component, depending on the radius of the explant. Rates of volumetric wear were calculated using Matlab (MathWorks, Natick, Massachusetts) using a previously validated program.22 The availability of the RedLux Artificial HipProfiler (RedLux Ltd, Southampton, United Kingdom) has also enabled further surface wear analysis on the most recent explants (n = 4), by scanning the surface of spherical objects and providing a three-dimensional image of the surface with the shape and location (geometric information) of the wear patch as well as information on the wear volume.23 Wear at the taper junction was determined using the coordinate measuring machine to perform several out-of-roundness traces at 0.5 mm height intervals on the internal surface of the tapers. Two unusual tapers were analysed as controls with maximum out-of-roundness of 8 μm and 10 μm, respectively.

**Statistical analysis.** All single variable hypothesis tests were conducted within a non-parametric framework, after confirming that data were not normally distributed. Group comparisons for continuous variables were made using the Mann-Whitney U test. Multifactorial analysis was conducted using logistic regression models with categorical
Table I. Patient demographics, radiological parameters and clinical outcome data for all patients and revised/awaiting revision patients

<table>
<thead>
<tr>
<th></th>
<th>All patients</th>
<th>Revised/awaiting revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hips (patients)</td>
<td>199 (185)</td>
<td>31 (28)</td>
</tr>
<tr>
<td>Mean age in years (range)</td>
<td>58.1 (29 to 77)</td>
<td>62 (49 to 71)</td>
</tr>
<tr>
<td>Number of women (%)</td>
<td>110 (58)</td>
<td>21 (68)</td>
</tr>
<tr>
<td>Mean follow-up in mths (range)</td>
<td>63 (32 to 83)</td>
<td>-</td>
</tr>
</tbody>
</table>

Indication (n, %)
- Primary osteoarthritis 144 (72.4) 20 (64.5)
- Avascular necrosis 22 (11.1) 3 (9.7)
- Rheumatoid arthritis 2 (1.0) 0
- Developmental dysplasia 20 (10.1) 4 (12.9)
- Trauma 8 (4.0) 2 (6.5)
- Others 3 (1.5) 2 (6.5)

Type of surgery (n, %)
- Primary 186 (93.5) 29 (93.5)
- Revision 13 (6.5) 2 (6.5)
- Number of bilateral cases 14 3

Component sizes in mm
- Median femoral head size (range) 46 (38 to 58) 46* (40 to 54)
- Median acetabular component size (range) 52 (44 to 66) 53* (52 to 56)

Clinical outcome
- Mean pre-operative OHS 21 (3 to 39)
- Mean latest OHS 45 (11 to 48) 17* (6 to 31)
- Number of painful MOMTHRs (excluding revisions & awaiting revision, %) 18/168 (11) 23/28* (82)

Radiological results
- Number of patients with radiological changes (%) 31 (16) 24/28* (86)
- Mean EBRA inclination in ° (range) 40.1 (19 to 57) 41.6* (51.5 to 56.0)
- Mean EBRA version in ° (range) 12.7 (-3.1 to 33.1) 12.3* (9.0 to 16.1)

Metal ion results
- Median Co levels in nmol/l (range) 136 (31 to 793) 192* (34 to 650)
- Median Cr levels in nmol/l (range) 63 (8 to 603) 87 (34 to 650)
- Median Mo levels in nmol/l (range) 9 (27) 9* (6 to 27)

* OHS, Oxford hip score; MOMTHR, metal-on-metal hybrid total hip replacement
† bilateral cases excluded
‡ non-adverse reaction to metal debris revisions excluded

Results
Two patients died from causes unrelated to their THR. There were 13 patients who were lost to follow-up. Further patient demographics are outlined in Table I.

A total of 17 hips (8.5%) required revision. The mean time to revision was 45.5 months (18 to 70), at a mean age at operation of 59.8 years (50 to 71). A total of 15 revisions were in women. Of the 17 revisions, 14 were the result of an ARMD, two due to deep infection and one to a peri-prosthetic fracture. The diagnosis of ARMD was made on MRI in nine cases, and clinically with subsequent histological confirmation after exploration of the hip and revision in five. Abundant necrotic caseous material was commonly found around the bone-component or bone-cement interfaces and often tracked anteriorly along the psoas tendon, with extensive bone loss and a spectrum of soft-tissue involvement and peri-implant fluid collections (Fig. 1). A further 14 patients are awaiting revision as indicated by radiological changes (n = 14) and/or pain (n = 9), along with high metal ion levels (n = 10). Details of all revision cases are summarised in Table II.

Survival. The cumulative survival rate for both acetabular and femoral components, with revision for any reason, was 97.0% (95% CI 93.4 to 98.6) at three years and 92.4% (95% CI 87.4 to 95.4) at five years (Fig. 2a), and for ARMD as the cause of revision 96.9% (95% CI 93.3 to 98.6) at three years and 93.6% (95% CI 89.0 to 96.3) at five years.
Analysis by gender showed that the cumulative survival rate in women remaining free of revision for any reason was 94.9% (95% CI 88.9 to 97.7) at three years and 88.1% (95% CI 80.4 to 93.0) at five years, and the corresponding rate for men was 100.0% at three years and 98.6% (95% CI 90.5 to 99.8) at five years (Fig. 2b).

Including those patients awaiting revision (i.e., all failures) as a worst-case scenario, the revision rate for any reason would be 15.1% at final follow-up, with a cumulative survival at five years of 89.6% (95% CI 83.9 to 93.4), and by gender, 12.5% at final follow-up with a cumulative survival at five years of 95.0% (95% CI 85.0 to 98.4) for men and 16.9% at final follow-up with a cumulative survival at five years of 87.8% (95% CI 79.9 to 92.8) for women.

Clinical analysis. All revision cases and nine of the 14 (64%) awaiting revision presented with symptoms. In two revision cases presentation was acute, with individual cases of dislocation and fracture (Vancouver B3 secondary to ARMD: no history of trauma and associated with a large volume of necrotic tissue in the cement-bone interface, with resorption and thinning of the cortex resulting in pathological fracture; Fig. 3b). In those not revised or awaiting revision, 17 patients (18 hips, 9%) had painful hips; 13 of the 17 (76%) were women. Mean pre- and post-operative Oxford hip scores for all patients and revision/awaiting revision cohort are outlined in Table I.

Radiological analysis. A common progressive spectrum of radiological findings was seen which involved initial scalloping and bone resorption of the medial calcar region, with progressive lucency in Gruen zones 7 and 1, along with scalloping and bone resorption in zones 1 and 3 around the acetabular component (Fig. 3). This spectrum of change was observed in ten of the 14 patients diagnosed with ARMD and in all those awaiting revision. In three cases peri-prosthetic fractures were present (two Vancouver AG, one acute fracture Vancouver B3). A further seven asymptomatic patients have non-progressive medial calcar resorption and are being investigated with metal artefact reduction sequence MRI.

Acetabular component size and EBRA analysis. Comparison of revision/awaiting revision versus non-revision cohorts showed no significant difference in acetabular component size (p = 0.77), inclination (p = 0.38) or version (p = 0.12) (Table I).

Metal ion analysis. There was a significant increase in Co levels in the revision/awaiting revision group (p = 0.001) compared with the non-revision cohort, but no significant rise in Cr or Mo metal ion levels (p = 0.14 and p = 0.22, respectively) (Table I).

Retrieval analysis. Black markings and deposits were visible at the trunnion/modular head interface (Fig. 4a). The stem had obvious pitting and evidence of corrosion along the surface which was more marked at the region of the proximal cement/stem interface and the tip of the stem (Fig. 4b).

The mean bearing surface wear between head and acetabular component was 1.86 mm/yr (SD 1.55) (+/- 2 SD). These values, and the geometric information gathered from
Table II. Details of the 17 revision patients (17 hips)

<table>
<thead>
<tr>
<th>Age</th>
<th>Patient (yrs)</th>
<th>Gender</th>
<th>Mths in situ</th>
<th>Indication for primary surgery*</th>
<th>Diagnosis</th>
<th>Establish­ment of diagnosis*</th>
<th>Component size (mm)</th>
<th>Painful pre-revision</th>
<th>Ver­ti­cal Inclina­tion (°)</th>
<th>Intra­operative fluid</th>
<th>Co levels (120 = 7 ppb)</th>
<th>Co levels (135 = 7 ppb)</th>
<th>Co levels</th>
<th>Cr levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60 F</td>
<td></td>
<td>32</td>
<td>OA</td>
<td>ARMD</td>
<td>Clinical</td>
<td>48</td>
<td>Yes</td>
<td>40.5</td>
<td>13.4 Prog</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>63 F</td>
<td></td>
<td>70</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>48</td>
<td>Yes</td>
<td>55.8</td>
<td>11.9 Prog</td>
<td>635</td>
<td>600</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>56 F</td>
<td></td>
<td>34</td>
<td>DDH</td>
<td>ARMD</td>
<td>MRI</td>
<td>48</td>
<td>Yes</td>
<td>49.2</td>
<td>15.2 Static</td>
<td>-</td>
<td>17 800</td>
<td>33 300</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>50 F</td>
<td></td>
<td>62</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>56</td>
<td>Yes</td>
<td>32.8</td>
<td>19.9 Prog</td>
<td>187</td>
<td>112</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>57 M</td>
<td></td>
<td>48</td>
<td>OA</td>
<td>Fracture</td>
<td>Clinical</td>
<td>50</td>
<td>Yes</td>
<td>34.7</td>
<td>3.9 None</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>59 F</td>
<td></td>
<td>57</td>
<td>Trauma (ORIF acetabulum)</td>
<td>Infection</td>
<td>Clinical</td>
<td>54</td>
<td>Yes</td>
<td>46.9</td>
<td>11.3 None</td>
<td>117</td>
<td>55</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>66 F</td>
<td></td>
<td>44</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>56</td>
<td>Yes</td>
<td>37.7</td>
<td>3.1 Static</td>
<td>412</td>
<td>58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>56 F</td>
<td></td>
<td>40</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>52</td>
<td>Yes</td>
<td>42.6</td>
<td>13.7 Prog</td>
<td>116</td>
<td>59</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>62 M</td>
<td></td>
<td>65</td>
<td>AVN (idiopathic)</td>
<td>Fracture</td>
<td>Clinical</td>
<td>58</td>
<td>Yes</td>
<td>41.5</td>
<td>9.6 Prog</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>63 F</td>
<td></td>
<td>68</td>
<td>OA</td>
<td>ARMD</td>
<td>USS</td>
<td>54</td>
<td>Yes</td>
<td>41.0</td>
<td>19.3 Prog</td>
<td>372</td>
<td>87</td>
<td>23 770</td>
<td>3020</td>
</tr>
<tr>
<td>11</td>
<td>64 F</td>
<td></td>
<td>48</td>
<td>OA</td>
<td>ARMD</td>
<td>USS</td>
<td>54</td>
<td>Yes</td>
<td>22.9</td>
<td>21.6 Prog</td>
<td>372</td>
<td>87</td>
<td>23 770</td>
<td>3020</td>
</tr>
<tr>
<td>12</td>
<td>56 F</td>
<td></td>
<td>31</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>50</td>
<td>Yes</td>
<td>35.1</td>
<td>33.1 Prog</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>62 F</td>
<td></td>
<td>52</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>52</td>
<td>Yes</td>
<td>38</td>
<td>12.9 None</td>
<td>171</td>
<td>42</td>
<td>85 229</td>
<td>3962</td>
</tr>
<tr>
<td>14</td>
<td>52 F</td>
<td></td>
<td>39</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>50</td>
<td>Yes</td>
<td>54.2</td>
<td>18.2 None</td>
<td>44</td>
<td>63</td>
<td>1053</td>
<td>1000</td>
</tr>
<tr>
<td>15</td>
<td>50 F</td>
<td></td>
<td>18</td>
<td>DDH</td>
<td>ARMD</td>
<td>MRI</td>
<td>54</td>
<td>Yes</td>
<td>40.1</td>
<td>20 Prog</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>68 F</td>
<td></td>
<td>30</td>
<td>OA</td>
<td>ARMD</td>
<td>MRI</td>
<td>50</td>
<td>Yes</td>
<td>46.9</td>
<td>10.3 Prog</td>
<td>34</td>
<td>88</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>45 F</td>
<td></td>
<td>78</td>
<td>OA</td>
<td>ARMD</td>
<td>Clinical</td>
<td>42</td>
<td>Yes</td>
<td>31.9</td>
<td>15.4 Prog</td>
<td>589</td>
<td>603</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* OA, osteoarthritis; DDH, developmental dysplasia of the hip; ORIF, open reduction and internal fixation; AVN, avascular necrosis

† USS, ultrasound scan

‡ EBRA, Einzel-Bild-Roentgen-Analyse

§ ARMD, adverse reaction to metal debris

Kaplan-Meier survival curves with 95% confidence intervals for a) all patients (revision for any cause) and b) all patients (by gender).

Redlux images, did not show abnormal wear volume, depth or position for the length of time the implants had been in place. The mean maximum out-of-roundness of the taper was 34.5 μm (SD 13.3) (± 2 SD; normal range 8 μm to 10 μm). A characteristic pattern was observed with two discrete regions of wear at polar opposites to each other on the margin of the edge of the trunnion (Fig. 4c). These findings indicate that increased wear at the trunnion/yearner interface and passive corrosion of the stem are the two main sources of metal ion debris.

Radiographs showing a) femoral lysis in zones 1 and 3 and zone 1 acetabular lysis (arrows) corresponding to sites of corrosion seen at retrieval at 75 months follow-up, b) widespread femoral and acetabular lysis (arrows) with associated peri-prosthetic fracture of the femur at 72 months follow-up, and c) extensive acetabular lysis (arrows) in the lateral view at 4 months follow up.

Fluid analysis. The mean metal ion levels in fluid taken at revision cases were 18 635 nmol/l Co, and 35 740 nmol/l Cr. The mean ratio of Cr/Co ions was 6.5:1 in the joint fluid, compared to 1:2 in whole blood.

Multifactorial analysis. The presence of pain, high whole blood Co levels and radiological changes were included in a multiple logistic regression analysis model to determine the strongest predictors of revision or impending revision. The presence of an isolated raised Co level in the absence of either symptoms or radiological changes was not predictive of failure (p = 0.675). However, the presence of pain (p < 0.001) and radiological changes (p < 0.001) in isolation were both significant predictors of failure.

Discussion
Outcomes from the National Joint Registry for England and Wales as well as independent series from centres in the United Kingdom have caused concern about the survival of metal-on-metal articulations. This large metal-on-metal hybrid THR series has demonstrated an unacceptably high level of early failure associated with extensive soft tissue and bony involvement. In the latest joint registry report the revision rate quoted for large-head metal-on-metal hybrid THR was 7.8% (6.6% to 9.3%) at five years. This is comparable to our revision rate of 8.5% (4.5% to 13.4%) at five years. More alarming, however, is the increase in the failure rate that is occurring with time (after two years in women and five years in men, Fig. 2b) both in this series and in the registry.

This high rate of failure might reasonably be thought to be the result of a mismatch between the head and stem taper/trunnion. However, it is becoming increasingly clear that the same pattern of failure is seen in similar devices supplied by a single manufacturer. There are no universally accepted standards relating to the testing of these devices, and as a result there are limited data available on the mechanical behaviour of large diameter modular heads on 12/14 trunnions. The testing undertaken on these devices before introduction to the market clearly did not identify the risk of excessive wear between the head and the stem. There is an opportunity for the engineers and the orthopaedic manufacturing industry to develop an
accepted testing methodology so that new devices can be tested appropriately before being implanted.

The majority of these failures have shown evidence of ARMD (14 of 17 revisions). It is now established that the reaction to metal debris may take several years to develop. This has been highlighted in this series by a mean time to revision of 45.5 months, with only one revision occurring within two years of implantation. It is therefore likely that we are underestimating the true rate of ARMD.

Pain was a prominent feature of failure but, unlike the recent series reported by Donell et al, was often accompanied by abnormal radiographs. The follow-up of these patients prior to the recall was planned at one and five years, and therefore risked not identifying patients until later in the failure process. We have learnt that the symptoms are often subtle and easily overlooked. In two revision cases the patients initially presented with mild lateral trochanteric pain and were treated conservatively for trochanteric bursitis. Later more intrusive symptoms prompted exploration to reveal, in one patient, complete destruction of the abductor insertion and a bald trochanter. Laterally based pain is a potential symptom of early failure and warrants further investigation.

Other established factors associated with early failure in hip resurfacing implants have included female gender, older age, high acetabular component inclination (> 50°) and version, small head size (< 50 mm) and high metal ion levels (Co > 7 ppb or 120 nm/l). It remains to be established whether these risk factors also apply to the large-head metal-on-metal hybrid THRs.

In this series being female was significantly associated with revision alone (p = 0.008), but not when combined with those awaiting revision (p = 0.157). The initial early revisions were predominantly in women, but at the latest review there had been a rise in failure in men. The cause of the high rate of early revision in women is not known, but indicates the multifactorial nature of the failure process. At present it is not possible to determine whether gender is an isolated risk factor in large diameter metal-on-metal hybrid THRs, but this will become clearer as the duration of follow-up increases.

High Co ion levels were also significantly associated with revision/awaiting revision, but unlike the failures of hip resurfacing, age, component positioning and component size were not significant risk factors for impingment or actual failure: 60% of the acetabular components in this cohort were implanted within the proposed ‘safe zone’ (inclination 40° ± 10°, version 15° ± 10°). Furthermore, within this cohort of failures or impending failure with raised Co or Cr levels, 80% of acetabular components were still placed within the safe zone. The use of ‘generic’ safe zones, however, must be interpreted with caution as each individual implant will have a spectrum of tolerance within which it functions optimally. There have been no previously reported series of either the large-head modular Birmingham or Adept hip systems, and therefore the optimum safe zone for these combinations is not known.

Retrieval analysis has identified the trunnion-head interface as a potential source of metal ion debris, with otherwise normal wear volume per year (including depth and position) at the articulating interface. The common wear pattern observed at the trunnions (= shape) suggests a mechanical cause from excess force at the interface. Burroughs et al have previously reported that torsional forces at the trunnion increase as head size increases when metal is tested on standard and highly cross-linked polyethylene. However, the observed pattern with two identical areas of wear at polar opposites around the taper circumference is more suggestive of ‘toggling’, rather than a rotational moment. Furthermore, the lubrication regime that will be integral to the magnitude of these forces will be different for metal-on-metal and metal-on-polyethylene articulations, and will be affected by several factors, including rim contact, impingement, acetabular component deformation, point loading, sliding distance and the duration of the bedding-in phase. These factors, along with the magnitude of early ion production, are therefore likely to be relevant in the currently unknown aetiology of these early failures.

A further source of metal ion production is passive corrosion of the stem surface. Analyses of metal particulate matter from tissues of failed metal-on-metal articulations have shown that Cr (in the form of chromium orthophosphate, a byproduct of corrosion) is the predominant ion. Joint fluid aspirated at revision surgery showed markedly elevated Cr levels compared to Co, a reciprocal finding to the ratio of Co and Cr in the whole blood of the same patients. This is a similar finding to that described by Langton et al and confirms the macroscopic retrieval findings that corrosion has also had a role in the production of metal ions in these failed cases.

What remains unclear is whether mechanical wear at the trunnion or passive corrosion of the stem is the predominant contributor to metal ion production. Elevated levels of metal ions in the blood have been implicated by hip resurfacing studies and retrievals where the articulating surfaces are the proposed primary sources of metal ions. We suggest that the smaller surface area of the trunnion will result in less metal ion production than the larger articulating surfaces. In this series, testing the metal ion levels using a threshold set at 120 nmol/l had a sensitivity of 83% and specificity of 52% for failure. In order to obtain 100% sensitivity the threshold would need to be lowered to 40 nmol/l (equivalent to 2.4 ppb). Multilogistic analysis further showed that an isolated high Co level at this threshold was not significantly predictive of failure compared to the presence of either radiological changes or symptoms (pain). This raises the concern that the threshold level of 7 ppb is too high for metal-on-metal hybrid THRs and may falsely reassure the surgeon that the implant is functioning well.

Without a specific diagnostic test the importance of a complete assessment of these patients including a clinical history, examination and standard plain radiographs, cannot be over-emphasised. Our early experience informs us
that one must intervene as soon as possible in patients with even mild symptoms, to avoid catastrophic complications.

This metal-on-metal hybrid THR series has demonstrated unacceptable high failure rates and a high occurrence of ARMD. Retrieval analysis has highlighted concerns over excess wear at the trunnion along with evidence of corrosion to the stem. Pain is a positive predictor of failure, and new subtle symptoms should not be overlooked. Metal ion levels remain a useful aspect of assessment, but in isolation are not specific or predictive of failure. Further work is necessary to determine the true aetiology of the high failure rates in large-head metal-on-metal hybrid THRs and to establish the mechanical forces and resultant effects on the 12/14 taper. With the increasing popularity of larger femoral heads this series highlights a need to develop international standardised testing regimes and evidence-based guidance for surgeons on the safe and appropriate use of large diameter modular heads on tapers of varying dimensions and geometries.

Supplementary material

A scatter plot showing the acetabular component inclination and version of those hips revised or awaiting revision, with reference to the ‘safe zone’, is available with the electronic version of this article on our website at www.jbjs.org.uk

Listen live

Listen to the abstract of this article at www.jbjs.org.uk/interactive/audio

The authors wish to thank Mrs A. Wakefield for her assistance in the data collection for this paper.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References


Chapter Ten
Taper junction failure in large-diameter metal-on-metal bearings

D. J. Langton, R. Sidaginamale, J. K. Lord, A. V. F. Nargol, T. J. Joyce

From Newcastle University, Newcastle upon Tyne, United Kingdom

Objectives
An ongoing prospective study to investigate failing metal-on-metal hip prostheses was commenced at our centre in 2008. We report on the results of the analysis of the first consecutive 126 failed mated total hip prostheses from a single manufacturer.

Methods
Analysis was carried out using highly accurate coordinate measuring to calculate volumetric and linear rates of the articular bearing surfaces and also the surfaces of the taper junctions. The relationship between taper wear rates and a number of variables, including bearing diameter and orientation of the acetabular component, was investigated.

Results
The measured rates of wear and distribution of material loss from the taper surfaces appeared to show that the primary factor leading to taper failure is the increased lever arm acting on this junction in contemporary large-diameter metal-on-metal hip replacements.

Conclusions
Our analysis suggests that varus stems, laterally engaging taper systems and larger head diameters all contribute to taper failure.

Keywords: Hip, Arthroplasty, Taper, Metal-on-metal, Large diameter, Adverse reaction to metal debris

Introduction
The last four years have seen an increasing number of reports of adverse soft-tissue reactions following metal-on-metal (MoM) arthroplasty of the hip. At first, there appeared to be no correlation between the incidence of these tissue reactions and either MoM total hip replacement (THR) or hip resurfacing. The 2011 report of the National Joint Registry for England and Wales (NJR) published failure rates of 29.0% for the Articular Surface Replacement (ASR; DePuy, Leeds, United Kingdom) THR at six years, in contrast to a 9.6% failure rate for the ASR resurfacing. A clear difference in the performance of the two systems despite identical bearing surfaces. Some observers have stated that the failure of the ASR THR can be linked directly to the flawed design of the ASR acetabular component, leading to higher friction that propagates distally, stressing the modular junction and leading to release of debris from the taper junction. However, joint registry reports from England and Wales and
Australia suggest that this issue is not specific to the ASR, but that there is a deeper underlying problem.6,8 Both registries have shown that MoM THRs in general have not performed as well as conventional bearing surfaces. As shown in the Australian registry,6 this difference in revision rates becomes more emphasised as diameter of the bearing increases. Garbuz et al,9 Langton et al10 and Beaulé et al9 have shown evidence that blood metal ion concentrations are elevated in THRs compared with their resurfacing counterparts in studies involving the Durom (Zimmer, Warsaw, Indiana), ASR and Conserve Plus systems (Wright Medical, Memphis, Tennessee), respectively. A recent prospective study10 in the United Kingdom comparing the Birmingham Hip Resurfacing (BHR; Smith & Nephew, Warwick, United Kingdom) with the BHR THR was terminated due to unacceptably high metal ion levels in patients receiving the THRs.

Retrieval studies have shown evidence of damage at the head–neck modular junction in patients suffering catastrophic tissue damage in association with minimal wear of the bearing surface.8,11 In the past, many authors have proposed that the mechanism of taper failure is primarily electrochemical in nature.10,12 However, previous analyses have been carried out largely using a simple visual quantification of “corrosion”. In this study we aimed to quantify the material loss from tapers, in order to identify risk factors for taper failure and in so doing propose a mechanism that may precipitate failure.

### Materials and Methods

#### Equipment

At our centre, we have carried out a continuing prospective investigation into the failure of MoM hip devices since 2008. All DePuy MoM THRs received at Newcastle University up to September 2011 underwent full volumetric and linear wear assessment of the femoral and acetabular bearing surfaces as well as the articular surfaces of the taper junctions. This was done using a coordinate measuring machine (CMM) (Legex CMM; Mitutoyo, Tokyo, Japan) with an accuracy of 0.8 µm. The technique we used to obtain volumetric measurements of the bearing surfaces of MoM components has previously been published.13,14 We used a custom designed Matlab programme (The Mathworks Inc., Natick, Massachusetts) to analyse the taper surfaces. Using gravimetric analysis as the benchmark measurement technique, our taper analysis was found to be accurate to approximately 0.2 mm³. DePuy components only were included in this analysis. The justification in limiting the research to DePuy products in the first instance was to avoid confounding factors including gross variations in metallurgy, engineering tolerances and material combinations of various stem–head combinations used by different manufacturers.

#### Patients and components

Only components with satisfactory radiographs suitable for analysis of acetabular component orientation using Ein-Bild-Roentgen-Analyse (EBRA) software (University of Innsbruck, Innsbruck, Austria)15 were included as well as those with a known reason for revision, mated femoral stem details and length of time in vivo. There were 111 tapers available for analysis in total. The majority of hips (n = 68, 61%) were obtained from University Hospital of North Tees, where three consultant hip surgeons (CT, RL and an author, AVFN) performed the primary and revision surgeries. The demographics of the patients from whom the components were taken are provided in Table I. In a minority of cases where femoral stems were available, the trunnions also underwent the same analysis. Implants were received from centres in the United Kingdom and United States. The majority of implants (n = 104) were revised secondary to adverse reactions to metal debris (ARMD), the diagnosis of which was based on a combination of clinical history, examination and macroscopic and histological appearance of tissues at revision surgery.2

In this study, the term “trunnion” refers to the intended articular area of the femoral stem (i.e., the “male” component of the modular junction). The term “taper” describes the area attached to the femoral head that is intended to contact the trunnion. It refers to either a sleeve that inserts into the femoral head (as with the ASR XL head system) or simply to the internal cone of the head in systems not using a sleeve adaptor (such as
the Pinnacle system (DePuy) that uses an “Articuleze” femoral head.

All the femoral heads in the study had been used in combination with DePuy stems that were manufactured from titanium (Ti)-aluminium (Al)-vanadium (V) alloy (Ti6Al4V). These stems were Corails and Summits (both DePuy).

As all but one of the implanted Pinnacle devices had a bearing diameter of 36 mm (the exception being 40 mm), the Pinnacle devices were initially analysed separately from the ASRs. The ASR bearing diameters varied in size, with a median of 22.75 mm (39 to 57). In this way two groups were created and analysed as separate entities in order to eliminate potentially confounding variables: 1) the ASR group (ASR bearing surface with Corail or Summit stems; n = 63); and 2) the Pinnacle group (Pinnacle MoM system with Corail or Summit stems; n = 48).

**Visual inspection.** In modular junctions with obvious macroscopic changes there was invariably a localised site of maximal damage where the taper had appeared to engage. The distance of this level of damage from the articular bearing surface we refer to from here on as the taper engagement level (TEL) (Figs 1 and 2). Using simple trigonometry and accounting for the anteroposterior plane only, we calculated the maximal possible lever arm acting on the TEL in the superior inferior direction. Figure 3 shows how this effective horizontal lever arm (HLA) distance was calculated. The lever arm distance is increased by an increasing head offset, increasing bearing diameter and an increasingly varus neck shaft angle.

**Statistical analysis.** The linear and volumetric wear rates of the tapers were examined using the Shapiro-Wilk test for normality and found to be non-parametric. Spearman rank correlation was therefore used to examine the relationships between taper wear rates and a number of variables: taper angle (obtained from unworn areas); bearing diameter; angles of inclination and anteversion of the acetabular component; the distance from the centre of rotation (COR) to the TEL; the HLA distance; volumetric bearing surface wear rates and nominal head offset. The effect of clearance on taper wear was also examined using the same method. A p-value < 0.05 was considered statistically significant.

**Results**

A total of 111 hips were available for analysis. There were 63 ASRs XL heads (12 Summit stems and 51 Corail
stems) and 48 Articuleze heads (47 Corails and one Summit). A total of 38 tapers (34%) exhibited no identifiable surface change on visual inspection. Volumetric and linear wear analysis showed little or no distinction between these tapers when compared with unused, sterile tapers used as control specimens. Wear depths were less than 5 μm in each case. These components were revised secondary to the effects of excessive bearing surface wear, unexplained pain or pain due to loosened femoral/acetabular components. Tables II and III show the differences in tapers found to have obvious TELS/surface change and those without. The majority of retrieved specimens (n = 73) were found to have grossly abnormal taper surfaces. The patterns of surface changes were remarkable in their similarity. Figure 1 shows the macroscopic appearance of a typical sample. Essentially an area of significant damage could be seen (and often even palpated) in a localised circumferential band that corresponded to the insertion of the base of the trunnion (Figs 1 and 2). Proximal to this band (in anatomical terms), the trunnion had left an imprint of its machining grooves (Fig. 2). Scanning electron microscopy (SEM) images confirmed this phenomenon (Fig. 4). CMM wear analysis showed that the most damaged area in visual terms corresponded to the area with the maximal wear depths (Fig. 2). This was identified as the TEL. The areas in which the CMM showed there had been material loss were analysed using the SEM. It appeared that the ridges formed by the trunnion grooves had been flattened somewhat in these regions. Multiple pits had developed. The pits were localised, approximately ten microns in diameter and appeared to be partially filled with inclusion bodies (Fig. 5). Due to the imbedded particulate matter, it appeared likely that the surface changes had occurred primarily due to a mechanical process. Energy dispersive X-ray spectroscopy (EDX) analysis showed that the pits were rich in chromium and the presence of small amounts of chlorides and oxides suggested that corrosion was occurring locally. The surface immediately surrounding these pits showed no changes from the manufactured alloy. There was no difference identified between the patterns of surface change between the ASR and Articuleze specimens.

**The effect of bearing surface wear, clearance and offsets.** Taper linear and volumetric wear rates appeared to be unaffected by variations in clearance or bearing surface wear rates. There was a trend towards increasing taper damage and increasing head offset (Table IV).
The effect of bearing diameter. When the two groups were compared directly, the ASR tapers were found to have significantly greater rates of volumetric and linear wear than the Pinnacle tapers (Table V).

Analysis of the samples as a whole. Figure 6 shows the significant relationship between HLA distance and linear wear rate of the tapers when all samples were included in the analysis (Spearman rank correlation = 0.527, p < 0.001). The same rank correlation using only the Articuleze group (to control for bearing diameter as a variable) was 0.472 (p = 0.002). For the ASR XL group it was 0.416 (p = 0.002).

The effect of orientation of the acetabular component. No significant relationship was identified between cup inclination or anteversion and taper wear (Table VI). This was consistent with the lack of correlation between surface wear and taper wear (Table IV).

Trunnion analysis. There were 11 Corail stems available for analysis. Volumetric analysis proved difficult. This was due to the apparently less tightly controlled manufacturing form of the trunnions (which was confirmed on sterile, unused specimens). Despite this limitation, wear depths appeared to be indistinguishable from manufacturing variation, in that there was no measurable wear over the intended articular area of the trunnions. We were unable to measure the trunnions base in seven of the 11 samples as damage had occurred during extraction. Of the four loose stems that had not suffered damage during explantation, there was no measurable wear. SEM analysis of the trunnions also identified no obvious areas of wear or corrosion although further investigation of the retrieved stems is ongoing.

Discussion
This paper contains an in-depth examination of the modular junction of failed contemporary MoM THRs. It is the largest of its kind in existence. In past research papers
taper junctions have been examined using visual scales. To our knowledge, no accurate quantification of volumetric material loss has previously been published. The results presented in the current paper show that significant volumetric material loss can take place at the modular junctions of modern large diameter THRs. This material loss can exceed that taking place at the bearing surface. The consistent pattern and location of maximal damage on the female taper is consistent with mechanical incompetence.

Is taper failure due to the MoM bearing surface? It is unquestionable that conventional THR is an extremely successful procedure. The 10-year survival of the most common hip prostheses used in Sweden is now over 95%. This is in contrast to the latest published results of large-diameter MoM THR systems that offer a 13.6% revision rate at seven years. The smaller 28 mm Metasul MoM bearing (Zimmer, Warsaw, Indiana) however appears to be functioning relatively successfully in a number of patients at long-term follow-up. It therefore seems unlikely from the evidence that the MoM bearing surface per se is the problem with the latest generation of MoM THRs. The stems associated with failure in this series are, without exception, titanium alloys. The practise of coupling a Ti stem with a CoCr taper has raised concerns of mixed material combinations leading to galvanic corrosion. However, Ti stems have been implanted with CoCr heads for many years with limited reports of gross clinical failures. It is unlikely either, therefore, that it is simply a mixed material coupling that is the root cause of failure. Indeed, previous reports of taper junction failures have included similar metal modular hips.

So why are tapers failing in the 21st century? There have been three obvious changes to the most recent designs of large-head MoM hips that we believe to be of paramount importance.

1. A change in trunnion dimensions. Trunnions have become shorter in length. The trade name of the DePuy taper of the standard group in this study is in fact the "Articuleze Mini Taper" (AMT). This is not a design change specific to DePuy. In fact, most commonly used trunnions are now only 10 mm to 12 mm in length. These changes were brought about to reduce the trunnion’s ‘skirt’ in order to increase the impingement-free range of movement. However, a reduction in length means that the base of the trunnion now often sits inside the taper. This increases the possibility of edge loading of the trunnion base. The subsequent increase in localised stress likely explains the circumferential pattern of surface damage seen on the standard stem taper surfaces in this investigation. Trunnions have also been slimmed down from 16 mm/14 mm diameter cones to the commonly used 14 mm/12 mm cones. A smaller diameter taper means less surface area for a successful interference fit and increases the potential for micromotion.

2. A change in trunnion surface. Most contemporary trunnions now have a ridged surface that has been machined into the material in order to accommodate ceramic heads. This makes financial sense for companies so they can manufacture one stem for multiple bearing surfaces. The machining grooves of the standard trunnion had left an imprint on the failed tapers in this study. We speculate that this plastic deformation of the taper surfaces leads to altered contact stresses and a potential for increased wear. The localised pits we observed may then allow secondary corrosion to take place over a greater surface area.

3. A trend for increasing head size. Larger diameter bearings have the potential to produce less wear and are less likely to dislocate. Understandably, these benefits proved irresistible to industry and surgeons alike. While head sizes increased through the years from 22 mm versions to a standard size male resurfacing femoral component of approximately 52 mm, there was no compensatory change in the morphology of the taper junction. In fact, as head sizes increased, taper junction sizes decreased (as described above). The findings of this retrieval analysis suggest that, while bearing diameter is significant, it is far from the only factor. In the ASR XL group in this study, head size was only weakly associated with taper wear rates. HLA distance proved to be the strongest predictor of taper wear rate in both groups. Consistent with this finding, when bearing diameter was eliminated as a variable (examination of the Articuleze group separately), the effect of an increased HLA distance became even more significant.

Frictional torque. An explanation for the damage observed at the modular junctions of large diameter hip systems could be the increased frictional torque that is generated by the bearing surfaces. This mechanism of failure has support from experimental studies involving metal-on-polymer articulations. However, we can offer no physical evidence of a torsional force from this series of retrieved MoM explants. We have described in our previous work that the thumb printed wear pattern consistently identified on the tapers matched with SROM stems. This pattern of surface change is contrary to the idea that the femoral components are turning on the trunnion. Moreover, in this study the site of maximal taper damage was always found on the opposite side of the head to the location of the wear scar of the head (generally the posteroinferior aspect of the taper). This important finding implies a toggle effect with the mechanical action itself causing damage to the taper surface and opening the junction to the potential corrosive effects of physiological fluids. Finite element analysis studies of the stress distribution on large diameter THRs are consistent with the location of damage identified on the tapers in this study. Another argument against a torque issue is the absence of a significant relationship between bearing surface wear rates and taper wear rates. It is possible, however, that frictional torque initially destabilises the
femoral component and then, combined with the joint reaction force, effectively screws it down the trunnion. This theory might explain why the trunnion bases have appeared to penetrate the taper surfaces so deeply at one level. Once the TEL is firmly established, it may be that a toggle effect then predominates.

**Manufacturing tolerances.** Small variations in the taper and trunnion angles may cause a defined TEL when the head is impacted at primary surgery. Tapers and trunnions are manufactured to a specific tolerance in terms of angles and dimensions. The tolerance band for the taper angle is ±0.069° according to the manufacturer's data. There is evidence from a previous finite element analysis (FEA) that even tiny variations in taper and trunnion angles could be critical. Shareef and Levine stated that 'the magnitude of micromotion...varied between 4 and 26% with increasing values of angle tolerance from zero to 1 minute on the male component'. One minute is equivalent to 0.0167°. The taper angle tolerance stated by DePuy is four times this value (Fig. 7). The same FEA report stated that 'off axis loading caused the female component to tip such that the micromotion on one side is roughly twice that on the other side'. This view is consistent with our findings of taper damage more pronounced on one side than the other. It is therefore likely that taper trunnion angle mismatch is important but something we cannot investigate further until more femoral stems are analysed.

In a perfect situation, the femoral component is impacted on the trunnion. A stable interference fit results, minimising localised stresses, the Morse taper functions adequately and there is no ingress of biological fluids. We propose a chain of events resulting in the damage observed at the taper junctions in this series. If we assume initially that the taper and trunnion angles are significantly mismatched the trunnion engages with the taper at the base of the trunnions. As a result, there is a localised concentration of stress at this level. The TEL may be put under greater stress if it is supporting a larger diameter head, is connected to a varus stem, or is in a laterally positioned (increased head offset). Micromotion ensues, opening up the taper junction to body fluids. Localised corrosion, secondary to removal of material by mechanical (wear) and chemical (corrosive) processes leads to metal ion/particulate release. Further damage at the TEL leads to penetration of the trunnion further into the taper potentiating the effects described above.

**Clinical implications of the mechanism of taper failure.**

As described above, it has been a matter of debate whether it is primarily an increase in frictional torque or an increase in moment arm that is the underlying problem with these devices. The answer to this question has real implications for the present and future of hip arthroplasty. If the primary problem is an increase in moment arm then all hard-on-hard bearing surfaces with taper configurations similar to those described here must be considered at risk of taper failure. In light of recently published literature, we have serious concerns that all large diameter bearings could indeed be affected in a similar way to those described in this paper. As, on the basis of commercial sensitivity, most orthopaedic manufacturers do not divulge their precise implant specifications we cannot comment in this paper on other manufacturers’ MoM devices. What can be stated is that ceramic-on-ceramic (CoC) devices were designed to mediate the TEL, with the aim being to protect the ceramic material from stresses that may cause fracture. The resultant reduction in HLA distance may also confer protection to the taper. Indeed, CoC devices should also be protected somewhat from the effects of frictional torque as the smooth wettable ceramic bearing surfaces have a lower coefficient of friction than MoM bearing surfaces. A ceramic taper could also be expected to be less vulnerable to the effects of wear and corrosion. In spite of this, our advice to surgeons would be to urgently re-evaluate their need to implant large diameter bearing surfaces, irrespective of the bearing surfaces, in these times of uncertainty.

The authors would like to thank Mr C. Tulloc and Mr R. Logishetty, Orthopaedic Surgeons, University Hospital of North Tees.

**References**


Funding statement:
- Funded via a British Orthopaedic Association Joint Action grant awarded in 2008

Author contributions:
- D.J. Langton: Study design, Data collection, Data analysis, Development of measuring techniques, Writing of manuscript
- R. Sigidanamale: Data collection, Data analysis
- A.V.F. Nargol: Data collection, Data analysis, Study design, Writing of manuscript
- J. K. Lord: Development of measurement technique, Data collection
- T.J. Joyce: Study design, Data analysis, Data collection, Development of analytical techniques, Writing of manuscript

ICMJE Conflict of Interest:
- None declared

©2012 British Editorial Society of Bone and Joint Surgery. This is an open-access article distributed under the terms of the Creative Commons Attribution license, which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.
Chapter Eleven
Blood metal ion testing is an effective screening tool to identify poorly performing metal-on-metal bearing surfaces


From University Hospital of North Tees, Stockton, United Kingdom

Objectives
The aims of this piece of work were to: 1) record the background concentrations of blood chromium (Cr) and cobalt (Co) concentrations in a large group of subjects; 2) to compare blood/serum Cr and Co concentrations with retrieved metal-on-metal (MoM) hip resurfacings; 3) to examine the distribution of Co and Cr in the serum and whole blood of patients with MoM hip arthroplasties; and 4) to further understand the partitioning of metal ions between the serum and whole blood fractions.

Methods
A total of 3042 blood samples donated to the local transfusion centre were analysed to record Co and Cr concentrations. Also, 91 hip resurfacing devices from patients who had given pre-revision blood/serum samples for metal ion analysis underwent volumetric wear assessment using a coordinate measuring machine. Linear regression analysis was carried out and receiver operating characteristic curves were constructed to assess the reliability of metal ions to identify abnormally wearing implants. The relationship between serum and whole blood concentrations of Cr and Co in 1048 patients was analysed using Bland-Altman charts. This relationship was further investigated in an in vitro study during which human blood was spiked with trivalent and hexavalent Cr, the serum then separated and the fractions analysed.

Results
Only one patient in the transfusion group was found to have a blood Co > 2 µg/l. Blood/serum Cr and Co concentrations were reliable indicators of abnormal wear. Blood Co appeared to be the most useful clinical test, with a concentration of 4.5 µg/l showing sensitivity and specificity for the detection of abnormal wear of 94% and 95%, respectively. Generated metal ions tended to fill the serum compartment preferentially in vivo and this was replicated in the in vitro study when blood was spiked with trivalent Cr and bivalent Co.

Conclusions
Blood/serum metal ion concentrations are reliable indicators of abnormal wear processes. Important differences exist however between elements and the blood fraction under study. Future guidelines must take these differences into account.

Keywords: Metal-on-metal, Chromium, Cobalt, Arthroplasty, Hip resurfacing, ARMD, Volumetric wear, Metal ions

Article focus
- Documentation of blood metal ion concentrations in a large group of healthy volunteers
- The relationship between serum and whole blood concentrations of chromium (Cr) and cobalt (Co) and volumetric wear of hip resurfacing components
- A description and investigation of the distribution of Cr and Co ions in the serum and whole blood fractions of patients with metal-on-metal (MoM) hip arthroplasty

Key messages
- There is a small variation in metal ion concentrations in the background population relative to the concentrations produced by MoM arthroplasty of the hip
CR and Co concentrations in the serum and whole blood fractions are closely related to the wear rates of explanted prostheses

Co appears to be a more reliable indicator of abnormally wearing prostheses

**Strengths and limitations**

- Most of the results have been drawn from a single area of the United Kingdom. It is unclear whether the results can be extrapolated to areas further afield
- There was a relatively small number of ‘normally wearing’ prostheses that had failed

**Introduction**

Adverse reaction to metal debris (ARMD) is a term used to describe a range of local pathologies seen in association with metal-on-metal (MoM) hips that include soft-tissue necrosis, large sterile joint effusions, metal staining of tissues, pseudotumours and osteolysis. Rates of ARMD-related failure of up to 49% have been reported at six years, but others have reported failure rates as low as 0.10%. Although the extent of soft-tissue lesions does not appear to be dose-related to metal debris exposure, there is accumulating evidence to show that such reactions are far more likely to develop if a MoM prosthesis is wearing at a greater rate than expected. There is currently no convincing evidence in the literature documenting a severe tissue reaction to a well-functioning (in tribological terms) MoM prosthesis.

At present there is no consensus as to what blood concentrations are indicative of an abnormally wearing metal hip. Nor is there consensus as to the interchangeability of serum and whole blood results for the detection of abnormal wear, or even whether chromium (Cr) or cobalt (Co) is the most useful element to measure. This has led to confusion among surgeons and patients. Normal background concentrations of Cr and Co have not even been documented in a large sample size of patients using modern analytical techniques. Furthermore, while the long-term effects of systemic exposure to high concentrations of Cr$^{6+}$ and Co$^{3+}$ metal ions are unknown, Cr$^{6+}$ (hexavalent chromium) is a proven carcinogen, and it is also unknown whether or not MoM joints generate this species of Cr.

The purposes of this study were: 1) to record background concentrations of blood Cr and Co concentrations measured using the latest techniques in a large number of healthy subjects; 2) to compare blood metal ion concentrations with the measured volumetric wear of retrieved prostheses in order to identify levels indicative of poorly functioning hip resurfacings; 3) to examine the distribution of Co and Cr in serum and whole blood of patients with MoM hip arthroplasties; and 4) to further understand the partitioning of Cr ions between the serum and whole blood fractions and investigate the potential release of hexavalent Cr.

**Materials and Methods**

**Background concentrations of Co and Cr in a healthy population.** A study was begun in 2007 to analyse the background environmental exposure to various heavy metals in the North of England, which has reported to the sponsor but not published data as of yet. We obtained the Cr and Co blood results from this study to use as a reference point to compare with those obtained from patients with MoM joints. Blood samples were taken from a random sample of informed volunteers from the National Blood Service (NBS), which is part of the NHS Blood and Transplant (NHSBT) service. The population under study was from the areas of Northumberland, Tyne & Wear, Cumbria and Durham. The total population is approximately 2.5 million people (2001 census data), although only a small proportion of the population donate blood (estimated to be around 4% by the NBS). The study population consisted of subjects aged between 17 and 70 years who had passed the screening health protocols for the NBS. Pregnant women and those with children up to nine months old were excluded, as were transient populations. It was not known whether individuals had metallic implants at the time of venesection. When a blood donation was taken, the first 8 ml to 12 ml were diverted into a sample pouch and then in to two BD Vacutainer Collection Tubes (Beckton Dickson, Franklin Lakes, New Jersey) containing K2EDTA (ethylenediaminetetraacetic acid). Cr and Co analysis was performed using inductively coupled plasma mass spectroscopy (ICPMS) (X-Series II; Thermo Electron Corporation, Bremen, Germany) with Collision Cell Technology (CCT) using 7.5% (v/v) hydrogen (H$_2$) in helium (He) as the collision gas, with in-sample switching between CCT and normal modes. ICPMS is currently accepted to be the preferred mode of blood metal ion measurement. Limits of detection (LoD) were 0.63 μg/l and 0.11 μg/l for Cr and Co, respectively. These analyses were carried out at a centre that participates in the Trace Element Quality Assurance Scheme (TEQAS). This scheme is a collaboration between seven centres in the United Kingdom that perform trace element analysis using the same techniques and regularly monitor results between units to ensure reproducibility. Additionally the laboratory takes part in the Quebec Multielement External Quality Assessment Scheme (QMEQAS). Both include Cr and Co and performance in these schemes is consistently within the acceptable ranges quoted. Ethical approval was sought and given through the National Research Ethics Service (NRES), part of the National Patient Safety Agency (ref: 07/Q0901/20).

**How do metal ion concentrations relate to volumetric wear of retrieved MoM hip resurfacing prostheses?** In 2008 a prospective study was commenced at the School of Mechanical and Systems Engineering at Newcastle University to analyse failed MoM hip prostheses. The study was conducted after approval from the Local
Research Ethics Committee of County Durham and Tees Valley. All mated (head and acetabulum) components retrieved from patients between 2008 and 2012 with failed unilateral hip resurfacings and pre-revision surgery whole blood samples were included in this analysis. Blood samples were taken using an intravenous catheter (Insyte-WTM; Becton Dickinson). After the catheter had been introduced, the metal needle was removed in order to avoid contamination from the needle. A second 5 ml were collected using a Venflon vacuum tube. Blood was placed into EDTA tubes. Concentrations of Cr and Co were determined using ICPMS at the Biochemistry Department, Royal Surrey County Hospital. The quantification limits for both elements were less than 0.2 μg/l and the within assay reproducibility was 2% at a concentration of 8 μg/l. This laboratory demonstrates excellent accuracy in international trace elements external quality assessment schemes and is one of the participants of the United Kingdom TEQAS. Exclusion criteria were: the presence of other metallic implants; loose CoCr backed components; gross immobility at the time of the blood test (University of California, Los Angeles (UCLA) activity score\(^\text{12}\) of 1) and or abnormal renal function tests (blood urea nitrogen and creatinine).

The bearing surfaces of the retrieved components underwent volumetric wear assessment using a Legex 322 coordinate measuring machine (CMM; Mitutoyo, Tokyo, Japan). The CMM has an accuracy of 0.8 μm. Between 4000 and 7000 points were taken from each component and the total material loss in volumetric terms was calculated using a dedicated Matlab program (Mathworks, Natick, Massachusetts) that has been validated previously.\(^\text{13}\)

The relationship between metal ion concentrations and bearing surface wear rates was examined using linear regression after the data had been log normalised. Receiver operating characteristic (ROC) curves were constructed to assess the sensitivity and specificity of different blood fractions and different metals to detect abnormal wear. 'Abnormal wear' was defined for the purposes of this paper after consideration of two factors:

1. Heisel et al\(^\text{14}\) showed that the volumetric loss during the ‘running in’ phase of hip resurfacings is approximately 1.5 mm\(^3\). Thereafter, wear rates are believed to decelerate and it is generally accepted that a well-functioning MoM bearing surface should produce < 1 mm\(^3\) of volumetric wear per year once the steady state is achieved.\(^\text{15}\) Our own analysis of the wear of explanted large-diameter MoM joints showed that this is a reasonable assumption.\(^\text{3}\)

2. In addition to our previous published examination of the accuracy of the method of volumetric wear assessment, we performed a further 30 gravimetric tests on a resurfacing head and acetabular component as we had further enhanced our scanning techniques since the previous publication. This showed that the median error was +0.034 mm\(^3\) per component and the limits of the measurement data were +0.530 to -0.283 (Fig. 1). The limits were calculated in a standard way for non-parametric data as follows: Lower limit of data (maximum under measurement) = quartile (Q) 1 - 1.5(Q3 - Q1); Upper limit (maximum over measurement) = Q3 + 1.5(Q3 - Q1).

Taking into account the above factors, and given the variation in patient mobility (patients with higher activity levels may subject their joints to 6 million cycles per year\(^\text{16}\)), we conducted several ROC analyses defining an ‘abnormal wear’ rate first at a value of 2.0 mm\(^3\)/year for the steady state phase (comfortably above what is generally considered normal) and then 3.0 mm\(^3\)/year (a figure based on our previous work\(^\text{15}\)) in an attempt to accommodate patients with high activity levels. We conducted further analyses in order to determine the effects of maximum over and under estimates of wear to give worst outcome sensitivity and specificity values.

**The relationship between Cr and Co concentrations in serum and whole blood samples of patients with MoM arthroplasty of the hip.** Metal ion analysis of whole blood and serum samples is carried out as part of routine follow-up of patients with MoM hip implants at the University Hospital of North Tees. We collected all available corresponding serum/whole blood samples taken from patients with MoM joints at the hospital up to 2011. Patients had been implanted with Articular Surface Replacement hip resurfacings (ASR; DePuy, Leeds, United Kingdom), Birmingham Hip Resurfacings (BHR; Smith and Nephew, Warwick, United Kingdom), the ASR XL total hip replacement (THR) (DePuy) or the Pinnacle MoM THR system (DePuy). All these devices are manufactured from a similar high-carbon CoCr alloy. Blood samples were collected as previously described. No patients were excluded.

Bland-Altman plots were constructed to assess the agreement between serum and whole blood concentrations of Cr and Co. The calculated difference between serum and whole blood concentrations was plotted for the relationship between Cr and Co concentrations in serum and whole blood samples of patients with MoM arthroplasty of the hip.
Statistical analysis. Windows SPSS v15.0 (SPSS Inc., Chicago, Illinois) was used for statistical analysis throughout and XLSTAT (Addinsoft, New York, New York) was used for graphical representation. A p-value < 0.05 was considered to denote statistical significance.

Results
Background metal ion concentrations. A total of 3042 patients gave samples. There were 1527 males and 1515 females with a mean age of 45 years (16 to 72). The median Cr concentration was 1.5 µg/l (0.6 (below detection limit) to 8.6), and for Co was 0.5 µg/l (0.3 to 6.7). The Shapiro-Wilk test for normality showed that neither Cr nor Co was normally distributed (p = 0.001 and p = 0.002, respectively). A total of 98 patients (3.22%) were found to have blood Cr concentrations > 2 µg/l, whereas only one patient (0.03%) was found to have a blood Co concentration > 2 µg/l. The majority of patients (2831 of 3042, 93.1%) had a Co < 1 µg/l. Figure 2 shows the distribution of Co concentrations. When patients were sub-divided by gender and age, median concentrations of Cr and Co in the various sub groups varied by no more than 0.1 µg/l.

How do metal ion concentrations relate to wear of MoM prostheses? In total, there were 91 retrieved resurfacings with corresponding pre-revision blood Co and Cr values. Of these, 13 patients did not have corresponding serum values, as the patients had been referred from hospitals where the serum fraction was not routinely tested. The indication for revision was ARMD in 82 hips (90.1%), loose titanium-backed acetabular components in two (2.2%), avascular necrosis (AVN) in two (2.2%), pain with an unknown cause in two (2.2%), and one case each of infection, painful impingement and uncomplicated femoral fracture (Table I).

Neither volumetric wear rates nor blood ion concentrations were normally distributed (p < 0.001, Shapiro-Wilk test).

Linear regression using logged values of bearing surface wear rates as the independent variable and logged blood/serum concentrations as the dependent variables returned adjusted R² values of 0.855 for blood Co (p < 0.001) (Fig. 3), 0.756 for blood Cr (p < 0.001), 0.813 for serum Co (p < 0.001) and 0.785 for serum Cr (p < 0.001). The equation of the best-fit line was used to normalise the logged values in order to translate the results into real clinical values. The results of the Co calculations are illustrated in Figure 4 and reported in Table II.

ROC analysis showed whole blood Co to have the most clinically appropriate sensitivity and specificity for the detection of abnormal wear. A whole blood Co concentration ≥ 4.5 µg/l appeared to be the most reliable clinical threshold value after the consideration of the sensitivity and specificity as well as the confidence limits. Whole blood Co ≥ 4.5 µg/l had a sensitivity of 90.4% (95% confidence interval (CI) 81.1 to 95.5) and specificity of 94.4% (95% CI 72.0 to 100.0) for detecting abnormal
Model VOI. 2, clinically sensitivity of 92.8% (95% wear of 'abnormal wear' was defined as that ≥ 3 mm³ per year, clinically the most useful threshold level appeared to be a blood Co concentration of 5.04 µg/l. This value showed a sensitivity of 92.8% (95% CI 83.7 to 97.2) and specificity of 95.5% (95% CI 76.2 to 100.0) (Table IV).

We performed further ROC analyses this time using the maximum possible over and estimates of wear from the wear calculations combined with the maximum and minimum reproducibility error from metal ion analysis. There were only marginal changes in the results. The full data can be seen in Table V.

As wear rates increased, there was a highly significant trend towards a disproportionate increase in serum Cr concentrations compared with the whole blood Cr concentration (Spearman Rank correlation = -0.646, p < 0.001).

The distribution of metal ions in whole blood and serum fractions: in vivo study. A total of 1048 patients gave blood samples. Patient demographics are included.

Table I. Explant patient demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Implants (n)</th>
<th>Mean age (yrs) (range)</th>
<th>Male:female (n)</th>
<th>Mean time to revision (mths) (range)</th>
<th>Device (n, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>91</td>
<td>54 (29 to 68)</td>
<td>38:53</td>
<td>54 (8 to 108)</td>
<td>ASR 73 (80.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BHR 12 (13.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Durom 3 (3.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Conserve Plus 2 (2.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cormet 1 (1.1)</td>
</tr>
<tr>
<td>Reason for revision (n, %)</td>
<td>82 (90.1)</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td>ARMD T</td>
<td>Avascular necrosis</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Loose acetabular component</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Unexplained pain</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Painful impingement</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Infection</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Uncomplicated femoral fracture</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>1 (1.1)</td>
<td>1 (1.1)</td>
</tr>
<tr>
<td>Mean blood concentration (µg/l) (range)</td>
<td>Chromium 14.7 (1.24 to 123.2)</td>
<td>16.0 (0.63 to 271.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cobalt</td>
<td>15.7 (1.9 to 186.7)</td>
<td>17.9 (0.96 to 235.4)</td>
<td>7.35 (0.50 to 138.1)</td>
<td>2 (2.2)</td>
</tr>
<tr>
<td>Mean serum concentration (µg/l) (range)</td>
<td>Chromium 17.9 (1.9 to 186.7)</td>
<td>15.2 (0.96 to 235.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cobalt</td>
<td>17.9 (1.9 to 186.7)</td>
<td>15.2 (0.96 to 235.4)</td>
<td>7.35 (0.50 to 138.1)</td>
<td>2 (2.2)</td>
</tr>
<tr>
<td>Mean volumetric bearing surface wear rate (mm³/year) (range)</td>
<td>Chromium 14.7 (1.24 to 123.2)</td>
<td>16.0 (0.63 to 271.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ASR, Articulating Surface Replacement; BHR, Birmingham Hip Resurfacing
† ARMD, adverse reaction to metal debris
‡ serum concentration values were only available for 78 patients

Graph showing the prediction of wear rates for given whole blood concentrations of cobalt (Co), using the equation derived from the regression model (Fig. 3) with values normalised to provide clinically relevant figures (Table II). The broken lines represent the 95% confidence interval.

Table II. Relevant clinical values of blood cobalt (Co) and their relationship to the volumetric bearing surface wear rate of the metal-on-metal joint (CI, confidence interval)

<table>
<thead>
<tr>
<th>Whole blood Co (µg/l)</th>
<th>95% CI rate of wear (mm³/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.47 to 0.64</td>
</tr>
<tr>
<td>1</td>
<td>0.77 to 1.16</td>
</tr>
<tr>
<td>2</td>
<td>1.26 to 2.10</td>
</tr>
<tr>
<td>3</td>
<td>1.68 to 2.96</td>
</tr>
<tr>
<td>4</td>
<td>2.07 to 3.79</td>
</tr>
<tr>
<td>5</td>
<td>2.43 to 4.58</td>
</tr>
<tr>
<td>10</td>
<td>4.00 to 8.28</td>
</tr>
<tr>
<td>15</td>
<td>5.35 to 11.7</td>
</tr>
<tr>
<td>20</td>
<td>6.58 to 14.9</td>
</tr>
<tr>
<td>30</td>
<td>8.80 to 21.1</td>
</tr>
<tr>
<td>40</td>
<td>10.8 to 27.01</td>
</tr>
<tr>
<td>50</td>
<td>12.7 to 32.7</td>
</tr>
<tr>
<td>100</td>
<td>20.9 to 59.0</td>
</tr>
<tr>
<td>150</td>
<td>28.0 to 83.4</td>
</tr>
<tr>
<td>200</td>
<td>34.4 to 107</td>
</tr>
<tr>
<td>250</td>
<td>40.3 to 129</td>
</tr>
<tr>
<td>300</td>
<td>46.0 to 151</td>
</tr>
</tbody>
</table>

Graph showing the linear regression of blood cobalt (Co) concentration and rates of surface wear (using logged values).

Fig. 3

Fig. 4

The distribution of metal ions in whole blood and serum fractions: in vivo study. A total of 1048 patients gave blood samples. Patient demographics are included.

VOL. 2, No. 5, MAY 2013
The distribution of metal ions in whole blood and serum fractions: in vitro study. Blood samples spiked with Cr$^{3+}$ showed a preferential increase in Cr concentration in the serum fractions. This was in contrast to blood samples spiked with Cr$^{6+}$, where there was a preferential increase in Cr in the RBCs (Fig. 7). Blood samples spiked with Co$^{2+}$ also showed preferential increases in Co concentrations in the serum fraction. The regression line equation for samples stored in EDTA tubes was calculated to be $[\text{Co}^{2+}]_{\text{serum}} = [\text{Co}^{2+}]_{\text{EDTA}}(1.6301 + 0.2323)$, which gave an $R^2$ value of 0.993.

The time from sampling to centrifugation had a very small effect on overall measured Cr concentrations and the distribution of the Cr species between the RBCs and serum, as did the type of tube used to collect the samples.

in Table VI. The Bland-Altman limits of agreements between serum and whole blood samples were -16.5 μg/l to 16.6 μg/l for Co, and -15.4 μg/l to 18.71 μg/l for Cr. The difference between Cr serum and whole blood became larger as Cr concentrations increased. The ratio of serum Cr to whole blood Cr stabilised at approximately 1.6 once an equivalent blood Co concentration of 20 μg/l was reached, the level above which gross metallosis (metal staining of peri-prosthetic tissues) is apparent (Fig. 6).$^2,17$ This relationship was almost identical to that observed in the resurfacing patients, implying there was no difference in the underlying mechanisms of metal ion transport between hip resurfacing and THR patients (Spearman correlation = -0.656, $P < 0.001$).

The graphs showing the sensitivity and specificity of whole blood cobalt (Co) and chromium (Cr) to identify abnormal wear defined as ≥ 2 mm/year (top row) and ≥ 3 mm/year (bottom row).

$^2$
Adverse reactions to metal debris (ARMD) are an increasingly recognised problem. There is a growing body of evidence to suggest that the vast majority of these reactions are seen in association with abnormally wearing MoM devices. There are currently no clear guidelines on how to interpret metal ion results, aside from preliminary recommendations issued by the Medicines and Healthcare Products Regulatory Agency (MHRA). A review on the current state of MoM hips also acknowledged that there is no current cut-off level at which poorly functioning hips can be identified.

There are several reasons why it is desirable to identify abnormally wearing joints:

1. ARMD may be a silent (pain-free) process. It is therefore inappropriate to design studies and formulate guidance where joints are labeled ‘well-functioning’ simply if the patient has no pain. True ‘control’ patients should also have evidence that their prosthesis is functioning well in a tribological sense.

2. Our previous work, which included assessment of the largest cohort of patients who had experienced failure of their joints secondary to ARMD, suggested that accelerated wear leaves patients at greatly increased risk of subsequent development of a catastrophic immune cascade, which can culminate in extensive soft-tissue injury and/or osteolysis. The apparent tolerance of some patients to higher concentrations of metal debris may only be temporary. Notably, in our previous work where we had failed to identify any BHR patients with ARMD in a single-surgeon comparison series, four asymptomatic patients in this group with elevated blood metal ions went on to be revised for gross osteolysis within two years of the study.

3. Normal, expected wear of MoM bearing surfaces is thought to lead to the removal of surface imperfections, generally considered as a self-polishing process occurring in vivo. It is believed that the resulting smoother surfaces can harness a naturally occurring lubricating film, which leads to improved wear characteristics. In contrast, abnormal wear processes lead to roughening of the bearing surfaces and impairment of lubrication. In this way, increased wear leads to further wear. Increased wear leads to greater host exposure to elevated concentrations of metal debris, and the long term systemic sequelae are unknown.

This paper provides data to facilitate the development of guidelines for the interpretation of blood results in order to identify abnormally wearing implants in vivo. It includes comprehensive wear analysis of the largest collection of failed contemporary MoM bearings at an independent centre. This paper also includes analysis of the largest

Table III. Sensitivity and specificity of whole blood cobalt (Co) and chromium (Cr) for identifying abnormal wear defined as a wear rate ≥ 2 mm²/year. The area-under-the-curve (AUC) value for Co was 0.971 (95% confidence interval (CI) 0.965 to 0.977) and 0.959 (95% CI 0.935 to 0.982) for Cr. The most clinically useful values after consideration of sensitivity and specificity are bolded (PPV, positive predictive value; NPV, negative predictive value)

<table>
<thead>
<tr>
<th>Whole blood (µg/l)</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>COBALT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.900</td>
<td>0.973 (0.898 to 0.998)</td>
<td>0.667 (0.435 to 0.837)</td>
<td>0.922</td>
<td>0.857</td>
</tr>
<tr>
<td>2.290</td>
<td>0.959 (0.880 to 0.990)</td>
<td>0.667 (0.435 to 0.837)</td>
<td>0.921</td>
<td>0.800</td>
</tr>
<tr>
<td>2.420</td>
<td>0.959 (0.880 to 0.990)</td>
<td>0.722 (0.487 to 0.876)</td>
<td>0.933</td>
<td>0.813</td>
</tr>
<tr>
<td>2.470</td>
<td>0.959 (0.880 to 0.990)</td>
<td>0.778 (0.541 to 0.913)</td>
<td>0.946</td>
<td>0.824</td>
</tr>
<tr>
<td>2.500</td>
<td>0.945 (0.862 to 0.982)</td>
<td>0.833 (0.598 to 0.948)</td>
<td>0.958</td>
<td>0.789</td>
</tr>
<tr>
<td>2.800</td>
<td>0.945 (0.862 to 0.982)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.972</td>
<td>0.800</td>
</tr>
<tr>
<td>3.210</td>
<td>0.932 (0.845 to 0.973)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.762</td>
</tr>
<tr>
<td>3.260</td>
<td>0.918 (0.828 to 0.964)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.727</td>
</tr>
<tr>
<td>3.540</td>
<td>0.904 (0.811 to 0.955)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.696</td>
</tr>
<tr>
<td>4.500</td>
<td><strong>0.904 (0.811 to 0.955)</strong></td>
<td><strong>0.944 (0.720 to 1.000)</strong></td>
<td>0.985</td>
<td>0.708</td>
</tr>
<tr>
<td>4.700</td>
<td>0.890 (0.795 to 0.945)</td>
<td>0.944 (0.720 to 1.000)</td>
<td>0.985</td>
<td>0.680</td>
</tr>
<tr>
<td>CHROMIUM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.720</td>
<td>0.945 (0.862 to 0.982)</td>
<td>0.722 (0.487 to 0.876)</td>
<td>0.932</td>
<td>0.765</td>
</tr>
<tr>
<td>5.800</td>
<td>0.945 (0.862 to 0.982)</td>
<td>0.778 (0.541 to 0.913)</td>
<td>0.945</td>
<td>0.778</td>
</tr>
<tr>
<td>6.000</td>
<td>0.945 (0.862 to 0.982)</td>
<td>0.833 (0.598 to 0.948)</td>
<td>0.958</td>
<td>0.789</td>
</tr>
<tr>
<td>6.280</td>
<td>0.932 (0.845 to 0.973)</td>
<td>0.833 (0.598 to 0.948)</td>
<td>0.958</td>
<td>0.750</td>
</tr>
<tr>
<td>6.860</td>
<td>0.932 (0.845 to 0.973)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.762</td>
</tr>
<tr>
<td>6.920</td>
<td>0.918 (0.828 to 0.964)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.727</td>
</tr>
<tr>
<td>7.180</td>
<td>0.904 (0.811 to 0.955)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.971</td>
<td>0.696</td>
</tr>
<tr>
<td>7.220</td>
<td>0.890 (0.795 to 0.945)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.970</td>
<td>0.667</td>
</tr>
<tr>
<td>7.630</td>
<td>0.877 (0.779 to 0.935)</td>
<td>0.889 (0.657 to 0.979)</td>
<td>0.970</td>
<td>0.640</td>
</tr>
<tr>
<td>7.850</td>
<td><strong>0.863 (0.763 to 0.925)</strong></td>
<td><strong>0.944 (0.720 to 1.000)</strong></td>
<td>0.984</td>
<td>0.630</td>
</tr>
<tr>
<td>8.250</td>
<td>0.849 (0.747 to 0.915)</td>
<td>0.944 (0.720 to 1.000)</td>
<td>0.984</td>
<td>0.607</td>
</tr>
<tr>
<td>8.370</td>
<td>0.849 (0.747 to 0.915)</td>
<td>1.000 (0.789 to 1.000)</td>
<td>1.000</td>
<td>0.621</td>
</tr>
<tr>
<td>8.420</td>
<td>0.836 (0.732 to 0.904)</td>
<td>1.000 (0.789 to 1.000)</td>
<td>1.000</td>
<td>0.600</td>
</tr>
</tbody>
</table>
Table IV. Sensitivity and specificity of whole blood cobalt (Co) and chromium (Cr) for identifying abnormal wear. The area-under-curve (AUC) value for Co was 0.979 (95% confidence interval CI) 0.968 to 0.989 and 0.958 (95% CI 0.930 to 0.980) for Cr. The most clinically useful values after consideration of sensitivity and specificity are bolded (PPV, positive predictive value; NPV, negative predictive value).

<table>
<thead>
<tr>
<th>Whole blood (μg/l)</th>
<th>Co</th>
<th>Cr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (95% CI)</td>
<td>0.727 (0.515 to 0.870)</td>
<td>0.955 (0.762 to 1.000)</td>
</tr>
<tr>
<td>Specificity (95% CI)</td>
<td>0.917 (0.842)</td>
<td>0.808</td>
</tr>
<tr>
<td>PPV</td>
<td>0.917</td>
<td>0.985</td>
</tr>
<tr>
<td>NPV</td>
<td>0.842</td>
<td>0.808</td>
</tr>
</tbody>
</table>

database of matched serum and whole blood samples in the published literature as well as the largest data set of Cr and Co blood concentrations in a healthy population. **Background blood metal ion concentrations.** The background metal ion data shown in this report is included to provide context to clinicians. The data obtained from these healthy volunteers are not intended to be interpreted as a control group, as they were neither age- nor gender-matched and it was unknown whether these patients had metallic implants in their bodies. Despite this, the results show little variation in metal ion concentrations with an even smaller variation in Co compared with Cr concentrations. We speculate that this may be due to dietary variations or other factors such as smoking. Co has been shown to vary dynamically in response to exercise and it is therefore unsurprising that this element appears to be the more reliable indicator of wear. Analysis of Co is more easily conducted than Cr, and our own review of the literature from the last five years shows blood median ion levels of Co are well-matched irrespective of the laboratory or the country in which studies were performed. The same cannot be said of Cr (Table VII). 

**Use of metal ions to detect abnormal wear.** A whole blood Co concentration ≥ 6.30 μg/l was found to be 93% sensitive and 96.0% specific for detecting abnormal wear (defined at the generous arbitrary value of 3 mm3/year). Serum Co concentrations ≥ 6.30 μg/l also showed high sensitivity and specificity for the detection of abnormal wear. We believe therefore that either serum or whole blood Co can be used for metal ion screening tests. However, the use of whole blood eliminates the need for the separation of serum, an extra step for a lab technician and one in which contaminants can be introduced, and also provides a truer representation of systemic exposure. As our results suggest, wear products appear to fill the serum compartment preferentially (see below). The measurement of whole blood metal ion concentrations therefore 'dilutes' the effect of transient increases in wear.

However, the major limitation in studies of this kind is that the data are drawn from are an inherently biased sample, in that all of the examined prostheses failed early. It would be ideal to examine and compare with a control group, such as those implants that have outlasted their recipient. This could only be done practically with the establishment of an autopsy bank.

**Interchangeability of whole blood and serum.** While serum and whole blood metal ion concentrations showed significant and strong correlation with each other, one fraction could not reliably predict the other; our results substantiate the conclusions of Daniel et al and Smolders et al who concluded that whole blood and serum concentrations cannot be used interchangeably due to the unacceptably wide limits of agreement.
Compared with these two studies, our current study contains a wider range of ion data due to the poor performance of the ASR prosthesis. This allowed us to observe the trend of Cr preferentially storing in the serum compartment as ion concentrations increased in vivo (Fig. 6). This phenomenon prompted us to design an in vitro study to investigate the distribution of Co and Cr ions in serum and whole blood with increasing concentrations of metallic load.

The results of the in vitro study showed that Cr\(^{3+}\) (trivalent Cr) ions had a stronger affinity for serum while Cr\(^{6+}\) (hexavalent Cr) ions showed a preference for RBCs. When blood was spiked with trivalent Cr in vitro, the in vivo distribution of Cr was replicated. This phenomenon provides indirect evidence that it is Cr\(^{3+}\) that is primarily released from MoM implants. These findings are further supported by the in vitro studies by Ordoñez et al.,\(^{34}\) who found Cr\(^{3+}\) to be associated with the serum protein transferrin in vitro, and by Merritt et al.\(^{35}\) who showed that Cr\(^{6+}\) ions generated from corrosion processes are preferentially taken into RBCs rather than serum. These results are also consistent with the in vivo findings of Walter et al.\(^{36}\) The results of the current study prompted us to examine the haematocrit values of the blood samples.
drawn from patients from whom the explants were retrieved. It was found that at lower levels of wear (< 5 mm³/year), haematoctrit concentrations significantly affected the serum and whole blood Cr concentration ratio. A multiple regression model using whole blood Cr:serum Cr as the dependent variable and haematocrit and wear rate as the explanatory variables returned an $R^2$ value of 0.345 ($p = 0.002$, with haematocrit coefficient = 0.480). This implied that at lower levels of wear, Cr concentrations were significantly affected by the pre-operative Cr concentration, which is likely to be stored preferentially in the RBCs. At higher levels of wear this relationship became progressively weaker, implying that the overall increase in Cr produced by the MoM device rendered the pre-operative Cr concentration insignificant. These factors warrant further discussion elsewhere but the findings reinforce the idea that serum and whole blood values are not interchangeable for clinical guidance, and this is particularly so for Cr concentrations. After a review of the literature and consideration of the results presented in this paper we propose a ‘cobalt ladder’ in order to guide the interpretation of blood Co results in mobile patients with a unilateral MoM resurfacing arthroplasty (Table VIII).²,¹⁷,³⁷

The Co values believed to be of concern as stated in this paper are similar to those proposed in the clinical study by Hart et al.,³⁰ in which a blood Co of 4.97 $\mu$g/l was shown to have a sensitivity of 63% and specificity of 86% for a failing MoM joint. How these values relate to the present study can be seen in Figure 5. With regard to serum values, Van Der Straeten et al.¹³ found that an upper value of serum Co of 4.0 $\mu$g/l should be regarded as abnormal for a unilateral resurfacing. Like Hart et al.,³⁰ these authors also found that this value had a much greater specificity (96%) than

Table VII. Previously reported median values of blood chromium (Cr) and cobalt (Co) in cohorts of patients with a Birmingham Hip Resurfacing (BHR) measured using inductively coupled plasma mass spectrometry

<table>
<thead>
<tr>
<th>Authors</th>
<th>Patients (n)</th>
<th>Implant</th>
<th>Laboratory</th>
<th>Mean time to venesection (yrs)</th>
<th>Median Cr (µg/l)</th>
<th>Median Co (µg/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daniel et al²⁸</td>
<td>26</td>
<td>BHR</td>
<td>Laboratory of Government Chemists Laboratories, Teddington, Middlesex, United Kingdom</td>
<td>4</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Langton et al²⁹</td>
<td>70</td>
<td>BHR</td>
<td>Royal Surrey County Hospital, United Kingdom</td>
<td>3.9</td>
<td>3.95</td>
<td>1.43</td>
</tr>
<tr>
<td>Hart et al³⁰</td>
<td>88</td>
<td>82 BHRs 6 Cormets (majority resurfacing)</td>
<td>Imperial Healthcare College, London, United Kingdom</td>
<td>3.5</td>
<td>2.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Holland et al³¹</td>
<td>53</td>
<td>BHR</td>
<td>Imperial Healthcare College, United Kingdom</td>
<td>10</td>
<td>1.74</td>
<td>1.67</td>
</tr>
</tbody>
</table>
sensitivity (22%). Again this is consistent with our data, which showed that this serum Co concentration was 88% specific for detecting abnormal wear.

Conclusions. Serum or whole blood Co and Cr concentrations are reliable indicators of the performance of a MoM bearing surface. There are significant differences between elements and blood fractions that must be taken into account during clinical assessment of patients. We recommend that whole blood Co should be the screening test of choice, and a level of 4.5 μg/l should be regarded as indicative of a poorly functioning joint. We emphasise that these results apply only to hip resurfacing. Total hip replacements have a further metallic interface between the head and the stem, which can lead to significant metal ion generation. The preferential concentration of Cr in the serum compartment likely indicates that trivalent Cr (Cr³⁺) is the predominant species released from MoM hip prostheses (resurfacing or THR), rather than the hexavalent Cr species (Cr⁶⁺).

References


Funding statement:

This study was funded via a grant from the British Orthopaedic Association/Joint Action

Author contributions:

- R. P. Sidaginamale: Writing the paper, Data collection and analysis, Study design
- T. J. Joyce: Data collection and analysis, Study design
- J. K. Lord: Data collection and analysis, Study design
- R. Jefferson: Data collection and analysis, Study design
- P. G. Blain: Data collection and analysis, Study design
- A. V. F. Nargol: Data collection and analysis, Study design
- D. J. Langton: Writing the paper, Overall study concept, Data collection and analysis

ICMJE Conflict of Interest:

None declared

©2013 The British Editorial Society of Bone & Joint Surgery. This is an open-access article distributed under the terms of the Creative Commons Attribution licence, which permits unrestricted use, distribution, and reproduction in any medium, but not for commercial gain, provided the original author and source are credited.
Chapter Twelve
The clinical implications of elevated blood metal ion concentrations in asymptomatic patients with MoM hip resurfacings: a cohort study

David J Langton, Raghavendra P Sidaginamale, Thomas J Joyce, Shonali Natu, Peter Blain, Robert Drysdale Jefferson, Stephen Rushton, Antoni V F Nargol

ABSTRACT

Objective: To determine whether elevated blood cobalt (Co) concentrations are associated with early failure of metal-on-metal (MoM) hip resurfacings secondary to adverse reaction to metal debris (ARMD).

Design: Cohort study.

Setting: Single centre orthopaedic unit.

Participants: Following the identification of complications potentially related to metal wear debris, a blood metal ion screening programme was instigated at our unit in 2007 for all patients with Articular Surface Replacement (ASR) and Birmingham MoM hip resurfacings. Patients were followed annually unless symptoms presented earlier. Symptomatic patients were investigated with ultrasound scan and joint aspiration. The clinical course of all 278 patients with 'no pain' or 'slight/occasional' pain and a Harris Hip Score greater than or equal to 95 at the time of venesection were documented. A retrospective analysis was subsequently conducted using mixed effect modelling to investigate the temporal pattern of blood Co levels in the patients and survival analysis to investigate the potential role of case demographics and blood Co levels as risk factors for subsequent failure secondary to ARMD.

Results: Blood Co concentration was a positive and significant risk factor (z=8.44, p=2×10^{-16}) for joint failure, as was the device, where the Birmingham Hip Resurfacing posed a significantly reduced risk of failure by 89% (z=-3.445, p=0.00005 (95% CI on risk 62 to 97)). Analysis using Cox-proportional hazards models indicated that men had a 66% lower risk of joint failure than women (z=-2.92419, p=0.0218, (95% CI on risk reduction 23 to 89)).

Conclusions: The results suggest that elevated blood metal ion concentrations are associated with early failure of MoM devices secondary to adverse reactions to metal debris. Co concentrations greater than 20 μg/l are frequently associated with metal staining of tissues and the development of osteolysis. Development of soft tissue damage appears to be more complex with females and patients with ASR devices seemingly more at risk when exposed to equivalent doses of metal debris.

ARTICLE SUMMARY

Article focus

- Current Food and Drug Administration guidance for the management of patients with MoM hips states that 'the utility of routine screening of asymptomatic patients using blood metal ion testing has not been established.'

- This study sought to document the clinical course of asymptomatic patients with elevated blood cobalt (Co) concentrations.

Key messages

- Elevated blood Co concentrations were associated with an increased probability of early joint failure secondary to the development of an adverse local tissue response.

- Male patients with low Co concentrations had a very low incidence of adverse tissue reactions and rationalisation of resources with regard to follow-up and repeat blood samples in this group may be considered.

Strengths and limitations of this study

- This study is the first to describe the clinical course of patients following their initial blood metal ion results. It includes three times the number of subjects as the research on which the current Medicines and Healthcare products Regulatory Agency (MHRA) guidance is based.

- The time at which blood samples were taken postoperatively was not standardised.

BACKGROUND

Modern metal-on-metal (MoM) hip resurfacing was developed in order to address the relatively poor survival of conventional prostheses in young, active male patients. Several clinical studies have now shown extremely encouraging survivorship, negligible dislocation rates and excellent functional outcome scores in this patient group. All materials used in hip arthroplasty, however, have their disadvantages. With MoM prostheses, it is the...
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

release of cobalt (Co) and chromium (Cr) metal debris from the components. Adverse reactions to metal debris (ARMD) have captured a lot of media attention recently but fluid collections and soft tissue damage were noted in the presence of increased hip joint fluid Co concentrations as early as 1975. These complications were documented in a report involving patients who had received one of the so-called ‘first generation’ MoM hips, the McKee. Thirty-five years on, we showed a significant increase in the wear rates of explanted fourth generation MoM hip components which had failed secondary to ARMD compared with control specimens. These findings were substantiated by Glyn-Jones et al in their studies on pseudotumour development.

Blood and serum Cr and Co concentrations are reliable indicators of wear rates in hip resurfacing arthroplasty. The relationship between metal ion concentrations and the development of soft tissue necrosis and osteolysis is not straightforward, however. We have previously examined the extent of local tissue damage in relation to blood and hip joint fluid ion concentrations and compared them with the wear rates of the corresponding retrieved components. While no patient was found to have developed a significant soft tissue reaction with a well-functioning bearing surface, in patients who did experience ARMD-related failure, the extent of soft tissue damage did not appear to be dose related to metal debris exposure. The overall results suggested that the observed soft tissue necrosis was the result of the development of a negative immune cascade which varied between patients.

It is quite likely that patients exposed to increased metal debris may show temporary ‘tolerance’ to the stimulus and a certain time period must elapse or a threshold exposure be reached before an immune response is established and symptoms develop. Several publications include descriptions of patients who were initially pain-free but went on to develop pain a number of years later. Furthermore, the increased metal ion levels in asymptomatic patients may be associated with underlying pathology, including osteolysis. It is well recognised that osteolysis can be silent, often only manifesting in the form of radiographic changes or pain secondary to a loosening of components.

These observations challenge recent Medicines and Healthcare products Regulatory Agency (MHRA) guidance, which does not recommend that patients with MoM hip resurfacing arthroplasties should undergo routine blood metal ion testing in the absence of symptoms ‘unless the patient cohort is of concern’. It states that a Cr or Co level greater than 7 μg/l ‘indicates the potential for soft tissue reaction’. But this guidance was primarily based on a cross-sectional study which compared patients with failed MoM hips to control patients with ‘well-functioning’ MoM hips. This study included asymptomatic patients with elevated blood metal ion levels in the ‘well-functioning’ control group. In fact, several patients assigned to the ‘well-functioning’ hip group had blood Co concentrations in excess of 5 μg/l and one was in excess of 50 μg/l. These patients then acted as the control group to determine whether blood metal ion testing was legitimate to identify problems in clinical practice. The postoperative time at which samples were taken was not controlled for, and nor was an attempt made to compensate for time as a confounding factor. Blood metal ion testing provides the surgeon with a clinical assessment of the performance of the prosthetic joint. We believe that the performance of the joint in a tribological sense should be included in the consideration of ‘well-functioning controls’ and some consideration should be given to the effect of exposure to high concentrations of metal debris over time.

To date, no studies thus far have documented the clinical course of asymptomatic patients with elevated blood metal ion concentrations. The aim of this piece of work was to address this deficiency.

PATIENTS AND METHODS

The senior author (AVFN) is a lower limb arthroplasty surgeon with extensive experience in MoM hip resurfacing arthroplasty. AVFN originally used the Birmingham Hip Resurfacing ((BHR), Smith and Nephew, Warwick, UK) from 2002, but in April 2004 he began to use the Articular Surface Replacement (ASR, Deput, Leeds, UK) exclusively. Our previously published work has described in detail the clinical and radiological outcomes of these ASR and BHR patient cohorts as well as identifying the variables leading to excessive metal ion release into the blood. In 2007, after we had observed a number of unusual tissue reactions in patients implanted with the ASR device at our centre, a clinical review and blood metal ion screening programme was started for all patients with MoM resurfacing hips under AVFN’s care. During these review appointments, Harris Hip and UCLA activity scores are documented. Patients are currently reviewed at a maximum of 12 months between appointments and blood tests are repeated.

This paper documents the clinical course of all patients who gave a blood sample between 2007 and 2010 who had no or ‘slight’ pain, no radiological abnormalities and a Harris Hip Score greater than or equal to 95 at the time of venesection. All patients therefore underwent a clinical review at a minimum of 2-year post-venesection. As blood metal ion testing and close follow-up of MoM hip patients have been routine at our centre since 2007, no ethical approval was required for this retrospective audit.

Revision surgery for ARMD

We have described ARMD in detail in our previous work. The diagnosis is made on the combination of clinical history, findings at revision and histological analysis of tissue retrieved at revision surgery. Reactions to metal debris are typically characterised by a combination
of revision findings which include: soft tissue lesions, which can be catastrophic and involve local neurovascular structures; the presence of abnormal (sterile) fluid, which can be copious and pus-like in nature; acetabular and/or femoral osteolysis associated with macroscopic deposition of metal (metallosis). At revision surgery the macroscopic appearance of the tissues was scored by the attending surgeon using a simple grading system\(^{10}\); 0, no soft tissue necrosis; 1, small, localised areas of (mild) tissue necrosis; 2, widespread (moderate) tissue necrosis, stability of the implant not obviously compromised and 3, widespread (severe) tissue necrosis with compromised stability of the implant. The tissues retrieved at revision also underwent a semiquantitative analysis by one consultant histopathologist with extensive experience in this area. For this paper, the cellular response was categorised as ‘histiocytic’, ‘lymphocytic/aseptic lymphocyte dominated vasculitis associated lesion (ALVAL)’ dominant or ‘mixed’ as described in our previous work.\(^{10}\) In order to categorise patients and provide context to the findings at revision, we divided the patients into subgroups according to the results of our work examining the relationship between wear rates of explanted prostheses and pre-revision blood metal ion analysis.\(^{8}\) These subdivisions are described in tables 1 and 2.

### Data analysis

**Event analysis of hip failure**

A statistical model was created to examine the relationship between a patient’s initial recorded blood Co concentration and the risk of the development of joint failure secondary to an adverse reaction to metal debris (ARMD). For the purposes of the prosthetic survival analysis, the endpoints were: the last documented clinical review; the patient undergoing revision surgery prior to March 2012 for any reason other than ARMD; or patient death. Individuals were censored if they had undergone revision surgery for ARMD. Specific inclusion/exclusion criteria are shown in table 3.

We used Cox-proportional hazards models to investigate risk factors for failure secondary to ARMD following the hip resurfacing procedure. The Cox model assumes that there is an underlying unspecified baseline hazard that stays constant through time, that is, influenced by covariates that mitigate or enhance the risk of an event. Our hypothesis was that the size\(^{17}\) and type of the device, patient age,\(^{19}\) sex\(^{19}\) and blood Co levels were associated with an enhanced risk of hip failure. The Cox model assumes that the hazard of the event remains fixed through time, that is, the effect of the individual covariates remains the same and proportional. We investigated non-proportionality in our models by plotting the Schoenfeld residuals and by correlating the residuals for each model with time for each covariate with a twosided \(\chi^2\) test. All models were fitted following the methodology of Therneau and Grambsch.\(^{20}\) We then used the best-fitting model to predict the most likely individual survival curves of MoM devices for men and women at three levels of blood Co (2, 5 and 10 \(\mu g/1\)), to illustrate the association between revision and blood Co.

### Table 1 Patient demographics

<table>
<thead>
<tr>
<th></th>
<th>&lt;1 (low wear)</th>
<th>1–2 (low/expected wear)</th>
<th>2–5 (equivocal wear)</th>
<th>5–10 (increased wear)</th>
<th>&gt;10 (excessive wear)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>57</td>
<td>95</td>
<td>76</td>
<td>25</td>
<td>46</td>
</tr>
<tr>
<td>Number of BHR hips</td>
<td>4</td>
<td>29</td>
<td>12</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Time from op to venesection</td>
<td>49 (44–52)</td>
<td>52 (44–68)</td>
<td>52 (42–68)</td>
<td>55 (52–58)</td>
<td>54 (42–68)</td>
</tr>
<tr>
<td>Number of ASR hips</td>
<td>53</td>
<td>66</td>
<td>64</td>
<td>23</td>
<td>39</td>
</tr>
<tr>
<td>Time from op to venesection</td>
<td>30 (4–67)</td>
<td>28 (7–62)</td>
<td>33 (12–73)</td>
<td>30 (12–67)</td>
<td>34 (12–70)</td>
</tr>
<tr>
<td>Mean age at primary (range)</td>
<td>55 (28–78)</td>
<td>55 (28–73)</td>
<td>57 (39–78)</td>
<td>59 (45–69)</td>
<td>54 (35–68)</td>
</tr>
<tr>
<td>Number of female hips (%)</td>
<td>15 (26.3)</td>
<td>29 (30.5)</td>
<td>26 (33.8)</td>
<td>15 (62.5)</td>
<td>18 (43.9)</td>
</tr>
<tr>
<td>Number (%) of hips revised for ARMD within 9 years of surgery</td>
<td>0</td>
<td>1 (1.1)</td>
<td>4 (5.3)</td>
<td>5 (20.0)</td>
<td>30 (65.2)</td>
</tr>
</tbody>
</table>

The patients have been subdivided according to the classifications described in table 2.

ARMD, adverse reaction to metal debris; ASR, Articular Surface Replacement; BHR, Birmingham Hip Resurfacing.
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

Table 3  Inclusion/exclusion criteria for the two parts of the study

<table>
<thead>
<tr>
<th>Patient inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event analysis of hip failure</td>
<td>No other MoM hip replacements</td>
</tr>
<tr>
<td>ASR or BHR implanted by AVFN</td>
<td>(no patients were excluded for renal function abnormalities)</td>
</tr>
<tr>
<td>Blood Co concentration recorded post hip resurfacing</td>
<td></td>
</tr>
<tr>
<td>Clinical review at a minimum of 2-year postblood test</td>
<td></td>
</tr>
<tr>
<td>Normal basic renal function test at the time of Co blood test (urea and creatinine)</td>
<td></td>
</tr>
<tr>
<td>Harris hip score ≥95</td>
<td></td>
</tr>
</tbody>
</table>

Mixed effects modelling of trends in blood Co

In order to examine the legitimacy of using the single initial blood Co result in the survival analysis, all available repeated Co results from single patients were collected (table 3). As blood samples were taken at different times postoperatively as repeated measures, a generalised linear mixed effect model (GLMM) was used to investigate the extent to which blood Co levels changed in relation to the time since the original hip replacement operation, with patient age and sex, the device used and bearing diameter of the implant as explanatory variables. Our hypothesis was that blood Co levels would increase with time since the operation, would differ between males and females and be independent of patient age. We hypothesised that the size of the implant would also impact on the levels of Co because smaller implants are known to wear (and hence release Co) more readily than larger ones. We adopted a progressive model building strategy following Pinheiro and Bates, first assessing the impact of all variables without the inclusion of random effects and then extending to include random effects and an autoregressive correlation structures as appropriate to allow for any autocorrelation in the response.

RESULTS

There were a total of 299 resurfacings in 278 patients. Of these, there were 246 ASRs and 53 BHR prostheses. Patient demographics can be seen in table 1. The mean (range) postoperative follow-up of the patients was 70 months (12–118). The mean (range) follow-up postvenesection was 36 months (2–63). Two patients died in the study period of unrelated causes.

Mixed effect modelling of trends in blood Co

In total, 205 patients gave repeat samples at a mean (range) of 27.3 months (6–52) following the initial test. Fifty-seven of these patients gave a third sample at a mean (range) of 10.6 months (6–21) following the second test. Blood Co concentrations were skewed to the right with a few patients having very high blood levels at all sampling points. The distribution appeared to be multimodal in both sexes (figure 1). Log transformed levels of blood Co in patients were not related to time since surgery (t=0.3291, p=0.7423); the type of device used (t=−1.225, p=0.2220), patient age (t=−0.1351, p=0.8926) or the size of the device used (t=−0.3084, p=0.7480) in the full model analysing trends through time. There was a suggestion that male sex was associated with reduced Co concentrations; patient sex (−1.967, p=0.0502). Simplification to a model with sex as the only explanatory variable suggested that there was some evidence that men had lower blood Co level than women (t=−2.240, p=0.0259) of analysis of log-transformed blood Co levels. The SD of the random effect for patient factors was 73% of the unexplained variation in the model, indicating that there was a large component of unmeasured variation attributable to the patient. Inclusion of an autoregressive correlation structure in the model gave a
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

marginally better model than one in which it was not included (L.ratio=3.737, p=0.0532).

As the measured Co concentration was not significantly related to time from surgery to blood, it was assumed to be legitimate to conduct survival analysis using the first sample for the patients with time from surgery to most recent follow-up/death/revision used in the analysis.

Revision cases

At the time of the writing, 41 joints had been revised (36 ASRs and 5 BHRs). All but one of these prostheses (an ASR revised secondary to avascular necrosis in a male patient) were revised secondary to ARMD. Only 25 of the ARMD prostheses were revised in female patients and 15 in male patients. Table 4 includes the modes of presentation of these patients.

Event analysis of hip replacement failure

As only two patients died during the course of the study, the impacts of competing risks were assumed to be negligible. Analysis of hip resurfacing failure using Cox-proportional hazards models indicated that men had a 66% lower risk of joint failure than women (z=−2.9419, p=0.00218, 95% CI on risk reduction 23 to 89). Inclusion of time since blood sample as a variable showed that it was not an important risk factor for failure (t=−1.454, p=0.1459). The level of blood Co (as a log transformed variable) was a positive and significant risk factor (z=8.44, p=2x10−16) as was the device, where the BHR posed a significantly reduced risk for revision by 89% (z=−3.445, p=0.00005, 95% CI on risk 62 to 97). An analysis of the Schoenfeld residuals for the model with only the significant variables model showed that the assumption of proportionality of hazard for this model were met (log(Co+1), χ²=8.72, p=0.350; sex, χ²=3.139, p=0.077 and device, χ²=1.591, p=0.207). Risk of failure was highest for females fitted with ASR and a high blood Co level, with the risk of avoiding failure dropping to zero after 10 years, suggesting that hip failure would be inevitable for a person with this set of characteristics. In contrast, failure was least likely for males fitted with the BHR device with low levels of blood Co for which the risk of failure was of the order of 5% over 10 years.

Predicted survival

We used the outputs from the best-fit Cox models identified in the event analysis to generate predicted survival curves for individual patients with certain characteristics. We then used the curves to predict the probability of avoiding revision at 5 and 7 years with associated CIs generated from the regression equations. Predicted survival curves of the replacement for men and women with blood Co concentrations of 2, 5 and 10 μg/l (figures 2–4 and table 5) illustrate the increased risk if the case was a woman, fitted with an ASR device and had elevated blood concentrations of Co. As an example, a woman with an ASR and 10 μg/l of Co in her blood has a 57.3% chance of avoiding revision by 7 year post primary replacement, while a man has a risk of 78.2%. Equivalent probabilities for BHR women and men with the same Co levels at 7 years are 94.2% and 97.4%, respectively. These indicate the increased risk of revision with ASR and high levels of Co.

DISCUSSION

This paper examines the association between elevated blood metal ion concentrations and early failure of hip resurfacing devices due to local ARMD in asymptomatic

<table>
<thead>
<tr>
<th>Table 4 Details of the revision cases in the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signs and symptoms</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Group 1 (n=21)</td>
</tr>
<tr>
<td>Increasing pain and moderate large fluid effusion</td>
</tr>
<tr>
<td>Group 2 (n=11)</td>
</tr>
<tr>
<td>Pain with small fluid effusion</td>
</tr>
<tr>
<td>Group 3 (n=4)</td>
</tr>
<tr>
<td>Acute pain with femoral collapse</td>
</tr>
<tr>
<td>Group 4 (n=4)</td>
</tr>
<tr>
<td>Minor discomfort and grossly elevated ions with mass (n=1)</td>
</tr>
<tr>
<td>Grinding sensation (n=1), gross femoral neck thinning (n=1), squeaking (n=1)</td>
</tr>
<tr>
<td>Avascular necrosis (non ARMD) (n=1)</td>
</tr>
</tbody>
</table>

ARMD, adverse reaction to metal debris; ALVAL, aseptic lymphocyte dominated vasculitis associated lesion; Co, cobalt.

Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

Figure 2  Predicted survival curves for hip replacements for two male and two female hypothetical individuals with different levels of blood cobalt at 2 and 5 µg/l. Device is the Birmingham Hip Resurfacing. Time period is in months. Survival curves shown with 95% CI.

patients. It provides the first evidence that blood metal ion tests can be used as a clinical indicator of the risk of early joint failure in asymptomatic patients.

The main weakness of the paper was the lack of standardisation of the time at which blood samples were taken postoperatively. This was because the work was carried out not as a piece of research but rather one of clinical need as more and more problems began to emerge with the ASR device. To analyse the effect of this variation in time from primary surgery to sampling, a GLMM was constructed. We were able to do this because a large number of patients had undergone repeat sampling at a mean of 2 years. No significant temporal variation was seen, with the implication being that, on average, very low levels tended to stay low and elevated levels remained high.

No preoperative blood samples were taken, although we do not accept that this was a significant limitation. A recent unpublished study involving 3042 healthy volunteers at the nearby regional blood transfusion centre showed that the median (first and third quartiles) blood concentration of Co was 0.5 µg/l (0.4 to 0.6). Only 0.46% of the concentrations were found to be above 1.5 µg/l and in only one sample (0.03%) was the Co concentration above 2 µg/l.

The results of the current study suggest that elevated blood Co levels are a matter of concern, even in asymptomatic patients. Female patients appear to be at greater

Figure 3  Predicted survival curves for hip replacements for two male and two female hypothetical individuals with different levels of blood cobalt of 2 and 5 µg/l. Device is the Articular Surface Replacement. Time period is in months. Survival curves shown with 95% CI.
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

Figure 4 Predicted survival curves for hip replacements for two male and two female hypothetical individuals with blood cobalt concentrations of 10 μg/l. Device is the Birmingham Hip Resurfacing (top two plots) and Articular Surface Replacement (bottom two plots). Time period is in months. Survival curves shown with 95% CI.

Figure 5 The incidence of osteolysis and/or moderate/severe soft tissue destruction at revision surgery. Patients have been grouped according to their preoperative blood result.
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings

Table 5  Predicted probabilities of risk of avoiding revision for patients with different blood Co concentrations 5 and 7 years after initial intervention

<table>
<thead>
<tr>
<th>Cobalt (µg/l)</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Probability</td>
<td>Lower CI</td>
</tr>
<tr>
<td><strong>Probability of avoiding ARMD at 5 years</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99.7</td>
<td>99.3</td>
</tr>
<tr>
<td>5</td>
<td>99.5</td>
<td>98.8</td>
</tr>
<tr>
<td>10</td>
<td>99.1</td>
<td>97.9</td>
</tr>
<tr>
<td>ASR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99.7</td>
<td>99.3</td>
</tr>
<tr>
<td>5</td>
<td>99.5</td>
<td>98.4</td>
</tr>
<tr>
<td>10</td>
<td>87.8</td>
<td>92.4</td>
</tr>
<tr>
<td><strong>Probability of avoiding ARMD at 7 years</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99.2</td>
<td>98.1</td>
</tr>
<tr>
<td>5</td>
<td>98.4</td>
<td>96.5</td>
</tr>
<tr>
<td>10</td>
<td>97.4</td>
<td>94.1</td>
</tr>
<tr>
<td>ASR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>99.2</td>
<td>98.1</td>
</tr>
<tr>
<td>5</td>
<td>98.4</td>
<td>96.5</td>
</tr>
<tr>
<td>10</td>
<td>78.2</td>
<td>67.0</td>
</tr>
</tbody>
</table>

ARMD, adverse reaction to metal debris; ASR, Articular Surface Replacement; BHR, Birmingham Hip Resurfacing.

-bearing surface can more readily stimulate an immune cascade. It is well documented that wear volume and particle morphology can determine the resulting peri- prosthetic immune response. Indeed, it was for this very reason that MoM joints were introduced. It was thought that the overall reduction in volumetric wear rate and the smaller size of particles liberated from MoM prostheses would avoid the initiation of macrophage (histiocyte)-driven osteolysis caused by polyethylene debris. Leslie et al. presented evidence that there is a significant difference in the morphology of ASR and BHR particles, with the median size of ASR particles being approximately half the size of BHR particles. If this is true, it would mean that ASR joints wearing at the same volumetric rate as a BHR joint expose the patient to a far greater number of particles.

When patients are exposed to very high levels of Co (>10 µg/l), the risk of development of osteolysis greatly increases in both ASR and BHR patients. The periprosthetic tissues of patients with Co concentrations in this category were riddled with macroscopically visible metal wear particles. Histological analysis, without exception, showed heavy infiltration of macrophages (histiocytes). This macrophage dominated response (similar to the response to polyethylene wear particles) is consistent with hip simulator data which have shown that particles released under harsh wear testing are much larger than those released under ideal wear conditions. Unfortunately, in this series of patients, only in those with advanced bony defects which required grafting were there any obvious changes on plain radiographs. This is quite likely due to the fact that opaque metal-laden debris is often deposited in the bony defects. We are now performing CT in patients with extremely elevated blood metal ion concentrations in order to better assess bony integrity and aid prerevision planning.

In summary, blood metal ion concentrations are a useful clinical tool, particularly in asymptomatic patients. They can be used as an indicator of the risk of the development of ARMD. Ion concentrations, if low, can be reassuring to both patients and surgeons and can also

Figure 6  Operative findings of a 55-year-old patient from another country. He had minimal discomfort but was not satisfied with his surgeon’s opinion that ‘there was nothing to be concerned about’. His preoperative blood cobalt concentration was 217 µg/l. Note the gross metal staining of the tissues (metallosis) and abnormal fluid. There was extensive acetabular and femoral osteolysis.
allow rationalisation of resources. Grossly elevated ion concentrations indicate the risk of early prosthetic failure and can be used to direct further investigations or implement closer follow-up. At our unit, which acts as a referral centre for the treatment of failed MoM joints, in total 40 patients with blood Co concentrations greater than 20 μg/l have undergone revision of their hip resurfacing so far. All were found to have macroscopic metal staining of the local tissues and 35 were found to have some degree of bone loss. In light of these findings, at our unit patients with grossly elevated metal ion concentrations are now offered revision surgery in the absence of symptoms. Figures 6 and 7 show two such examples. Figure 6 shows the intraoperative finding of a 50-year-old patient with an ASR with ‘slight’ pain only who had been reassured by his primary surgeon and sought a second opinion. Figure 7 shows the retrieved femoral neck of a patient with a BHR who also had minimal symptoms but elected for revision on account of elevated blood metal ion concentrations.

Contributors DL, RS, SN AVFN and SR conceived the idea of the study and were responsible for the design of the study. DL, RS, TJ, SN, PB, AVFN and RJ were responsible for the acquisition of the data, and all authors contributed to the interpretation of the results. The initial draft of the manuscript was prepared by DJ and then circulated repeatedly among all authors for critical revision and approval.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests DL and AVFN have received one-off funding for travel/accommodation reimbursement from Zimmer, Smith and Nephew, Depuy/ Finsbury to attend educational orthopaedic conferences. He also works as an unpaid consultant to Wright Medical. TJ, SN and AVFN is an expert witness in litigation proceedings with regard to failed MoM joints. AVFN worked formerly as a consultant for DePuy.

Ethics approval This study, as stated in the manuscript, was carried out as one of clinical need rather than research.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES


Figure 7 The femoral neck and prosthesis of a male patient in this study with a Birmingham Hip Resurfacing who was found to have blood cobalt of 155 μg/l on routine screening. He had no symptoms but elected for surgery 2 years later when he developed progressive discomfort in his hip. There were no obvious changes on plain x-rays. As can be seen on the right, there was a large cavity in the femoral neck filled with metal-stained, caseous material.
Elevated metal ion concentrations in asymptomatic patients with hip resurfacings


The clinical implications of elevated blood metal ion concentrations in asymptomatic patients with MoM hip resurfacings: a cohort study

David J Langton, Raghavendra P Sidaginamale, Thomas J Joyce, et al.

*BMJ Open* 2013 3:
doi: 10.1136/bmjopen-2012-001541

Updated information and services can be found at:
http://bmjopen.bmj.com/content/3/3/e001541.full.html

These include:

**References**
This article cites 26 articles, 3 of which can be accessed free at:
http://bmjopen.bmj.com/content/3/3/e001541.full.html#ref-list-1

**Open Access**
this is an open-access article distributed under the terms of the creative commons attribution non-commercial license, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non-commercial and is otherwise in compliance with the license. see:
http://creativecommons.org/licenses/by-nc/2.0/ and
http://creativecommons.org/licenses/by-nc/2.0/legalcode.

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

Evidence based practice (84 articles)
Surgery (51 articles)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
Reflective Summary of the Published Works

On reflection, it is obvious that, as time passed, the quality of the research improved. The quality of the earlier publications in this body of work was affected in part due to my inexperience as a researcher but also as a result of the clinical environment. The enclosed publications emerged from an evolving clinical situation rather than a classical, hypothesis driven form of research. An inherent weakness, therefore, of some of the publications is the retrospective nature of the studies. Although the implantation of the ASR was carried out as part of a prospective study, this body of work contains no prospective randomised trials. A prospective randomised trial comparing the ASR to existing conventional forms of arthroplasty and the best performing metal on metal device would have been the gold standard means of assessing the device in terms of functional outcome, metal ion release and rates of ARMD.

Chapter One

In addition to the comments above, “The Influence of Age and Sex on Early Clinical Results After Hip Resurfacing: An Independent Center Analysis” could have been improved specifically by the use of Cox proportional hazards modelling such as was used in the later publication “Adverse reaction to metal debris following hip resurfacing THE INFLUENCE OF COMPONENT TYPE, ORIENTATION AND VOLUMETRIC WEAR”.

Chapter Two

The tests described in “Cup Anteversion in Hip Resurfacing: Validation of EBRA and the Presentation of a Simple Clinical Grading System” were carried out with relatively limited resources. Were this study to be repeated today the fresh frozen cadavers at Newcastle Surgical Training Facility could have been used to more closely replicate the true clinical setting. It would also have been preferable to use a portable coordinate measuring machine to assess acetabular orientation as the gold standard measurement technique.

Chapter Three

“The effect of component size and orientation on the concentrations of metal ions after resurfacing arthroplasty of the hip” reported on the relationships between bearing diameter, cup orientation and metal ions. Ways to improve the methodology of the study would have been to identify the patients to be involved in the study preoperatively. A power analysis would have given an estimate of the number of patients required for each bearing size and a retrospective analysis of cup orientations previously achieved by the surgeon would have facilitated these calculations. Considerations could have been given towards matching patients by weight and sex. Post operatively, blood samples would have been taken at regular follow up periods. Furthermore it may have been preferable to take a number of repeat samples to assess whether there was significant variation in ion concentrations depending on the time of the day or week on which venesection was
performed. Clearly, however, this would have been extremely disruptive to the participants’ lives as well as draining in terms of resources.

The statistical analysis in the existing study could also perhaps have benefitted from log normalisation of the ion data to perform multiple regression tests to investigate the interaction of different variables on ion release.

Chapter Four

“Blood metal ion concentrations after hip resurfacing arthroplasty: A COMPARATIVE STUDY OF ARTICULAR SURFACE REPLACEMENT AND BIRMINGHAM HIP RESURFACING ARTHROPLASTIES.”

This comparative study would have provided the highest quality of evidence if it had been conducted using the same methodology as described above in chapter three, with patients being randomised to receive either the ASR or BHR prosthesis.

Chapter Five

On reflection, in chapter five, the title of the paper “Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement A CONSEQUENCE OF EXCESS WEAR” is somewhat misleading as it implies that causation rather than association has been proved. The intention of the study was primarily to compare groups with high blood metal ion concentrations (and, by extension, abnormally wearing prostheses) to those with low blood ion concentrations in terms of rates of adverse tissue reactions. While the study showed convincing evidence that patients implanted with prostheses wearing at higher wear rates were more likely to develop tissue reactions, it can only strictly be termed an associated rather a causative factor in the development of those reactions.

Stronger evidence could have been obtained by means of a randomised prospective trial. Patient assessments could have been carried out by an experienced independent surgeon, unaware of the device which had been implanted and blinded to post-operative blood metal ion results. Confounding effects of patients’ subjective symptomatology could have been eliminated by the use of widespread screening ultrasound and/or MRI scans.

Chapter Six

“Volumetric wear assessment of failed metal-on-metal hip resurfacing prostheses.”

“Reducing Metal Ion Release Following Hip Resurfacing Arthroplasty.”

The studies involving explant analysis which are included in this thesis were conducted prospectively. However, explant studies introduce an issue of confounding by their very nature. It is hoped that the drawbacks of these types of studies have been expressed clearly in the published papers. The main limitations are brought about by the low number of well-functioning hips from a tribological point of view. This is essentially unavoidable given the association between accelerated wear and early failure.

Chapter Seven
"Reducing Metal Ion Release Following Hip Resurfacing Arthroplasty."

This study suffers from similar limitations, and thus could have been optimised using the same methodology, as described in the paper "Blood metal ion concentrations after hip resurfacing arthroplasty: A COMPARATIVE STUDY OF ARTICULAR SURFACE REPLACEMENT AND BIRMINGHAM HIP RESURFACING ARTHROPLASTIES".

Chapter Eight

"Adverse reaction to metal debris following hip resurfacing: THE INFLUENCE OF COMPONENT TYPE, ORIENTATION AND VOLUMETRIC WEAR."

The quality of evidence generated from this study could have been optimised by the collaboration of the surgeons from the outset. The surgeons would have followed a set protocol of indications to perform the procedure, with patients randomised to receive an ASR, a BHR or a Conserve Plus prosthesis. It would have been preferable for each surgeon to have an equal amount of experience in the use of each device in order to eliminate confounding effects of the learning curve. Patients would have undergone the same follow up procedures and blood samples drawn at the same intervals post operatively. Patient review and clinical assessment would have been carried out by medical staff without knowledge of the device implanted or the surgeon who had carried out the procedure. In order to screen for ARMD, regular ultrasound and MRI scans would have been carried at regular follow up periods, irrespective of symptomatology.

Chapter Nine

"Accelerating failure rate of the ASR total hip replacement."

"High failure rates with a large-diameter hybrid metal-on-metal total hip replacement: CLINICAL, RADIOLOGICAL AND RETRIEVAL ANALYSIS."

The studies described in this chapter could also have been improved by prospective randomised controlled studies as detailed in the reflective summary of chapter eight.

Chapter Ten

"Taper junction failure in large-diameter metal-on-metal bearings."

This study, as in those described previously, has the unavoidable limitation of the skewed nature of the sample set. However, that does not mean that the taper junctions which were analysed had all performed poorly. Indeed, a significant proportion of the THR constructs in this study exhibited no significant changes at this interface.

Chapter Eleven

"Blood metal ion testing is an effective screening tool to identify poorly performing metal on metal bearing surfaces."

Without the constraints of ethics, the ideal way to conduct this study would have been to enrol an equal number of male and female patients who were deemed suitable for a resurfacing procedure. At a set time point, for example two years post operatively, the patients would undergo removal of
their devices in order to conduct full volumetric wear analysis with intention to compare these results to pre revision blood ion concentrations. Clearly this idealised situation will never be acceptable from an ethical point of view. For this reason, studies of this kind will always suffer the limitation of relatively small numbers of low wearing bearing surfaces available for examination.

Chapter Twelve

“The clinical implications of elevated blood metal ion concentrations in asymptomatic patients with MoM hip resurfacings: a cohort study.”

Ideally, this study would have reported on the longitudinal results of the prospective randomised study detailed in the reflections in chapters four and five. Repeated regular cross sectional scans would have provided valuable information on the natural history of fluid collections and masses associated with MoM hips.