Analysis of wear of retrieved metal-on-metal hip resurfacing implants revised due to pseudotumours


From the Nuffield Orthopaedic Centre, Oxford, England

The presence of pseudotumours, which are soft-tissue masses relating to the hip, after metal-on-metal hip resurfacing arthroplasty has been associated with elevated levels of metal ions in serum, suggesting that pseudotumours occur when there is increased wear. We aimed to quantify the wear in vivo of implants revised for pseudotumours (eight) and of a control group of implants (22) revised for other reasons of failure.

We found that the implant group with pseudotumours had a significantly higher rate of median linear wear of the femoral component at 8.1 μm/year (2.75 to 25.4) than the 1.79 μm/year (0.82 to 4.15; p = 0.002) of the non-pseudotumour group. For the acetabular component a significantly higher rate of median linear wear of 7.36 μm/year (1.61 to 24.9) was observed in the pseudotumour group compared with 1.28 μm/year (0.81 to 3.33; p = 0.001) in the other group. Wear of the acetabular component in the pseudotumour group always involved the edge of the implant, indicating that edge-loading had occurred.

Our findings are the first direct evidence that pseudotumour is associated with increased wear at the metal-on-metal articulation. Furthermore, edge-loading with the loss of fluid-film lubrication may be an important mechanism of generation of wear in patients with a pseudotumour.

With the introduction of the metal-on-metal (MoM) bearing in hip resurfacing, there has been a rapid increase in the number of such procedures performed. These bearings have better wear properties than conventional metal-on-polyethylene implants. This has the potential to reduce substantially wear-induced osteolysis as the major cause of failure. Other proposed advantages of MoM hip resurfacing arthroplasty over conventional total hip replacement (THR) include bone conservation, greater stability of the implant and assumed easier revision. The data from the National Joint Registry of England and Wales and the Australian National Joint Registry have indicated that hip resurfacing procedures account for 35% and 26%, respectively, of all primary hip arthroplasties performed in patients younger than 55 years of age. However, there is a growing concern regarding the occurrence of periprosthetic soft-tissue lesions in patients who have had this hip resurfacing arthroplasty. These cause considerable symptoms and require revision operations in a high proportion of patients, the outcome of which is poor. Since these lesions are often difficult to distinguish morphologically from necrotic tumours, the term inflammatory 'pseudotumour' has been used to describe them. They have also been termed bursae, cysts, or inflammatory masses.

The presence of soft-tissue pseudotumours in patients with MoM hip resurfacing arthroplasty has been associated with elevated levels in the serum of cobalt (Co) and chromium (Cr) ions, the principal elements in the CoCr alloy used in these implants. The concentrations of Co and Cr ions in the blood have been considered to be potential surrogate markers of the wear of MoM-bearing, since it is not possible to measure the wear of bearing couples radiologically in these implants. This assumption is supported by recent measurements of wear in retrieved MoM prostheses. Elevated concentrations of serum metal ion were significantly correlated with greater linear wear of the femoral component. Since wear is positively correlated with the elevation of concentrations of metal ions in vivo, this suggests that pseudotumours in patients with contemporary MoM resurfacing implants are associated with increased wear at the articulation. Furthermore, edge-loading, a phenomenon whereby the femoral component comes into contact with the edge of the acetabular component, has recently been suggested to be a
possible mechanism which leads to increased wear in these implants. However, the direct assessment of wear and the occurrence of edge-loading have not been previously investigated in patients with pseudotumours. Our aims were to quantify the wear of MoM hip resurfacing arthroplasty implants revised for pseudotumours using direct measurement, to compare this measurement with that for a control group of these implants revised for other reasons of failure, and also to establish whether the edge-loading phenomenon occurred in implants revised for pseudotumours.

### Materials and Methods

**Selection of retrieved implants.** We investigated two groups of retrieved implants. Group 1 consisted of implants retrieved from patients who had undergone revision because of the presence of a pseudotumour. Group 2 contained implants retrieved from patients who had undergone revision because of other causes of failure. The study had to be performed on representative samples in each implant group to ensure a valid generalisation of the findings. This was particularly important in selecting implants for the pseudotumour group, because the presence of a pseudotumour is not a well-established indication for revision. In order to minimise selection bias, the retrieved implants were selected using the following criteria. We studied the MoM hip resurfacing arthroplasty implants of a contemporary design which were currently in use in order to reflect the design of implants in patients with pseudotumours in whom the elevated concentrations of metal ions had been measured. The diagnosis of pseudotumour was confirmed according to the following requirements. First, the presence of a pseudotumour should have been documented pre-operatively in the revised resurfaced hip. Secondly, the stated indication for revision was solely the presence of pseudotumour and thirdly, the final diagnosis of the histopathological report of the tissue specimen by the musculoskeletal pathologist was a pseudotumour, in accordance with the published histological features for this entity.

The non-pseudotumour group was limited to patients undergoing revision for fracture of the femoral neck and infection. These two diagnoses represented two instances of failure which could be reliably established from the operating notes and/or the histopathological reports. Other reasons for failure such as loosening of a component could have had many underlying causes and thus, the diagnosis was often difficult to establish. Exclusion criteria included implants which had been severely damaged at removal since these would have led to an unreliable measurement of roundness and hence estimation of wear. Implants in the pseudotumour group which have shown a discrepancy between the operative indication and the histopathological report were also excluded. This ensured that the diagnosis of a pseudotumour was confirmed by medical records, the operating notes and the final histopathological report.

**Collection of retrieval implants.** Approval from the local Research Ethics Committee was obtained before implants were accessed from the archived holdings at the authors’ institution, which contained details of retrieved implants obtained from revision procedures. In total, 30 of these implants were investigated. There were eight hip resurfacing arthroplasties retrieved from patients who had undergone revision because of a pseudotumour (group 1) and 22 retrieved from those who had a revision for a fracture of the femoral neck or infection (group 2). A summary of the implants in each group is given in Table I. Although seven of the eight prostheses in the pseudotumour group contained both the acetabular and the femoral components, the acetabular components were not available for analysis in four cases of fracture of the neck of the femur since they had not always been replaced when the femoral component had been revised.
Study: power estimation. Because of the limited number of available retrieved implants which met the selection criteria, it was important to estimate the power of the study. This was done using Altman’s nomogram.18 For the linear rate of wear of MoM hip resurfacing arthroplasty, a significant difference of interest was 10 μm, which represented a twofold increase from the reported steady-state rate of wear of 5 μm for MoM bearings.2 Based on repeated measures taken during a pilot study on five implants, the SD of the expected measurement was estimated to be 8 μm. Thus, a standardised difference of 1.2 was calculated. The statistical level of significance was set at a p-value $\leq 0.05$. Using Altman’s nomogram,18 the sample size in our study (n = 30) estimated the power to be 0.90. This implied that the study had a probability of 90% of correctly concluding that there was a difference of 10 μm. However, since the nomogram is used under the assumption that samples of equal size are required, the unequal samples in our study would have reduced the power.

Assessment of linear wear using a roundness machine. All implants were catalogued by an identification number to ensure patient anonymity. The components were not autoclaved between removal and examination. The bearing surfaces of the retrieved components were first inspected with the naked eye under bright light to assess qualitative changes in the bearing surfaces. The linear wear of the retrieved femoral and acetabular bearing surfaces was then assessed using a Taylor-Hobson Talyrond 290 Roundness Machine (Taylor Hobson Ltd, Leicester, United Kingdom) in a blinded fashion at the Smith and Nephew Implant Development Centre (IDC) (Leamington Spa, United Kingdom).

The roundness testing machine measures both partial and full circles. The machine was used to measure the roundness of the femoral and acetabular components in several planes. Multiple equatorial roundness profiles were then taken from the skirt of the femoral component and the edge of the acetabular component towards their polar regions to locate the maximum wear on each. This was done in a sequential manner in increments of 5 mm. The wear patch was identified as a discontinuity in the characteristic manufactured profile. Once the area of wear had been located, further equatorial roundness profiles were taken in increments of 1 mm away from the area in order to identify the maximum wear value. Polar measurements perpendicular to the equatorial axis were also taken at the site of maximum wear. The location of the wear scar, the maximum depth of wear, was recorded in degrees from the centre of the polar region on the prosthesis. The maximum linear wear, recorded in micrometres (μm), was measured by subtracting the wear profile from an ideal circle. For each of the components, the mean linear wear rate was defined as the maximum linear depth of the wear scar divided by the duration of the implant in vivo (years).

Definition of edge-loading. Edge-loaded components were defined as acetabular components which showed the maximum area of wear crossing over the edge of the implant (Fig. 1). Non-edge-loaded components were defined as acetabular components which showed the maximum area of wear on the device occurring within the hemispherical bearing surface of the component.

Statistical analysis. The distribution of the data was assessed for normality. The Mann-Whitney non-parametric test was used to calculate the level of statistical significance for the differences in the non-normally distributed linear wear and the wear rate between the pseudotumour and non-pseudotumour groups. Pearson’s correlation coefficient was used to assess the strength of the association between the linear wear rate and the time to revision. The incidence of edge-loading in each group was compared using Fisher’s exact test. A p-value $\leq 0.05$ was considered to be significant. The SPSS statistical software programme version 13.0 (SPSS Inc., Chicago, Illinois) was used to perform the analyses.
Results

Qualitative assessment. Artefacts indicating damage after retrieval were evident in a number of implants as isolated scratches or dents. In all cases, fine scratches were observed on the polar regions of bearing surfaces of the femoral head and the acetabular component. There was no evidence of stripe wear such as that reported in ceramic-on-ceramic bearing couples.19 This may be explained by the ability of MoM bearings to self-polish by wearing down the surface surrounding a scratch, which may hide any distinctive visible stripe on the head. Thus, there was minimal wear damage visible to the naked eye in all the implants.

Linear rate of wear. In comparison with the non-pseudotumour implant, those with a pseudotumour had a significantly higher rate of median linear wear of the femoral component of 8.1 μm/year (2.75 to 25.4) compared with 1.79 μm/year (0.82 to 4.15) (Mann-Whitney U test, p = 0.002; Fig. 2) and a significantly higher rate of median linear wear of the acetabular component of 7.36 μm/year (1.61 to 24.9) compared with 1.28 μm/year (0.18 to 3.33) (Mann-Whitney U test, p = 0.001).

Similarly, differences were also measured in absolute wear values. The median absolute linear wear was significantly higher in the pseudotumour implant group, 21.05 μm (2.74 to 164.80) compared with 4.44 μm (1.50 to 8.80) for the femoral component (Mann-Whitney U test, p = 0.005) and 14.87 μm (1.93 to 163.68) compared with 2.51 μm (0.23 to 6.04) for the acetabular component (Mann-Whitney U test, p = 0.008). In all cases, the maximum wear occurred in localised zones.

There was a poor linear correlation between the rate of linear wear and the time in vivo. Pearson’s correlation coefficient (r) was 0.33 for the femoral component, and 0.32 for the acetabular implant (Fig. 3). In addition, there was no significant difference in the rate of linear wear between the fracture and infection subgroups within the non-pseudotumour group (Mann-Whitney U test, p = 0.41).

Edge-loading. This was observed in all the acetabular components in the pseudotumour group. By contrast, it was seen in only one acetabular component in the non-pseudotumour group. The deepest wear was observed well within the bearing surface for the rest of the latter group. The difference in the incidence of edge-loading between the two groups was statistically significant (Fisher’s exact test, p = 0.03).

There was no significant difference in the mean size of the femoral component between the pseudotumour and non-pseudotumour groups (47 mm vs 49 mm, Mann-Whitney U test, p = 0.75).

Discussion

Wear debris from MoM bearing surfaces is generated by mechanical wear, surface corrosion or a combination of both. It consists of both insoluble particles and metal ions, the latter disseminating into the systemic circulation.20 Pseudotumours were not detected in patients who had normal Co and Cr levels. In the light of the recently reported positive correlation between elevated concentrations of serum metal ions and the greater wear of MoM hip resurfacing implants,14 this result suggests that pseudotumours occur when there is increased wear at the MoM articulation. However, direct measurement of wear in these implants has not been previously investigated in patients with pseudotumours.

Our findings have shown that MoM resurfacing implants revised for pseudotumour had significantly greater linear wear of both the femoral and acetabular components compared with a control group of implants revised for other reasons of failure. Implants in the pseudotumour...
group had up to a fourfold increase of the median rate of linear wear of the femoral component and more than a fivefold increase of the acetabular component compared with implants in the non-pseudotumour group.

The median rates of linear wear of the femoral components measured in the pseudotumour group in our study were similar to those recently reported by Witzleb et al. in eight Birmingham Hip Resurfacing (Smith and Nephew) femoral components revised for non-pseudotumour-related failure. In their series seven implants were revised within 15 months of the initial operation. It is difficult to compare these findings with our results since the wear values during the run-in phase are higher than the steady-state values. In these findings with our results since the wear values during the first-year run-in phase to the third-year steady-state has fact, a fivefold reduction in the annual rate of wear from the run-in phase are higher than the steady-state values. In our study was during or beyond the third year of implantation, being 3.6 years for the pseudotumour group and 2.3 years for the non-pseudotumour group.

Potential confounding factors which may have led to increased rates of wear in the pseudotumour group include the relative differences in time from implantation to retrieval, the size of the femoral component and the heterogeneity of subgroups in the non-pseudotumour group.

There was a poor correlation between the linear rate of wear and the time in vivo for implants in both groups. Furthermore, there was no significant difference in sizes of the femoral component between the groups. Lastly, the linear rates of wear between the two sub-groups, namely fracture of the femoral neck and infection, in the non-pseudotumour implants did not differ significantly. Therefore, significantly greater linear rates of wear measured from the retrieved MoM implants revised because of pseudotumour support the findings in the in vivo study of elevated concentrations of metal ions in serum in patients with pseudotumours, thereby confirming that the presence of pseudotumours in patients with contemporary MoM resurfacing implants is associated with increased wear at the MoM articulation.

A characteristic of the morphology of the wear patch on the acetabular components in the pseudotumour group was that maximum wear was always at the edge of the surface of the implant, indicating edge-loading of the bearing. This suggested that the contact zone of the bearing was predominantly positioned at the edge of the implant. Thus, it is likely that the wear mechanism was not that of a normal wear patch expanding with time over the edge of the device, but rather that in which the head component was inadequately covered by the acetabular component from the outset. Edge-loading has been recently suggested as a possible mechanism which leads to increased wear in MoM resurfacing implants because it has the potential to disrupt the favourable fluid-film lubrication that occurs in these bearing surfaces. Since edge-loading and the thickness of the lubricating fluid-film influence the extent of generation of metal ions, these conditions would lead to increased levels of metal ions in hip aspirates and serum, such as those measured in patients with pseudotumours. Therefore, edge-loading with an associated loss of fluid-film lubrication may be the dominant mechanism of wear generation in these patients.

There are several limitations to our study. The method of measurement of linear wear was limited to estimating the local deepest wear and not the mean wear of the component. However, wear was found to be highly localised in the MoM implant couples. The topographical surface shape of the wear scar was not mapped. It would have been informative to characterise fully the corresponding wear scars located on the femoral components. Anatomical orientation of the implants was difficult without landmarks provided at the time of revision since MoM resurfacing implants, unlike stemmed THRs, are symmetrical in appearance without distinctive geometrical markers which can be used for orientation. Since the retrieved implants were not marked with in vivo orientation at the time of retrieval, information regarding the specific anatomical orientation of the wear scar could not be assessed. Despite these limitations, the pattern of edge-loading was clearly demonstrated on the acetabular components. Complete good-quality plain radiographs were only available for fewer than half of the selected retrieved implants. Therefore, the radiological inclination of the acetabular component, which has been reported to influence wear in MoM bearings, could not be reliably assessed.

---

**Scattergram showing the linear wear rates of all the retrieved implants plotted against time in vivo.**

---

<table>
<thead>
<tr>
<th>Component</th>
<th>Linear Wear Rate (µm/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femoral</td>
<td>0.0, 1.5, 3.0, 4.5, 6.0, 7.5</td>
</tr>
<tr>
<td>Acetabular</td>
<td>0.0, 1.0, 2.0, 3.0, 4.0, 5.0</td>
</tr>
</tbody>
</table>

Pearsons correlation coefficient (r):
- Femoral component: r = 0.33
- Acetabular component: r = 0.32
It was assumed in our study that the rate of wear was linear with time. Data from hip simulator studies and clinical measurement of metal ion levels suggest a biphasic rate of wear with an initial high run-in followed by lower steady-state wear. Although most implants in both groups were from or beyond the third year of implantation, several in the non-pseudotumour group had failed earlier, within the higher wearing run-in period. This may have led to underestimation of the wear difference between the implant groups. The wear rates are also sensitive to material parameters (carbon content, alloy processing and heat treatment), design parameters (radial clearance), and manufacturing parameters (surface roughness and sphericity). There are differences in designs, materials and manufacturing processes between the three types of MoM hip resurfacing implants measured in our study. This confounding factor could be controlled if only one type of implant from a single manufacturer was measured. However, this would have further reduced the available number of retrieved implants, thereby reducing the power of the study. Moreover, a proportionally similar number of types of implant was studied in each group, thus minimising this confounding effect on comparative values.

Significantly higher linear rates of wear were measured in the retrieved femoral head and acetabular components revised because of pseudotumour than in those revised for other reasons. This supports the in vivo findings of elevated concentrations of metal ions in patients with pseudotumours. The results of our study therefore provide the first direct evidence to confirm that the presence of pseudotumours in patients with contemporary MoM resurfacing implants is associated with increased wear at the MoM articulation. Thus, the soft-tissue pseudotumour, a clinical complication associated with a high rate of revision, may represent a logical biological reaction to an increased burden of wear debris. Furthermore, highly localised wear occurring because of edge-loading with an associated loss of fluid-film lubrication may be the dominant wear mechanism in the patients with pseudotumours since no pseudotumours were observed in non-edge-loaded devices. Further in vivo investigations are required to evaluate the risk factors of edge-loading such as the positioning of the component as a mechanism responsible for increased wear in these patients. Such evidence-based knowledge would be crucial for the selection of patients, the surgical technique and the design of future hip resurfacing implants in order to ensure the long-term survivorship of the prosthesis.

References