HIP

What advantage is there to be gained using large modular metal-on-metal bearings in routine primary hip replacement?

A PRELIMINARY REPORT OF A PROSPECTIVE RANDOMISED CONTROLLED TRIAL

The aim of this study was to investigate the possible benefit of large-head metal-on-metal bearing on a stem for primary hip replacement compared with a 28 mm diameter conventional metal-on-polyethylene bearing in a prospective randomised controlled trial.

We investigated cemented stem behaviour between these two different bearings using Einzel-Bild-Röntgen-Analyse, clinical and patient reported measures (Harris hip score, Western Ontario and McMaster Universities osteoarthritis index, Short Form-36 and satisfaction) and whole blood metal ion levels at two years. A power study indicated that 50 hips were needed in each group to detect subsidence of > 5 mm at two years with a p-value of < 0.05.

Significant improvement (p < 0.001) was found in the mean clinical and patient reported outcomes at two years for both groups. Comparison of outcomes between the groups at two years showed no statistically significant difference for mean stem migration, clinical and patient reported outcomes; except overall patient satisfaction which was higher for metal-on-metal group (p = 0.05). Metal ion levels were raised above the Medicines and Healthcare products Regulatory Agency advised safety level (7 μg per litre) in 20% of the metal-on-metal group and in one patient in metal-on-polyethylene group (who had a metal-on-metal implant on the contralateral side). Two patients in the metal-on-metal group were revised, one for pseudotumour and one for peri-prosthetic fracture.

Use of large modular heads is associated with a risk of raised whole blood metal ion levels despite using a proven bearing from resurfacing. The head-neck junction or excess stem micromotion are possibly the weak links warranting further research.

Metal-on-metal (MoM) articulations have shown good results over the last ten years for hip resurfacing,1,2 with low metal ion levels reported on long-term follow-up.1-5 The perceived advantages of these bearings are better range of movement, reduced dislocation rates,6,7 and improved wear properties.8 Modular heads were introduced to address the complication of femoral neck fracture9,10 by inserting a stem with a large modular femoral head and retaining the acetabular component. The use of a large bearing11 (≥ 36 mm) on a stem gained popularity. Previous reviews on MoM total hip replacement (THR) comprised smaller head diameters and monobloc femoral components but to the best of our knowledge there is only one prospective randomised controlled trial (RCT) comparing large-head MoM with metal-on-polyethylene (MoP) THR.12

Collarless polished tapered femoral stems are used widely in THR with long-term outcomes documented13,14 but always using conventional bearings. The addition of MoM bearing may have an effect on the behaviour of the stem as there are documented differences in frictional torque of MoM compared with MoP bearings;15,16 there is also an increased sliding distance of the bearing couple and moment arm from the tapered trunnion to the joint line. The British Hip Society (BHS),17 British Orthopaedic Association (BOA)18 and National Institute of Health and Clinical Excellence (NICE) guidelines19 have recommended limited use of new constructs and monitoring through prospective RCTs. We aimed to monitor the introduction of a large-head modular MoM articulation by comparing it with the current standard 28 mm diameter MoP bearing on a collarless polished tapered cemented stem.

Patients and Methods

This is a preliminary report from a single-centre, single-surgeon prospective RCT comparing two different articulations and two arms of randomisation. Ethical approval was
obtained. Patients who underwent a primary THR under the care of the senior author (JPH) were approached for participation. The following exclusions were applied: patients who were considered unsuitable for the use of an uncemented press-fit acetabular component (at the discretion of the senior author), those unable to answer questionnaires for cognitive reasons (Standard Mini Mental State Examination < 12 of 30), those who refused consent and young, active males below the physiological age of 65 years, who were suitable for resurfacing in accordance with NICE guidelines, with a resurfaceable femoral head. After screening, written consent was obtained. Randomisation was on an intention to treat basis to one of the two arms the day before surgery. Patients were assigned either a large MoM or standard MoP bearing surface in combination with a collarless polished tapered cemented stem (CPCS; Smith & Nephew, Hull, United Kingdom). It was unethical to blind patients to the combination of implants, therefore follow-up evaluations were unblinded. All operations were performed using a standard posterior approach. Antibiotic prophylaxis and thromboprophylaxis protocols (foot pumps and anti-embolism stockings with additional chemical agents for higher risk cases) were similar for both the groups. Post-operative rehabilitation was similar for both groups and full weight-bearing mobilisation was allowed for all patients, with no precautions other than avoidance of low seating for the first six weeks.

We attempted to undertake a THR where the only variable was the bearing. The collarless polished tapered cemented stem used was compatible with the Birmingham hip modular head system (Smith and Nephew) (Fig. 1). It is a triple tapered polished stem available in standard or high offset (only standard offset was used in this study) with 12/14 morse taper for modular head engagement. A Birmingham hip acetabular component with a corresponding large modular cobalt chrome head without a taper sleeve was used in patients who were randomised to the MoM arm of study. A Reflection acetabular component (Smith & Nephew) without holes or screws, with a standard polyethylene liner (Fig. 1) was used in those randomised to the MoP arm.

The primary outcome was to compare the migration of the femoral component between groups. This was done using Einzel-Bild-Röntgen-Analyse (EBRA; University of Innsbruck, Innsbruck, Austria) software with previously validated techniques. Secondary outcome was to compare the pre-operative clinical and patient-reported outcomes of pain, function and satisfaction and those at three months, one year and two year follow-up between the two groups. Clinical outcomes comprised the Harris hip score (HHS) collected by an independent physiotherapist (JRR). Patient-reported outcomes were assessed using the Western Ontario and McMaster Universities osteoarthritis index (WOMAC), the Short-Form 36 health survey (SF-36), comorbid medical conditions, height and weight measures, the Hip disability and Osteoarthritis Outcome Score (HOOS) and patient satisfaction. Satisfaction data were recorded using a validated measure of satisfaction. This questionnaire includes four questions asking about satisfaction with overall outcome, pain relief, ability to perform activities of daily living and ability to participate in leisure activities. Responses are on a four point Likert scale, which ranges from very satisfied to very dissatisfied. Complications were recorded using a standard proforma.

Radiological analyses. Analyses of anteroposterior (AP) views of the pelvis and lateral views of the hip, were done pre-operatively and at each follow-up by an independent assessor (RB). The EBRA-FCA (femoral component...
analysis using EBRA\textsuperscript{22} method was used to study comparability between three pairs of radiographs; reference points were marked on the femoral component and the program calculated stem subsidence.\textsuperscript{22} The importance of acetabular orientation has been established in relation to raised metal ions\textsuperscript{32} therefore EBRA was also used to measure acetabular inclination and anteversion\textsuperscript{33,34} on AP pelvis radiographs, as validated by Langton et al.\textsuperscript{35}

**Metal ion analysis.** In accordance with the Medical Devices Alert issued by the Medicines and Healthcare products Regulatory Agency (MHRA) in April 2010,\textsuperscript{36} whole blood cobalt (Co) and chromium (Cr) ion levels were measured on all patients after one year. These tests were analysed at the Clinical Chemistry department, Imperial Healthcare College NHS Trust, London (participating laboratory of the Trace Element External Quality Assessment Scheme).

**Statistical analysis.** All analyses were performed with SPSS software version 17 (SPSS Inc. Chicago, Illinois). A sample size calculation prior to enrolment showed 40 hips were required in each arm of the study to detect a stem subsidence of > 0.5 mm at two years with a power of 0.90 at a p-value < 0.05. Allowing for a loss to follow-up of 5% at each assessment time, we aimed for 50 hips in each arm of study. In order to determine whether the data were significantly different from the normal distribution, a Kolmogorov-Smirnov test was used. If p < 0.05, the data were treated as non-parametric. To compare continuous variables two sample t-tests were used for parametric data and Mann-Whitney U test was used for non-parametric data. All tests were two tailed and a significance level of p < 0.05 was maintained throughout. Chi-squared tests were used to compare categorical data. Comparisons of migration of the stem, and of anteversion and inclination of the acetabular component between the groups were done using t-tests. Differences in outcomes for HHS, WOMAC and SF-36 scores between the groups were analysed with repeated measures analysis of variance (ANOVA) and patient satisfaction between the groups was compared using chi-squared analysis. Pearson’s and Spearman’s rho tests were used to study the correlation between metal ion levels and bearing surfaces. All tests were two-tailed and a significance level of 0.05 maintained.

**Results**

There were 95 patients in the study (100 THRs; five patients had bilateral THR and took part in the study twice). Their mean age was 64.4 years (42 to 82) and their mean body mass index (BMI) was 29 (22 to 46). There was no significant difference between the groups for age, gender, BMI and side of operation (Table I).

No difference was found in mean femoral stem subsidence between the MoM and MoP groups at one year (0.75 mm (0.10 to 3.90) versus 0.73 mm (0.10 to 3.00); t-test, p = 0.922) and two-year follow-up (1.35 mm (0.10 to 6.20) versus 1.38 mm (0.10 to 5.20); t-test, p = 0.93).

<table>
<thead>
<tr>
<th>Femoral head sizes (mm) (n)</th>
<th>MoM (n = 50)</th>
<th>MoP (n = 50)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>(0)</td>
<td>50</td>
<td>(100)</td>
</tr>
<tr>
<td>38</td>
<td>(1)</td>
<td>2</td>
<td>(2)</td>
</tr>
<tr>
<td>42</td>
<td>(18)</td>
<td>46</td>
<td>(46)</td>
</tr>
<tr>
<td>48</td>
<td>(1)</td>
<td>50</td>
<td>(50)</td>
</tr>
<tr>
<td>54</td>
<td>(1)</td>
<td>50</td>
<td>(50)</td>
</tr>
</tbody>
</table>

* t-test, unless otherwise stated
† chi-squared test

The mean pre-operative HHS improved significantly at two-year follow-up in both MoM (42.1 (20.1 to 72.8) to 87.5 (67.8 to 100); ANOVA, p = 0.001) and MoP groups (46.4 (22.2 to 63.6) to 88.1 (52.6 to 100); ANOVA, p = 0.001). However, there was no statistically significant difference between groups at two years (ANOVA, p = 0.329).

There was a significant improvement in the mean WOMAC scores for MoM (pain 31.4 (4.4 to 70) to 90.5 (56.2 to 100); function 31.3 (6.4 to 66.2) to 80.9 (57.6 to 100); stiffness 39.8 (12.1 to 100) to 85.1 (64.1 to 100); ANOVA, p < 0.001) and MoP groups (pain 35.1 (10.1 to 80) to 85.8 (56.3 to 100); function 34.7 (14.7 to 83.2) to 79.2 (52.1 to 100); stiffness 40.1 (12.6 to 87.5) to 74.3 (45.8 to 100); ANOVA, p < 0.001) from pre-operative to two years. There was no significant difference between groups in terms of the mean WOMAC scores for pain (ANOVA, p = 0.296), function (ANOVA, p = 0.303) and stiffness (ANOVA, p = 0.093) at two years.

There was a significant improvement in the mean SF-36 scores from pre-operative to two years for both groups (all domains, ANOVA, p < 0.001) except general health (ANOVA, p = 0.277). There was no significant difference (p = 0.34) between groups for all domains. The MoM group reported better overall satisfaction than MoP group (100% versus 91.7%; ANOVA, p = 0.05) at two years. There was no difference between groups in terms of satisfaction reported for pain relief (chi-squared, p = 0.219), activities of daily living (ANOVA, p = 0.081), recreational activities (ANOVA, p = 0.658) and HOOS components for squatting (ANOVA, p = 0.714), running (ANOVA, p = 0.782), jumping (ANOVA, p = 0.204), twisting/pivoting on loaded leg (ANOVA, p = 0.684) and walking on uneven surface (ANOVA, p = 0.342).
At two years, the mean whole blood levels of Cr for the MoM and MoP groups were 2.78 μg/L (0.30 to 7.85) and 0.79 μg/L (0.25 to 4.5) respectively (t-test, p < 0.001) (Fig. 2). The mean levels of Co for the MoM and MoP groups were 5.21 μg/L (1.2 to 14.2) and 1.61 μg/L (0.24 to 8.4) respectively (t-test, p < 0.001) (Fig. 3). A total of eight of 40 results of patients in the MoM group reported metal ion levels (either Cr or Co) higher than the MHRA recommended action level of 7 μg/L\(^{36}\) compared with only one of 29 results in the MoP group. The patient in the MoP group who had high metal ion levels had a bilateral THR with MoM bearing surface on the other side.

There was significant association between serum Cr levels and acetabular component inclination in MoM patients (r = 0.389, p = 0.014); but no significant association between serum Cr levels and acetabular component anteversion (r = -0.027; p = 0.23), Co levels to acetabular
component inclination \( (r = -0.141, p = 0.19) \) and Co levels to acetabular anteversion \( (r = -0.075, p = 0.34) \). There was no significant association between serum Cr levels and femoral head size \( (r = 0.102, p = 0.26) \) or serum Co levels to femoral head size \( (r = 0.119, p = 0.32) \).

**Complications.** At two years, three patients had died from unrelated causes. Four patients underwent revision. Two hips were revised in the MoP group for recurrent dislocation after trauma. Two hips in the MoM group were revised, one for pseudotumour formation and another for a peri-prosthetic fracture, both at four-year follow-up.

The patient with pseudotumour (62 year old female) had presented with unexplained pain; serum metal ion analysis revealed a Cr level of 3.85 \( \mu \)g/L and a Co level of 8.1 \( \mu \)g/L. Radiographs did not show any obvious abnormality, the component alignment was satisfactory (Fig. 4); MRI scans showed evidence of a soft-tissue mass (Fig. 5). During revision there was extensive soft-tissue destruction but no evidence of metal staining. Volumetric wear analysis of the bearing surface and the internal surfaces of the tapers using a co-ordinate measuring machine (Legex 322; Mitutoyo, Hampshire, United Kingdom) with an accuracy of 0.8 \( \mu \)m, showed no material loss from the bearing surface or corrosion of the posteromedial stem, but some wear was observed at the trunnion and the taper exhibited wear to a depth of 6 \( \mu \)m with a circumferential pattern of wear maximal supero-anteriorly (Fig. 6).

The patient with a peri-prosthetic fracture (72-year-old female) had presented with pain in the hip after twisting her ankle and was found to have a fracture of the greater trochanter at two-year follow-up. She was treated non-operatively and gradually improved, although the fracture did not unite and about four years after her surgery radiographs revealed progression of the fracture to the tip of the stem (Fig. 7). The three-year whole serum metal ion levels were Cr of 4.29 \( \mu \)g/L and Co of 1.04 \( \mu \)g/L. During revision there was no evidence of metal staining, but pseudotumour fluid was present with soft-tissue destruction secondary to
adverse reaction to metallic debris and stem corrosion was noted. Wear analysis showed little wear of the trunnion.

Discussion
Improved materials and designs of THR have allowed the use of larger bearings providing increasing range of movement with enhanced stability, low wear and better restoration of gait. Large MoM THR evolved directly from the success of MoM hip resurfacing. However, there is little evidence in relation to the suitability of simply exchanging the conventional MoP bearing on a stem for a large MoM head. Clinical, functional and metal ion results for standard diameter MoM THR have been presented, but there is a lack of outcome studies on large head MoM THR.

We found no difference in the behaviour of the polished tapered cemented collarless stem as detectable by EBRA between the two bearing surfaces (MoP and MoM) up to two years. Previous safety studies have shown that this stem has similar migration properties to the more established Exeter stem and therefore the results can probably be extrapolated to the other polished taper cemented stems.

The clinical and patient-reported outcomes and the complication rates in both groups were similar at two years. Thus the functional outcome after large head MoM THR is similar to that of MoP THR. Although we found higher overall satisfaction rate (p = 0.05) in the MoM group this may be confounded by non-blinding. Trials comparing 28 mm MoM THR with 28 mm MoP THR, have reported similar outcomes. As hip resurfacing has shown better short term outcomes than conventional THR, one might expect these results to translate to large head MoM THR, but from our results it does not appear so. This may be because the trial was offered to a standard patient population and patients who would be suitable for resurfacing, with an expected higher level of activity, were excluded.

There was a significant difference in number of patients with whole blood Co or Cr levels above the MHRA ‘safe level’ at two years in MoM versus MoP groups. From the MoM group, two patients underwent revision and had an adverse reaction to metal debris and stem corrosion at revision. The source of metal wear debris in the first case seemed predominantly from the morse taper of the head-neck junction. The second case had minimal trunnion wear, but significant stem corrosion. It is possible that despite a relatively innocuous trochanteric fracture two years earlier, the stem became unstable causing local tissue reaction to stem wear debris and ultimately, progression to a peri-prosthetic fracture.

Unlike studies of hip resurfacing we did not find direct correlation between the Co-Cr whole blood metal ion levels and the orientation of the acetabular component or to the size of the femoral head. We did note the low positive correlation between the Cr levels and acetabular component inclination, but this was not reflected in the Co levels. This might be because the anteversion and inclination of the acetabular component for most patients was within the acceptable range (Fig. 8). It is well known that the head-neck junction is a source of release of metal ions because of fretting and corrosion. We believe that a cause of high metal ion levels in the MoM group may be the increased torque on the trunnion, with possible secondary wear. Garbuz et al in a study comparing large head MoM THR with MoM resurfacings concluded that the cause of high metal ion in the MoM THR group was related to the modularity of the attachment of the femoral head to the stem. Similar to our findings they have noted that the high metal ion levels have lasted beyond two years post-operatively, unlike following resurfacing where it tends to plateau beyond one year.

The 7th annual report of the National Joint Registry has reported that the five-year revision rate was the highest for large head MoM THR at 7.8%. The preliminary report from this study suggests that using this large head combination bearing in primary uncomplicated THR gives no clear functional benefit. Two patients required revision in each arm of the study and only long-term follow-up will reveal any potential survival benefit for the MoM bearing. Although there were no clinical adverse features in patients with elevated ions they will, as recommended, be kept under surveillance.

These results should be extrapolated to other MoM bearing designs with caution due to differences in metallurgy (cast versus heat-treated), diometrical clearance, arc of head cover, acetabular component thicknesses relating to deflection on insertion and differences in head-neck tapered trunnions. Extrapolation to uncemented titanium stems must also take account of the differing materials. If stems are stable and the bearing proven, the source of the raised ions is unclear. However the tapered trunnion may, by the process of elimination, be a target for future investigation and research.
Our study shows that functional outcome is no better after large MoM hip replacements, although this group had a higher overall satisfaction rate. We believe that any benefits of large MoM hip replacement are outweighed by the associated risk of high metal ion levels, as highlighted in the MHRA alert.

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


