A prospective randomised controlled trial comparing three alternative bearing surfaces in primary total hip replacement


From St Michael’s Hospital, Toronto, Ontario, Canada

The ideal bearing surface for young patients undergoing total hip replacement (THR) remains controversial. We report the five-year results of a randomised controlled trial comparing the clinical and radiological outcomes of 102 THRs in 91 patients who were < 65 years of age. These patients were randomised to receive a cobalt–chrome on ultra-high-molecular-weight polyethylene, cobalt–chrome on highly cross-linked polyethylene, or a ceramic-on-ceramic bearing. In all, 97 hip replacements in 87 patients were available for review at five years. Two hips had been revised, one for infection and one for peri-prosthetic fracture.

At the final follow-up there were no significant differences between the groups for the mean Western Ontario and McMaster Universities osteoarthritis index (pain, \( p = 0.543 \); function, \( p = 0.10 \); stiffness, \( p = 0.99 \); Short Form-12 (physical component, \( p = 0.878 \); mental component, \( p = 0.818 \)) or Harris hip scores (\( p = 0.22 \)). Radiological outcomes revealed no significant wear in the ceramic group.

Comparison of standard and highly cross-linked polyethylene, however, revealed an almost threefold difference in the mean annual linear wear rates (0.151 mm/year versus 0.059 mm/year, respectively) (\( p < 0.001 \)).

The ideal bearing surface for young active patients undergoing total hip replacement (THR) remains controversial. Younger patients have a longer life expectancy and place greater demands on the prosthetic components. The choice of implant, and particularly the bearing coupling, is of great importance when considering the eventual need for revision surgery.

The most common mode of failure in THR is aseptic loosening due to particulate debris generated by wear of ultra-high molecular-weight polyethylene (UHMWPE). Osteolysis has been reported in up to 60% of young active patients with UHMWPE bearings at five years. In an attempt to avoid the problems caused by wear debris, different bearing surfaces have been developed, such as highly cross-linked polyethylene (HXLPE), which is intended to reduce linear and volumetric wear compared with standard polyethylene. Hard bearing surfaces such as ceramic-on-ceramic (CoC) and metal-on-metal (MoM) have also been developed in an attempt to address the problems of osteolysis, but have their own inherent limitations.

It has been shown that HXLPE exhibits markedly less wear than conventional UHMWPE, both in vitro and in vivo. Similarly, it has been reported that MoM and CoC hip replacements produce even less volumetric wear and therefore have a reduced reaction to wear debris. MoM bearings produce 20 to 200 times less volumetric wear than conventional metal-on-polyethylene bearings. However, there are concerns about increased metal ion levels in serum, whole blood, plasma or erythrocytes with these bearings, especially in young women, as fetal exposure to metal ion levels in pregnancy has been demonstrated. CoC bearings have outstanding wear performance, but complications such as fracture or squeaking remain a concern.

In order to investigate the ideal bearing coupling for patients < 65 years of age we studied the clinical and radiological outcomes of a group of patients randomised to receive a CoC, cobalt–chrome on UHMWPE or cobalt–chrome on HXLPE bearing surface.

Patients and Methods

Between February 2003 and March 2005, 91 patients (102 hips) with a mean age of 52.7 years (19 to 64) were operated on at a single centre by one of three surgeons (EB, EHS, JPW). All patients received an uncemented femoral component (Synergy; Smith & Nephew, Memphis, Tennessee) with an...
uncemented acetabular component (Reflection; Smith & Nephew). Patients were randomised using opaque envelopes and a computer generated block randomisation scheme (block size = 4), provided by the study coordinator on the day of surgery, to one of three possible bearing surface combinations: a 28 mm diameter cobalt–chrome head with UHMWPE liner (CP group, 36 hips); a 28 mm diameter cobalt–chrome head with HXLPE liner (HXLPE group, 32 hips); or a 28 mm diameter ceramic head and ceramic acetabular liner (CoC group, 34 hips). Patients and the staff managing post-operative care were blinded to the implant inserted.

Inclusion criteria consisted of patients aged between 18 and 65 years undergoing THR for primary or secondary osteoarthritis. Exclusion criteria included a history of hip joint sepsis, primary or secondary malignancy in the involved hip, acute fracture of the femoral neck, and bone deficiency requiring the use of autologous or allograft bone for either acetabular or femoral reconstruction.

A standardised pre-operative evaluation was carried out in all patients between one and two weeks pre-operatively, usually at the time of pre-admission. This included history, registration of demographics and comorbidities, and completion of the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) and Short-Form 12 (SF-12) self-reported questionnaires. Patients were also asked to categorise their level of activity on a three-level scale (light, moderate or heavy).

Additionally, all patients had a physical examination and assessment using the Harris hip score (HHS). Digital pre-operative radiological evaluation included an antero-posterior (AP) pelvic radiograph, as well as AP and lateral radiographs of the involved hip. Each patient was reviewed clinically and radiologically at three, 12, 24, 36 and 48 months after surgery. WOMAC, SF-12 and radiological assessments were undertaken at each review. At 60 months a blinded reviewer undertook an evaluation of the HHS.

An independent blinded reviewer (MRE) undertook the radiological assessment. However, it is obvious that the analyser was able to distinguish the CoC bearing from the other two types of articulation. Digital pelvic and AP and lateral hip radiographs were used. All patients received a 28 mm femoral head, and this was used to scale the images using digital templating software (EndoMap v2.01; Hectec GmbH, Niederviebach, Germany).

Acetabular inclination was measured using the transischial line as reference, and anteversion by the validated method described by Widmer. Any osteolysis or radiolucent lines (RLLs) > 1 mm were assessed using the zones of Gruen, McNeice and Amstutz and of DeLee and Charnley for the femoral and acetabular components, respectively. Linear wear was measured in the most recent radiographs at a minimum of five years. The validated technique described by Dorr and Wan was used on the scaled images. The total wear rate was determined by comparing the AP radiograph obtained six weeks after surgery with that obtained at last follow-up. The annual wear was calculated by dividing total wear by the years of follow-up.

This study was approved by the hospital’s Ethics Review Committee in addition to the University of Toronto Ethics Review Committee.

**Statistical analysis.** The three groups were verified for comparability of baseline measurements, including age, gender, primary diagnosis, level of education, the hip being replaced, surgeon, WOMAC, HHS and SF-12 scores, using appropriate univariate techniques (including Mann-Whitney U and chi-squared tests). Analysis of variance (ANOVA) was undertaken using the Kruskal-Wallis (KW) test to assess differences at the final (60 month) review with regard to the HHS, WOMAC and SF-12 scores and the liner wear between groups. A p-value < 0.05 was considered significant for all analyses.

**Results**

The three groups were comparable in terms of baseline and pre-operative parameters (Table I). Primary osteoarthritis was the pre-operative diagnosis in 67 patients (66%) (Table II).

A total of 97 THRs in 87 patients were available for review at five years’ follow-up. Two hips in two patients had been revised (one for infection at 18 months and one for peri-prosthetic fracture four months post-operatively). Two patients (three hips) were lost to follow-up.

At the 60-month follow-up the mean WOMAC, SF-12 and HHS scores (with the exception of the SF-12 mental component) had improved significantly compared with pre-operative values in all three groups (Table III). At five years there were no statistical differences in the clinical outcomes using the HHS, WOMAC or SF-12 scores (Table IV; Figs 1 and 2).

At five years there were no failures due to aseptic loosening. Radiological analysis revealed non-progressive radiolucency > 2 mm involving the femoral component in one patient in the CP group, and in the acetabular component of one patient in the HXLPE group.

Analysis of liner wear, however, revealed significant differences between the groups. The mean total wear at five years was 0.869 mm (SD 0.535) for the CP group, 0.329 mm (SD 0.209) for the HXLPE group and 0.035 mm (SD 0.073) for the CoC group (p < 0.001, Kruskal-Wallis) (Fig. 3). The mean annual liner rate of wear was 0.151 mm/year (SD 0.09), 0.059 mm/year (SD 0.014) and 0.001 mm/year for the CP, HXLPE and CoC groups, respectively (p < 0.001, Kruskal-Wallis) (Fig. 3). Considering each group separately, we found that the CP group had significantly higher annual and mean five-year wear rates than the HXLPE group (p < 0.001 for both, Mann-Whitney U test). Furthermore, the HXLPE group had significantly higher annual and mean five-year wear rates than the CoC group (p < 0.001 for both, Mann-Whitney U test).

Three patients with a CoC bearing reported squeaking that had started between three and five years post-operatively.
Radiological assessment revealed no abnormalities, and in all three patients the acetabular component had been placed within the safe zone of 40° (± 10°) of inclination and 15° (± 10°) of version, as defined by Lewinnek et al. Despite the noise, no patients required revision of their THR.

**Discussion**

As expected from the randomisation, pre-operative assessment did not reveal any significant difference between the groups for the WOMAC, SF-12 or HHS assessments.

Patients in all groups demonstrated significant improvement in the scores at the five-year follow-up (with the exception of the SF-12 mental component), but despite the different bearing surfaces, no statistically significant differences were found among the three groups regarding the five-year patient-based or functional outcomes. However, significant differences were recorded for the final liner wear.

Previous studies have shown that HXLPE significantly reduces wear compared with UHMWPE. In a prospective randomised trial, Geerdink et al reported that, at a mean follow-up of eight years (7 to 9), HXLPE had a significantly reduced annual wear rate (0.088 mm/year (sd 0.03)) compared with UHMWPE (0.142 mm/year (sd 0.07)). In another prospective randomised study, McCalden et al found that, at a mean follow-up of 6.8 years, there were no differences between the two polyethylene groups with regard to the HHS, WOMAC or SF-12 scores, but there was a mean linear wear of 0.003 mm/year for HXLPE, whereas UHMWPE had a mean linear wear rate of 0.051 mm/year. Likewise, Mutimer et al reported that, at a mean follow-up of 5.5 years (4.1 to 7), HXLPE had a mean wear rate of 0.05 mm/year and UHMWPE had a mean wear rate of 0.26 mm/year.

We compared UHMWPE, HXLPE and CoC articulations in a randomised manner and found mean wear rates of 0.151 mm/year (sd 0.09) for UHMWPE compared with 0.059 mm/year (sd 0.041) for HXLPE (p < 0.001). These results are similar to those of previous randomised trials and support the superiority of HXLPE over conventional UHMWPE in reducing wear rates in the medium term. Osteolysis around a well-positioned THR is frequently a result of cellular response to wear debris. It has
been shown that wear rates > 0.2 mm/year can cause 100% failure of THRs, and wear rates > 0.15 mm/year can put the hip at risk for aseptic loosening, at a maximum of follow-up of 25 years.45 Osteolysis is uncommon with a wear rate < 0.1 mm/year.46 With different head sizes such a penetration rate will produce different volumes of wear. For example, for 32 mm heads such a wear rate corresponds to a volumetric wear rate of 80 mm³/year, and for 28 mm heads, as in our series, 62 mm³/year.

We wanted to compare polyethylene on cobalt–chrome bearings with CoC bearings, as the exceptional hardness and smoothness of the latter have shown minimal wear.

Table III. Pre-operative and 60-month post-operative scores (WOMAC, Western Ontario and McMaster Universities osteoarthritis index; SF-12, Short-Form 12; HHS, Harris hip score) between the three groups (CP, cobalt–chrome on conventional ultra-high molecular-weight polyethylene; HXLPE, highly cross-linked polyethylene; CoC, ceramic-on-ceramic)

<table>
<thead>
<tr>
<th>Score</th>
<th>Pre-operative</th>
<th>60 months</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean WOMAC (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP Pain</td>
<td>47.57 (15 to 75)</td>
<td>91.65 (35 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Stiffness</td>
<td>38.60 (12.50 to 75)</td>
<td>87.04 (37.5 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Function</td>
<td>42.82 (20.6 to 86.8)</td>
<td>83.05 (44.1 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>HXLPE Pain</td>
<td>39.84 (10 to 75)</td>
<td>92.39 (60 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Stiffness</td>
<td>35.00 (12.5 to 75)</td>
<td>87.50 (50 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Function</td>
<td>35.34 (10.3 to 72.1)</td>
<td>82.67 (55.9 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>CoC Pain</td>
<td>42.73 (10 to 75)</td>
<td>86.17 (55 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Stiffness</td>
<td>41.13 (12.5 to 75)</td>
<td>86.88 (50 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Function</td>
<td>40.86 (5.9 to 75)</td>
<td>89.03 (35.3 to 100)</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

Table IV. Comparison between the outcome scores (WOMAC, Western Ontario and McMaster Universities osteoarthritis index; SF-12, Short-Form 12; HHS, Harris hip score) at the five-year follow-up for the three groups (CP, cobalt–chrome on conventional ultra-high molecular-weight polyethylene; HXLPE, highly cross-linked polyethylene; CoC, ceramic-on-ceramic). There were no statistically significant differences between the scores

<table>
<thead>
<tr>
<th>Score</th>
<th>CP</th>
<th>HXLPE</th>
<th>CoC</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean WOMAC (range)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>91.65 (35 to 100)</td>
<td>92.39 (60 to 100)</td>
<td>86.17 (55 to 100)</td>
<td>0.543</td>
</tr>
<tr>
<td>Stiffness</td>
<td>87.04 (37.5 to 100)</td>
<td>87.50 (50 to 100)</td>
<td>86.88 (50 to 100)</td>
<td>0.99</td>
</tr>
<tr>
<td>Function</td>
<td>83.05 (44.1 to 100)</td>
<td>82.67 (55.9 to 100)</td>
<td>89.03 (35.3 to 100)</td>
<td>0.102</td>
</tr>
<tr>
<td>Mean SF-12 (range)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>48.02 (14.29 to 63.45)</td>
<td>47.92 (30.76 to 64.8)</td>
<td>49.44 (23.93 to 59.41)</td>
<td>0.878</td>
</tr>
<tr>
<td>Mental</td>
<td>52.30 (20.44 to 65.77)</td>
<td>55.49 (42 to 63.38)</td>
<td>54.49 (32.18 to 64.74)</td>
<td>0.818</td>
</tr>
<tr>
<td>Mean HHS (range)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>47.82 (20.55 to 67.19)</td>
<td>54.49 (32.18 to 64.74)</td>
<td>0.096</td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td>47.06 (22 to 63)</td>
<td>87.88 (61 to 98)</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>CoC</td>
<td>51.89 (23 to 90)</td>
<td>91.47 (63 to 100)</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>CoC</td>
<td>45.69 (23 to 90)</td>
<td>91.03 (61 to 100)</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
</tbody>
</table>

* Mann-Whitney U test

* Kruskal-Wallis test
rates and limited osteolysis in the long term, even in young active patients. In hip simulator testing, a CoC articulation produced wear rates of $1 \mu m/yr$. In a previous randomised trial from our institution, comparing a CoC with a ceramic on conventional polyethylene bearing we showed that CoC wear was as low as $0.02 mm/yr$. In our current study the annual wear rate for the CoC group was even lower, at $0.0067 mm/yr$ (SD 0.014), supporting the excellent behaviour of the ceramic bearing at the five-year follow-up period. However, three of 34 patients (8.2%) with ceramic bearings reported squeaking. The incidence of squeaking after CoC THR has been reported to vary between 0.7% and 20.9%. The squeaking, which started three to five years post-operatively, has been previously described. Walter et al reported that squeaking occurred at a mean of 17 months post-operatively, and Mai et al reported that squeaking usually occurred 12 to 30 months post-operatively, although it may be delayed until 48 months.

We acknowledge the limitations of this study. First, the number of patients in each group was smaller than initially planned. It had been our intention to recruit 45 patients per group to limit the likelihood of a type I error to 0.05 with a power > 95%, but unfortunately we were unable to recruit the full complement. Nevertheless, rates of wear between the groups were markedly different, and it seems unlikely that results would have changed with a larger number of patients. Second, in this study the assessment of wear involved the technique of Dorr and Wan, using two-dimensional digital clinical radiographs and templating software. It is true that many studies have shown the alternative Livermore technique to be more accurate, even though there are studies that support the opposite. We chose to use the Dorr and Wan technique because it is simple to use and we have used it successfully in previous studies. Additionally, both the techniques of Livermore and Dorr and Wan were described to measure wear on metal/UHMWPE articulations, and it is easy to modify the Dorr and Wan method for CoC articulations, but not easy to use the Livermore method on radiographs of CoC THRs.

Our main concern is the accuracy of measurements in cases of extremely low rates of wear, such as those recorded...
in the CoC group of patients. Many methods of measuring linear wear on conventional or digital radiographs have been described, and this may explain the differences in wear rates that have been reported for the same articulations in different studies. Although radiostereometric analysis is the most accurate tool for in vivo assessment of polyethylene wear, it is expensive and not widely available.\(^5\)

To the best of our knowledge, this is the first prospective randomised controlled trial comparing the clinical and radiological results of patients treated with these three bearing surfaces. We have included to this study younger patients (< 65 years; mean age 53 years) who are more active and place greater demands on the prosthetic components. Despite this, there were no statistically significant differences in clinical outcome between the three groups. However, this study clearly demonstrated that the rate of wear of the conventional UHMWPE was nearly three times greater than that of HXLPE. The CoC bearings demonstrated the lowest rate of wear, but were associated with the generation of noise in some patients.

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
31. Beltamy N, Buchanan WW, Goldsmith CH, Campbell J, Sittt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthri-


