The Journal of Bone & Joint Surgery

Journal Club: 4 May 2012
Organiser: Dr. Jemma Kerns
Reviewer: Barian Mohidin
Senior Attendee: Dr. Vivek Mudera (Institute Director of Education)
Institute of Orthopaedics and Musculoskeletal Science Journal Club
Royal National Orthopaedic Hospital, Stanmore

Nikolaou VS, Edwards MR, Bogoch E, Schemitsch EH, Waddell JP.
A prospective randomised controlled trial comparing three alternative bearing surfaces in primary total hip replacement
J Bone Joint Surg [Br] 2012;94-B:459-65.

Introduction
The introduction section of the paper started by explaining the controversy that surrounds bearing surfaces for total hip replacements (THRs). It highlighted the importance of taking revision surgery into account when deciding on the bearing surface combination. The authors explained that prostheses do not last forever, and that the decision as to which bearing surface combination to use is further complicated by younger patients having longer life expectancies and placing greater demands on their prosthesis. Therefore, the choice of the bearing surface combination is heavily influenced by the need for revision. The authors clearly stated that aseptic osteolysis is the most common form of failure of THRs but failed to explain exactly how the wear debris lead to aseptic osteolysis. Although this is a topic that is widely studied and not perfectly understood, a number of mechanisms have been proposed and there is now a general understanding of how wear debris reach the bone-implant interface, and how that subsequently triggers an inflammatory cascade, resulting in the release of TNF-α, a cytokine widely believed to influence osteoclast development. It would have been appropriate to briefly mention the basic understanding and acknowledge the limitations of current knowledge.

The different bearing surfaces were described with references for further explanation. The authors mentioned that highly cross-linked polyethylene has superior wear resistance to conventional ultra-high molecular-weight polyethylene. Moreover, the authors acknowledge that both metal-on-metal and ceramic-on-ceramic show even better wear rates than the previous two bearing surfaces. However, the authors correctly point out that there are currently serious concerns about metal-on-metal hip replacements because of increased metal ion levels in the blood of patients with these replacements. This concern is particularly important in young women as foetal exposure to metal ions has been shown. The authors also noted the drawbacks of ceramic-on-ceramic bearing surfaces: complications such as spontaneous fracture and squeaking. The study lacked a hypothesis, however, the aim was clearly stated and this was to study the clinical and radiological outcomes of a group of patients less than 65 years of age and randomised to receive a ceramic-on-ceramic (CoC), cobalt-chrome on ultra-high molecular-weight polyethylene (UHMWPE) or cobalt-chrome on highly cross-linked polyethylene THR (HXLPE).

Patients and Methods

In a period of approximately two years, 102 hips were operated on from 91 patients. The mean age was 52.7 years. All operations were primary total hip replacements (pTHRs) and they were all done in one hospital by one of three surgeons (co-authors). Each pTHR consisted of an uncemented femoral component and an uncemented acetabular component from Smith & Nephew. Patients were randomised to three intervention groups, each differing by their bearing combinations: 28mm cobalt-chrome head with an ultra-high molecular-weight polyethylene liner, 28mm cobalt-chrome head with a highly cross-linked polyethylene liner, or a 28mm ceramic head with a ceramic acetabular liner. Both the patients and the staff involved in postoperative care were kept blind to the type of bearing surface used. The inclusion criteria were young patients, whom they defined as aged between 18 and 65 years, who were undergoing pTHR. Anyone with a history of hip joint sepsis, primary/secondary hip malignancy, acute fracture of neck of femur, history of autograft/allograft bone was excluded from the study.

Preoperative assessment included history, demographics, comorbidities, physical examination, Harris hip score (HHS), Western Ontario and McMaster Universities osteoarthritis index (WOMAC), Short-Form 12 (SF-12), AP and lateral radiographs. Each patient was also asked to categorise their level of activity as light, moderate or heavy. Patients were assessed clinically and radiologically at 3, 12, 24, 36, 48 and 60 months after surgery. The outcomes measured were the WOMAC, SF-12, HHS and the annual linear wear. Although not explicitly stated, it can be inferred that the former three outcomes make up the primary objective and the latter could be regarded as a secondary objective. The assessor was also kept blind to the bearing surface combination used in each patient. However, the authors acknowledge that the assessor was able to distinguish on the radiographs the CoC bearing surfaces from the other two types of articulation, which used a cobalt-chrome (metal alloy) head component. Differences at 60 months between the three groups in the HHS, WOMAC, SF-12 and the linear wear were assessed using the Kruskal-Wallis test. A p-value of less than 0.05 was considered statistically significant. The study was approved by the University of Toronto Ethics Review Committee.

Results

The three groups were judged to be comparable as differences in baseline characteristics between the groups were statistically insignificant. After 60 months, 97 pTHRs in 87 patients were successfully followed-up. Revision THR was required in two hips from two patients (one for infection and one for peri-prosthetic fracture). Additionally, two patients (three hips) were lost to follow-up. Three patients with the CoC bearing surface reported squeaking starting between 3-5 years postoperatively. At 60 months, the mean WOMAC, SF-12 (except for the mental component) and HHS had improved significantly when compared to preoperative scores in all three intervention groups. Moreover, the differences between the three groups in each of the three primary outcome measures was statistically insignificant so despite the bearing surface used, the patients would still score similarly in the WOMAC and the SF-12 and the HHS. This was mentioned succinctly in few words and was made clear only by the supporting figures and tables. Furthermore, significant differences between the three groups were found with respect to the annual linear wear. The annual linear wear was significantly lower in the CoC group than in the HXLPE group ($p<0.001$), which in turn had significantly lower annual wear rates than the UHMWPE group ($p<0.001$).

Discussion

The findings in this study, with regards to the annual linear wear, were consistent with previous studies. There is now a significant body of evidence that supports the superiority of HXLPE over UHMWPE in wear

rates in the medium term. The authors wisely pointed out that aseptic osteolysis is uncommon with linear wear rates of less than 0.1 mm/year whereas a 100% failure rate has been identified when linear wear rates exceed 0.2 mm/year. The CoC bearing surface produced exceptional wear results; however, three of the 34 patients (8.2%) reported squeaking. Limitations of the study include the number of patients in each group. The authors intended to recruit 45 patients per group so as to give a power greater than 95%. Instead, they managed 36 patients in the UHMWPE group, 32 in the HXLPE group and 34 in the CoC group. Moreover, wear assessment was done using the Dorr and Wan technique. However, the literature favours the Livermore technique which is believed to be more accurate. The authors also acknowledge that radiostereometric analysis is the most accurate tool for *in vivo* assessment of polyethylene wear. The authors drew two firm conclusions from the results. First, HXLPE is more wear-resistant than UHMWPE. Second, CoC has the best wear resistance of the three; however, it does have serious drawbacks evidenced by 8.2% of patients complaining of squeaking.

Study strengths

The study was prospective and randomised. Although the study did not require a control, it could be inferred that the conventional cobalt chrome-on-UHMWPE bearing surface was the standard against which the other two combinations were being compared. There was a good sample size even though the authors fell short of their 45 patients per group target. Despite missing out on 95% power, the authors still had enough 'n' numbers to power the study to at least 80% conveying a maximum type II error of 0.2. The outcome measures used were a good mix of patient reported outcome measures (WOMAC and SF-12), which gave the patient perspective, and physician-completed assessments (HHS), which gave the 'expert' perspective. Each outcome measure offered an alternative viewpoint. Whereas the HHS placed emphasis on physical aspects such as the range of motion and the level of deformity, the WOMAC was more concerned with the practical implications such as the ability to go shopping independently and carry out domestic duties. In comparison, the SF-12 placed greater emphasis on the emotional and physiological wellbeing of the patient and had a series of questions assessing their mental state. The introduction section of the paper was very strong and scientifically sound. The methodology was written with clarity in mind and each step was explained in adequate detail. Overall this study was well-designed and has many strengths.

Improvement points

The overwhelming majority of the patients in this study had osteoarthritis. Few patients had any other disease that required a THR. Therefore, I think it would have been better to only study the patients with osteoarthritis. I say this for two reasons. Firstly, patients with other diseases are not contributing much to the sample size and so it would make for a more meaningful comparison if the authors used patients with the same disease. Secondly, patients with osteoarthritis are likely to score differently in the WOMAC, SF-12 and HHS to those with, for instance, rheumatoid arthritis or avascular necrosis. Therefore, the change in the scores of the WOMAC, SF-12 and HHS are likely to be different after THR, in patients with different diseases. Patients were asked to categorise their level of activity as light, moderate or heavy during the study. I do not think this adds anything of value to the study as each category is subjective and there are only three categories to choose from.

Moreover, Dr. Mudera wisely pointed out that the five year measures are not very informative as THRs are expected to perform for at least 10-15 years. Therefore, it would be expected that no significant differences be found between the three types of bearing surfaces at five years, as each is still performing at an optimal level. He suggested that the study would have benefited from repeating the outcome measures at 10 years, 15 years and 20 years. The discussion section of the paper focussed greatly on the wear results and very

little was mentioned of the WOMAC, SF-12 and HHS outcome measures. Considering this was the primary objective, the authors could have discussed how it related to other work and further explored the implications of their findings. In comparison, the wear results were thoroughly examined with a detailed discussion and comparison to previous work.