Journal Club: 14 July 2010
Chairman: Mr H Sandhu
Organiser: Mr M Sangster
United Hospital, Bath

Metal-backed acetabular components with conventional polyethylene: A review of 9113 primary components with a follow-up of 20 years.
G Hallen, E Dybvik, O Furnes, LI Havelin
J Bone Joint Surg [Br] 2010;92-B:196-201

Reviewers: Mohammed As-Sultany and Nirad Joshi

Introduction

It has long been recognised that the longevity of a total hip prosthesis depends mostly on the survival of its acetabular component. Until about 20 years ago, these components were only ever cemented, but the presence of significant fixation failure due to osteolysis promoted the design and use of uncemented acetabular implants.

Methods

Originality

Available data from trials and reports on the survivorship of uncemented acetabular implants have been unfavourable after a 10-15 year follow-up. The authors of the current study allude to the fact that these implants may have failed due to wear of polyethylene, osteolysis and/or aseptic loosening. This study supersedes those before it because it has a much larger patient population and a longer follow-up period of 20 years.

Aims and hypotheses

The aim of the study was clearly stated as evaluating the performance of primary uncemented metal-backed acetabular components. Although a hypothesis was not stated, the authors mention the findings of other similar studies, which tended to favour cemented acetabular components.

Ethics

All patients included in the Norwegian Arthroplasty Register since 1987 gave their written consent to the collection and use of the data.

Study type / level of evidence

This was a population-based prospective observational study presenting level IV evidence data.

Preliminary statistics

The authors correctly used the Kaplan-Meier survivorship method with associated 95% confidence intervals. Cox multiple regression analyses were also done to adjust for variables such as age, gender, diagnosis and femoral head material and size.

Subject recruitment and criteria

The authors obtained the data from the Norwegian Arthroplasty Register, which was established in 1987. Over a 20-year follow-up period (1987-2007) a total of 110,991 primary THRs had been registered. Only THRs used in more than 400 hips, had modular metal-backed acetabular components, standard UHMWPE and femoral heads made of stainless steel, CoCr alloy or alumina ceramic were included in this study.

Interventions

The above criteria were met by seven different acetabular component designs, which were implanted in 9113 hips of 7937 patients. Although all seven designs were modular and used a metal-backing made of titanium alloy, the components had different coating (porous coating and hydroxyapatite coating) and used different forms of fixation (threaded, press fit +/- screw-fixed). Also, the study does not provide information on the number of hips, which were totally uncemented or hybrids. Furthermore, the polyethylene liners underwent different methods of sterilisation such as gamma or beta irradiation in air, gamma irradiation in inert atmosphere, ethylene oxide and gamma irradiation in vacuum.

Outcome measures

The primary endpoints included revision of the acetabular component due to aseptic loosening and for any cause, which comprised dislocations, infections, polyethylene wear and osteolysis. The secondary endpoints considered results of polyethylene-related failures and material of the femoral head used

Results

The survival analysis of aseptic loosening beyond 12 years ranged from 94% to 100% for all implants apart from Atoll, which had a survival of 81%. In comparison to the Tropic implant (used as the reference), the Atoll implant demonstrated a three times higher risk of revision due to aseptic loosening. From the data in Table III and Figure 1, it is apparent that 91% of all revisions occurred in the Tropic (159 out of 433) and Atoll (234 out of 433) groups. Analysis for the second primary endpoint of acetabular component revision due to any cause beyond 12 years showed poorer results with survival ranging from 90% (Trilogy implant group) to 74% (Atoll implant group). Data presented in Table IV shows that the contemporary acetabular implants with relatively shorter follow-up mainly underwent revision due to dislocation and infection. On the other hand, the older implants with longer follow-up time commonly underwent revision due to polyethylene wear and osteolysis. Figure 3 clearly shows that acetabular component revision occurred mainly due to aseptic loosening and polyethylene wear after about seven years.

The secondary endpoint of acetabular component revision and the material of the femoral head showed alumina to perform better than stainless steel and CoCr alloy as they had relative risks for revision of 1.9 (95% CI 1.6 to 2.2) and 1.8 (95% CI 1.4 to 2.3) respectively.

Discussion

Strengths

Overall, this is well conducted prospective study on a large population of patients with a 20-year follow-up period, which is comparably longer than previously published work. All of the data was extracted from the Norwegian Arthroplasty Register, which is a well-established and reliable database. The initially outlined aim was fulfilled by broadly showing the long-term survivorship of commonly used implants between 1987-2007. Presentation of the data in both Tables and graphs was suitable, with subsequent clear and appropriate statistical analysis.

Limitations

The authors demonstrate good insight into various limitations of this study. The observational nature of the study makes the level of presented evidence weak and prone to bias. The design of the study dictates the inclusion of earlier acetabular implant designs that are no longer used in current practice. The relatively limited data on contemporary acetabular implant designs was confounded by the exclusion of THRs, which were used in less than 400 hips. It is important to note that the study combines data on acetabular implants, which had variable methods of fixation, different coating and used liners that underwent different sterilisation techniques. These may have all influenced the survivorship of the implants. In addition, the study does not clearly state how many of the hips included in the study were totally uncemented or hybrids. It has been shown from previous studies (Hernandez et al, 1994) that polyethylene wear is more common in total uncemented hips than hybrids. Overall, we think that this study only highlights the poor survivorship of two designs (Atoll and Tropic) and makes a generalized conclusion that uncemented acetabular components perform unsatisfactorily. This is despite the fact that some of the contemporary designs, which are still in current practice, have shown excellent results in the medium term follow-up.

The fact that no radiological evaluation was carried in this study highlights that patients with radiological evidence of osteolysis and potentially subtle clinical signs were not included. The various limitations in this study and the improvements of currently used implant designs, means one cannot conclude that the overall performance of metal-back cementless acetabular components is unsatisfactory. Furthermore, the inclusion of data on largely outdated implants may weaken the relevance of obtained results to current orthopaedic practice.

Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: A consequence of excess wear

D. J. Langton, S. S. Jameson, T. J. Joyce, N. J. Hallab, S. Natu, and A. V. F. Nargol
J Bone Joint Surg Br, Jan 2010; 92-B: 38 - 46.

Reviewers: Marshall Sangster, Priyan Landham

Summary
Follow-up study of metal-on-metal joint replacements performed by a single surgeon at one centre.

Level of Evidence
IV

Originality and Design
The authors refer much of their own and other previous published work on metal-on-metal hip failures. Bowser JG, et al (2009) have published work on wear patterns in metal-on-metal bearings. The influence on cup orientation on the wear performance of metal bearings discussed in this paper has previously been published by Angadji et al (2009). The authors published a prospective single-surgeon series on the results of Articular Surface Replacement (ASR) resurfacings in 214 hips in January 2010.

Methods

Study Design
All patients undergoing metal-on-metal arthroplasty implants:

- Birmingham Hip Resurfacing (BHR, Smith & Nephew)
- Articular Surface Replacement (ASR, DePuy)
- Articular Surface Replacement Total Hip Replacement(ASR THR, DePuy)

Patients were enrolled in a prospective study monitoring “clinical, radiological and functional results”. Patients were followed up at 6 weeks, 3 months, 6 months and then annually. Harris Hip score and UCLA activity scores were produced at one year for all patients.

Digital radiographs were performed immediately post op and then at each visit. Radiographic parameters measured by two of the authors included inclination angle and anteversion angle of the acetabular component and articular contact patch to articular rim were calculated for all patients. Chromium and cobalt metal ions were measured in whole blood and serum at a minimum of 12 months post operatively from June 2007 onwards.

A small group of patients presented with groin pain (symptomatic group) and underwent revsion surgery (or were awaiting revision at the time of publication).

At revision, joint fluid was aspirated for metal ion analysis, soft tissue was excised for histopathological examination, and the explants underwent analysis. Patients also underwent metal allergy testing.

Results

155 BHRs were implanted between 2002 and 2004. From 2004 until January 2009, 505 ASRs were implanted (418 ASRs and 87 ASR THRs). 16 patients presented with groin pain and all had an ASR bearing *in situ*. 14 of these patients were women. 13 patients underwent revision surgery and 3 were awaiting revision. Inflammatory markers (WBC, ESR and CRP) were normal in these 16 patients and aspiration revealed sterile effusions. They were diagnosed as having "ARMD", an adverse reaction to metal debris that encompasses "joint failures associated with pain, a large sterile effusion of the hip and or macroscopic necrosis/metallosis". Comparison between the ARMD hips and the asymptomatic ASRs revealed significant differences in terms of femoral size, anteversion angle, whole blood and serum Cr and Co. There was no significant difference in inclination angle of the acetabular components between these two groups. Histological examination of the tissue samples revealed widespread histiocyte infiltration with areas of tissue necrosis consistent with the description of ALVAL (aseptic lymphocyte-dominated vasculitis-associated lesion). Explant analysis showed high out-of-roundness values when compared with the low out-of-roundness values for a revision for an uncomplicated fracture and suggest that this deformation is not caused by the manufacturing process or implantation. Only one ARMD patient showed metal reactivity which was a mild response to Al, Mo and Ni.

Discussion and Conclusions

The authors conclude that the group of patients that developed ARMD had "significantly smaller components, significantly higher acetabular component anteversion, and significantly higher whole blood concentrations of blood and joint chromium and cobalt ions than asymptomatic patients did".
They state that from previous studies, the closer the articular contact patch lies to the rim of the acetabular component, the greater the edge loading and therefore the more articular wear that takes place. The ASR component provides a smaller arc of coverage than the BHR and therefore for matched components there is more likely to be edge loading, and therefore increased wear. Women are more likely to be affected as they have a reduced mean joint size and increased mean acetabular inclination and femoral anteversion from anatomical studies. The latter leading to posterior impingement and microseparation, potentially increasing wear and producing higher ion concentrations.

Limitations

Although this is a prospective study, it is essentially a follow up study or audit of one surgeon's practice in one hospital and therefore constitutes level IV evidence.

There is no true control group and there are no inclusion/exclusion criteria with no hypothesis being tested.

The main limitation is the small number of patients in the study. Only 16 patients developed ARMD and it is difficult to draw conclusions form such a small case series of patients.

With regards to the Explant analysis, the out-of-roughness measurements in the ARMD group are compared with one explanted prostheses (revised for fracture) and so this is not an adequate control. There was no mention of possible manufacturing issues but this could have serious implications.

The BHR group is included in the study to demonstrate there is no "genetic predisposition" in the "same geographical pool" to ARMD. Again, with the ethnic diversity in the UK and the migration of people between regions, this is not a logical conclusion to draw.

There is also the issue of follow-up bias and whether the adequate time has elapsed between surgery and assessment which will be little more than a year in some cases. It is also unclear from the study, the extent to which prosthetic infection was excluded as a cause of pain. Whilst the inflammatory markers were normal and the joint effusions sterile, it is not stated clearly whether tissue was sent for microbiological examination. They have also not stated either way whether they had any joint injections in their study numbers.

Despite a complex histological description we felt it was difficult to ascertain if patients labelled as having an adverse reaction in the presence of pseudo-tumour well described in other papers on metal on metal reactions.

Unanswered questions

1. If as the authors state, that levels of blood metal ions are a "surrogate marker of *in vivo* wear". Table 1 shows serum and whole blood metal ions in BHRs are comparable with ASRs and ASR THRs, why then do BHRs not go on to have problems with ARMD? The authors have not clearly described why there has been no failure of the BHR when ion levels have been very high in some cases as well as other factors they have linked with ASR failure.

2. The authors say there is less edge loading due to greater coverage of the acetabular implant but then why are the metal ions concentrations raised? One or both of these hypotheses is wrong or there is another reason?

3. Is there a difference in the metal ions between the two implants and hence different reactivity?

4. Why do not all ASRs with small components and anteverted positions go on to ARMD?

Summary

This paper was chosen to discuss due to the topical issue of metal ions and failure of metal on metal articulations.
We appreciated the author's honesty in publishing this work and the importance of acknowledging metal on metal failure.

Following a long discussion in the journal club we felt this is a useful study that shows a worrying trend to early failure due to increased wear for smaller sizes, more anteverted acetabular components and with women more at risk.

We did feel that larger randomised control studies are needed or multi-centre data from joint registries need to be accessed to draw firm conclusions.

The group also raised a number of un-answered questions that we felt needed to be answered not necessarily in this paper but for future research on metal on metal bearings.

A prospective, randomised trial comparing closed intramedullary nailing with percutaneous plating in the treatment of distal metaphyseal fractures of the tibia
J Guo, N Tang, H Yang, T Tang
J Bone Joint Surg [Br] 2010;92-B:984-8

Reviewers: Andrew Tasker, Hideki Nagata and Alexandru Mertic

Introduction

The optimum treatment of distal tibial fractures remains controversial. Soft tissue coverage is poor and the fractures often have an intra-articular extension. Minimally invasive fixation techniques aim to achieve early rehabilitation whilst preserving soft tissues but require adequate distal fixation and fracture reduction.

Methods

Study Design

A single centre, multi-surgeon, prospective, randomised trial comparing the treatment of distal tibial fractures by reamed intramedullary nail fixation (IMN) and minimally invasive plate osteosynthesis (MIPO) with a locked compression plate (LCP).

Recruitment and Criteria

Patients were recruited over a two and a half year period from 2005. Inclusion criteria were Orthopaedic Trauma Association (OTA) classified 43-A fractures with a distal fragment of at least 3 cm.

Exclusion criteria included pathological fractures, displaced intra-articular extension, those requiring fibular fixation for syndesmotic disruption, Grade II- III open fractures and diabetics.

Randomisation

Method not specified.

Methods

IMN performed reduction using skeletal traction. Reduction method for MIPO not detailed but referenced. Fibular diastasis assessed intra-operatively. Perioperative care and follow up standardised. Progression of weight bearing status was made on an individual basis during follow up.

Outcomes

Operative time, Intra-operative imaging time, wound problems, time to radiographic union, metal work removal and recovery / return to mobility (assessed with the AOFAS).

Statistics

No power calculation performed. Odds Ratios were calculated with 95% confidence intervals. Parametric outcomes were assessed with unpaired independent t tests. Non-parametric data was assessed with Chi squared and Fisher's exact test. Level of significance set at p value of <0.05.

Hypothesis

That IMN would confer recovery benefits over MIPO

Results

Completeness data

111 patients were randomised. 57 underwent IMN, 54 MIPO. 19 were excluded intraoperatively due to fibular fracture fixation (10 IMN, 9 MIPO). 76% achieved 1 year follow up, equally distributed between both interventions.

Number of patients excluded prior randomisation not specified.

Analysis

Results clearly displayed in tables.

- *Intra-operative outcome measures*: The mean radiation time and operating time were significantly longer in the MIPO group (3.0 vs 2.12mins and 97.9 vs 81.2mins respectively)
- *Post-operative outcome measures:* There was no statistical difference between the two groups with regards to time to bony union (IMN 17.66 weeks, MIPO 17.59 weeks), clinical outcome scores (pain, function, alignment, AOFAS). There was more wound problems with the MIPO group (6 vs 3) although this was deemed not significant and they all went onto heal after 6 weeks.
- *Removal questionnaire:* A similarly high proportion of the patients went onto have implant removal (84.1% IMN vs 92.7 % MIPO).

Conclusions

There was no clinical or radiological difference between the two groups at one-year follow up. Therefore the authors state that both IMN and MIPO can be used safely to treat a distal metaphyseal tibial fracture of OTA type 43-A.

Intra-operative time and radiation exposure were significantly shorter with IMN and the authors have adopted this as their preferred technique.

Relationship with other studies

The authors state that AOFAS scores were similar to other studies although their scores were lower. They had similar rates of infection in the MIPO group with other studies but a much higher implant removal rate (52% vs 92.7%).

Limitations

The authors recognize the delay to surgery with the MIPO of up to 10 days may affect the clinical outcome. They also state that the study is underpowered (24 % drop out rate) and that a longer follow-up period is required.

Operations were performed by senior surgeons but stage of MIPO operative learning curve was not stated. The number of patients with Gustilo type 1 fractures was not specified. Method of randomisation was not stated. The number of patients eligible for the study but excluded was not stated. No reference to local or systemic complications bar wound problems was made, for example venous thromboembolism.

The removal questionnaire was not outlined or validated and this may have unduly influenced the high rates of metalwork removal.

IMN is associated with anterior knee pain yet this is not mentioned in the paper or reflected in the AOFAS.

One might have expected earlier weight-bearing in the IMN group but there was no consistent weight-bearing protocol post operatively for either group. This may have influenced time to union and obviated some of the benefits of IMN.

Lack of a power calculation exposed the study to type II error. A larger study group, earlier initiation of weight-bearing and complete follow up may have found significant differences.

Level of Evidence

Level 1b

Conclusions Justified & Relevance

We feel that the conclusions made from this study are justified. The study is a unique prospective randomised comparative study with high numbers for a single centre with strict inclusion criteria. There was heterogeneity in treatment group demographics. The validated AOFAS outcome score was used and appropriate to the study population.

It concludes that both techniques went onto bony union by 17 weeks with comparable outcomes and without significant morbidity. IMN has a shorter operative and radiation exposure time. These findings are of clinical relevance.

We are concerned by the high rates of metalwork removal and that the study lacked power to show a significant difference in wound problems.

The impact of pre-operative obesity on weight change and outcome in total knee replacement: A prospective study of 529 consecutive patients
MM Dowsey, D Liew, JD Stoney, PF Choong
J Bone Joint Surg [Br] 2010;92-B:513-20

Reviewers: Mark Kemp, Shahid Punwar and Julia Blackburn

Introduction

Most patients presenting for total knee replacement (TKR) are obese. Obesity has been linked to a worse outcome after TKR.

Methods

Originality

A number of previous studies have suggested that weight loss does not occur after TKR but these have been limited in sample size.

Ethics

No institutional review board approval was sought

Aims and hypotheses

The three aims of the study were as follows:

1. To establish the rate of clinically significant weight change in patients 12 months after primary TKR
2. To compare the clinical details and characteristics of patients who lost or gained weight after primary TKR with those who did not
3. To compare the clinical and functional outcome between obese and non-obese patients after TKR

The hypotheses were not clearly stated but were presumed to be:

1. Is there a patient profile in which weight loss or gain can be predicted following primary TKR?
2. Is there a significant difference in outcome between obese and non-obese patients following primary TKR?

Methods

Study type/level of evidence
This was a level II prospective cohort study.

Subject recruitment
All patients admitted for a primary TKR between Jan 2006 and Dec 2007 were eligible for the study. They received standardised treatment and rehabilitation according to hospital protocol.

Exclusions
Patients were excluded from the study if they were undergoing a revision procedure, surgery for neoplastic disease or if they were unable to provide informed consent

Preliminary statistics
A power calculation was performed to determine an appropriate sample size.
Patients were divided into groups according to their BMI for statistical analysis: non-obese < 30 kg/m^2, obese 30 kg/m^2 to 39 kg/m^2 and morbidly obese ≥ 40 kg/m^2.

Results

Outcome measures
1. Significant weight change (+/- 5% of pre-op weight):
There was no difference in the median weight change between the non-obese and obese groups at 12 months following TKR.
73 pts (14%) lost weight, 340 (65%) remained unchanged, 108 (21%) gained weight.
Only 40 (12.6%) of the obese/morbidly obese group lost weight
Increasing age was significantly associated with weight loss at 12 months but otherwise weight loss or weight gain could not be predicted by any other pre-morbid characteristic or by improvements in function and physical health

2. Functional and quality of life (QoL) outcomes
Median International Knee Score (IKS) at 12 months improved in all weight groups but was lower in obese compared to non-obese patients.
There was a 10 point difference in the median IKS change between non-obese and obese/morbidly obese groups, with the change in the functional component of the IKS particularly affected.

No statistical difference in the median pre-operative Short Form 12 (SF-12) scores (physical and mental) between obese and non-obese patients
At 12 months the physical component score improved to a significantly greater degree in the non-obese and obese compared with morbidly obese patients.
No significant difference in the median mental component score in any weight group

3. Adverse events

An adverse event occurred in 20 (35.1%) of morbidly obese, 59 (22.1%) in obese and 30 (14.2%) in non-obese. This was a statistically significant difference (chi-squared p=0.001).

Using the BMI as a continuous variable the odds ratio (OR) for the risk of an adverse event was 1.084 for every unit increase in BMI, adjusted for age and gender

There was also found to be a higher re-admission rate observed in the obese groups

Accuracy and clarity of graphs/tables

There is a very large amount of original data included in the tables which at times we found confusing and difficult to interpret

We wonder whether table III has been mislabelled and should read 'weight gain' in the headings of the second and third columns rather than 'weight loss'

Table II only appears to be displaying the obese and morbidly obese patients although this is not clear in the text or in the title of the table

Valid/reliable

The objective component of the IKS was performed by the consultant/registrar in the clinic. This may have introduced measurement bias as this was not blinded.

Missing data

Four patients died before follow-up at 12 months and 4 patients did not complete the questionnaires at 12 months

Completeness/length of follow-up

Patients followed up to 12 months- improvements in function and QoL occur within 3 months of TKR and therefore 12 months is a reasonable time period over which to expect some weight loss to have occurred

Complete follow-up data available for 521 out of the 529 patients (98.5%) recruited to the study

Statistics

Function, QoL and weight change for BMI groups were compared in a univariate analysis using the rank-sum test. Weight-change groups were compared in a univariate analysis using the chi-squared test for categorical data and the rank-sum test for numerical values. Kruskal-Wallis one way analysis was also used (analysis of variance on ranks when comparing all 3 groups).

These tests seem appropriate to the data.

Conclusions/discussion

Study aims fulfilled/Hypotheses answered?

Yes. They showed that there is no significant change in weight in any group 12 months following primary TKR (with patients tending towards gaining rather than losing weight). Other than age being a predictor of weight loss, they were unable to identify a patient profile in which weight loss or gain can be predicted following primary TKR. They did however demonstrate a significant difference in outcome between obese and non-obese patients.

Conclusions justified
Yes

Relevance
Very relevant to modern day practice with an ever increasing obese population

Insight into shortcomings
They have discussed some but not all of the confounding factors (such as the technical difficulty of the procedure in the obese). They have not specifically discussed any possible shortcomings in their study

Relationship with existing knowledge
They found in this study that despite comparable SF-12 physical health scores before surgery, these scores were significantly lower in morbidly obese patients at 12 months after surgery. The only other study that has examined QoL outcome using SF-12 after TKR specifically in morbidly obese patients found no significant difference in the change in the SF-12 physical component score between morbidly obese and non-morbidly obese patients. This study therefore contradicts the conclusions of that study which felt that morbidly obese patients had the same benefit after TKR as all other patients. Other than this, generally their results support the existing literature.

References complete?
Yes

Pros
- Large sample size
- Identified potential confounding factors and individually analysed these
- Free from selection bias
- Very little loss to follow-up

Cons
- Measurement bias: surgeons in the outpatient department were not blinded to the patients when completing the objective components of the IKS
- They do not go into any detail about the functional outcome and QoL scores (IKS and SF-12)
- No mention of grade of surgeon performing the procedures and no real discussion regarding the technical aspects of the operation and how these are affected by degree of obesity (increase in complications in the obese groups may be solely a reflection of the technical difficulty of the procedure in these patients)
- Confusing tables that are not explained adequately and appear to be mislabelled