

## RESEARCH

# Comparative assessment of implantable hip devices with different bearing surfaces: systematic appraisal of evidence

 OPEN ACCESS

Art Sedrakyan *associate professor*<sup>1,3</sup>, Sharon-Lise T Normand *professor*<sup>2</sup>, Stefan Dabic *fellow, FDA*<sup>3</sup>, Samantha Jacobs *fellow, FDA*<sup>3</sup>, Stephen Graves *professor*<sup>4</sup>, Danica Marinac-Dabic *director of epidemiology, FDA*<sup>3</sup>

<sup>1</sup>Weill Cornell Medical College, New York, NY, USA; <sup>2</sup>Harvard Medical School and Harvard School of Public Health, Boston, MA, USA; <sup>3</sup>Office of Surveillance and Biometrics, Center for Devices and Radiological Health, FDA, Silver Spring, MD, USA; <sup>4</sup>University of Melbourne and the Australian Orthopaedic Association's National Joint Replacement Registry in Adelaide, Australia

## Abstract

**Objective** To determine comparative safety and effectiveness of combinations of bearing surfaces of hip implants.

**Design** Systematic review of clinical trials, observational studies, and registries.

**Data sources** Medline, Embase, Cochrane Controlled Trials Register, reference lists of articles, annual reports of major registries, summaries of safety and effectiveness for pre-market application and mandated post-market studies at the United States Food and Drug Administration.

**Study selection** Criteria for inclusion were comparative studies in adults reporting information for various combinations of bearings (such as metal on metal and ceramic on ceramic). Data search, abstraction, and analyses were independently performed and confirmed by at least two authors. Qualitative data syntheses were performed.

**Results** There were 3139 patients and 3404 hips enrolled in 18 comparative studies and over 830 000 operations in national registries. The mean age range in the trials was 42-71, and 26-88% were women. Disease specific functional outcomes and general quality of life scores were no different or they favoured patients receiving metal on polyethylene rather than metal on metal in the trials. While one clinical study reported fewer dislocations associated with metal on metal implants, in the three largest national registries there was evidence of higher rates of implant revision associated with metal on metal implants compared with metal on polyethylene. One trial reported fewer revisions with ceramic on ceramic compared with metal on polyethylene implants, but data from national registries did not support this finding.

**Conclusions** There is limited evidence regarding comparative effectiveness of various hip implant bearings. Results do not indicate any advantage for metal on metal or ceramic on ceramic implants compared with traditional metal on polyethylene or ceramic on polyethylene bearings.

## Introduction

Every year over 700 000 joint replacements are performed in the United States alone,<sup>1</sup> of which over 270 000 are hip replacements. The annual volume of hip joint replacement is projected to double over the next decade. Moreover, these surgeries are expected to become more expensive with total costs tripling in just five years.<sup>2</sup> While joint replacement is a successful operation and deals with a great public health burden, substantial numbers of patients who receive hip implants require revision surgery within 10 years to replace the implant because of infection, dislocation, wear, instability, loosening, or other mechanical failures.<sup>3-6</sup> The bearing/articulating surface is designed to endure the contact stress and is one of the key design factors to reduce the complications and the chance of revision. Hip implants with metal femoral heads with polyethylene cups as articulating surfaces are associated with low rates of revision.<sup>7</sup> National registries throughout the world continue reporting low risk of revision with this traditional bearing.<sup>8-10</sup> On the other hand, rapid growth of technology introduced several alternative bearings to the market that aimed to further reduce implant wear and subsequently the time to revision surgery. These alternatives include metal on metal and ceramic on ceramic bearings.<sup>11</sup>

Correspondence to: A Sedrakyan, Weill Cornell Medical College, 402 East 67th Street, Suite 223, New York, NY 10065, USA [ars2013@med.cornell.edu](mailto:ars2013@med.cornell.edu)

Extra material supplied by the author (see <http://www.bmj.com/content/343/bmj.d7434?tab=related#webextra>)

**Appendix 1:** Search terms

**Appendix 2:** Details of registries

**Appendix 3:** Identification of studies for inclusion

Metal on metal bearings are particularly attractive to surgeons as they allow use of larger femoral heads (>32 mm v <32 mm) and supposedly reduce the risk of dislocation and improve the functional outcomes in younger patients.<sup>12 13</sup> They were quickly adopted by surgeons and often used even in older patients. In one study, one out of three older patients undergoing hip surgery received metal on metal hip implants.<sup>2</sup> Recently, however, the United Kingdom regulatory agency (Medicines and Healthcare Products Regulatory Agency) alerted the public about severe cases of metallosis (accumulation of metal ions in the tissues) related to the release of metal ions from the implants,<sup>14</sup> and the British Orthopaedic Association developed specific recommendations.<sup>11</sup> Furthermore, in August 2010 Johnson and Johnson recalled over 93 000 metal on metal implants called "ASR."<sup>15 16</sup> The recall received widespread coverage in the *New York Times*,<sup>17</sup> and the *British Medical Journal* published papers related to device regulation, clinician involvement, and the need to develop evidence.<sup>16 18</sup>

In the US, the Food and Drug Administration has been closely monitoring reports of failure of hip implants related to various bearing surfaces and, in November 2009, initiated a comprehensive evaluation and synthesis of evidence of reported outcomes for approved implants. We systematically reviewed the evidence to determine the short and long term outcomes reported by patients undergoing hip replacement and the rates of revision after use of implants with various bearings.

## Methods

### Identification of studies

We worked with the Food and Drug Administration to identify the summaries of safety and effectiveness for all pre-market application trials and relevant Food and Drug Administration mandated post-market studies reporting comparative information for hip implants. Summaries of safety and effectiveness of all ceramic on ceramic hip implants are publicly available as these were the only pre-market application hip replacement devices. We then identified all relevant publications related to pre-market application trials in Medline, Embase, and the Cochrane Controlled Trials Register to learn about long term results (any follow-up) reported. Additionally, we performed a comprehensive search of all other publications in Medline, Embase, and the Cochrane Controlled Trials Register from January 1995 to June 2011. We used all MeSH terms corresponding to hip bearing surface as well as any text words that were applicable to locate the studies. The systematic search strategy is outlined in appendix 1 on bmj.com. We also searched the reference lists of trials and reviews for additional studies. Finally, we reviewed the online annual reports of all registries (within and outside the US) that report information from their registry (see appendix 2 on bmj.com for details).

We limited our study to papers in English. Our inclusion criteria were clinical studies, only adults enrolled, reporting any one of the clinical outcomes of interest (any functional outcomes or revisions, or both), and conventional hip replacement.

### Abstraction of data

The data were independently abstracted and checked multiple times by two abstractors (SD, SJ). The senior author (AS) resolved any discrepancies. Both SD and SJ were trained by using a large learning sample at the stage of abstract screening and manuscript review. One of the senior investigators (AS) reviewed a 20% random sample of excluded studies and all included abstracts for quality control. The process led to 100% agreement among the three authors.

We abstracted information on study design, quality, number of hips (patients), date of procedure, age, sex, number of surgeons and centres, diagnosis (osteoarthritis or avascular necrosis), mean follow-up, percentage follow-up, and information on the manufacturer. The main outcome measures included any functional outcome and occurrence of revision.

### Methodological quality

To evaluate quality of included studies we used four important criteria that lead to conclusions of overall perception of the risk of bias. As masking/blinding of surgeons in these trials is not applicable, we considered "randomised" study description; description of a correct randomisation procedure; masking of patients and outcome assessors; intention to treat analysis; and allocation concealment (methods such as sealed envelope, central telephone) including the time of the announcement of the allocation (in the operating room v before). We used the STROBE criteria to assess the quality of the observational studies.<sup>19</sup> Quality was classified as high (low risk of bias), intermediate (moderate risk of bias), and low (high risk of bias). No formal tool or score was developed. The high quality classification was based on adequate description of randomisation or comparator group, units of randomisation (patient v hip), masking (participants, outcome assessors), and allocation concealment. If studies did not report any of the criteria and described potentially problematic designs (such as questionable exclusion of patients, not describing the group balance regarding prognostic factors), we categorised these studies as low quality. Other classifications (moderate, moderate to high, moderate to low) were based on overall clinical and critical judgments of proximity to high or low quality. Thus, the decision to classify studies into the three categories was based on overall perception of risk of bias after evaluation of the quality criteria, descriptions related to design, and presentations of data.

### Publication bias

While publication bias is applicable to any topic area, we believe that comparative surgical trials in orthopaedics are relatively rare, and papers for such comparisons are likely to be submitted and accepted by journals regardless of the findings. Another possible publication bias might be related to selective reporting of outcomes. Many studies did not report all outcomes of interest to us (particularly revision surgery).

### Qualitative and quantitative analyses

We evaluated clinical heterogeneity (diversity) among the trials by abstracting data on included populations and evaluating definitions of outcomes (often not reported). Functional outcomes included disease specific Harris hip score and general quality of life measures (SF-12). We collected data on number of events (categorical data), scores (continuous data), relative risks (when reported) and their 95% confidence intervals, and P values. We report evidence for alternative bearings (such as metal on metal and ceramic on ceramic) compared with traditional bearings (such as ceramic on polyethylene and metal on polyethylene) bearings. We also report any comparative evidence between two traditional bearings (such as ceramic on polyethylene versus metal on polyethylene) and between alternative bearings (ceramic on ceramic versus metal on metal). Formal meta-analysis was considered in one instance of relatively complete reporting of functional outcome. The results from each study are expressed as a mean difference with 95% confidence intervals and combined with a random effect method.

Statistical heterogeneity was evaluated with  $I^2$  estimates. Rev Man 5.1 software was used for statistical analyses.

## Results

### Studies and populations

We identified and reviewed 3254 abstracts (see appendix 3 on bmj.com). We identified 18 randomised trials and comparative observational studies.<sup>20-40</sup> One study (Capello and colleagues) included both randomised and non-randomised arms and is represented twice in the tables.<sup>30-32 37</sup> There were 3139 patients and 3404 hips enrolled in 18 comparative studies (table 1). The mean age ranged from 42 to 71, and only five out of 18 comparative studies reported standard deviation. The proportion of women was 26-88%; four studies did not report information on sex.<sup>28 34 36 39</sup> None of the studies reported age specific or sex specific analyses, and none reported race or patients' comorbidity, such as diabetes and obesity. Only 13 reported information on diagnosis, and osteoarthritis was the most prevalent diagnosis in 12 studies (range 41-100%). One study predominantly enrolled patients with avascular necrosis,<sup>35</sup> which was the second most common diagnosis in the trials. Reported outcomes included functional outcomes such as disease specific Harris hip score, general quality of life measure (such as SF-12), and revision surgery. Follow-up ranged from three months to 8.1 years (table 1).

### Methodological quality

Four studies were classified as moderate to high quality, five studies as moderate quality, and six studies as low quality. Subgroup analyses based on quality were not informative because of homogenous results and relatively few studies contributing to each comparison. We report the quality information to illustrate the overall lack of appreciation of design and analysis reporting in the comparative studies of bearing surface.

### Patients' outcomes: Harris hip score

The Harris hip score (with higher scores indicating better outcomes) was the most commonly used functional outcome measurement and was reported by 16 out of 18 studies (table 2). Only 10 studies, however, reported both baseline and follow-up measurements; six reported only postoperative scores.

*Metal on metal v traditional bearing surfaces*—All studies had follow-up measurements, but one study did not report preoperative scores (table 2). Four studies compared metal on metal bearing with metal on polyethylene at two years. The reporting was relatively complete in these studies, and the scores were combined. The combined estimate favoured metal on polyethylene bearing. There was a significant 2.4 points higher score (95% confidence interval 4.5 to 0.3) associated with metal on polyethylene compared with metal on metal bearing (fig 1). Three studies reported Harris hip scores beyond two years: two studies compared metal on metal with metal on polyethylene and another with ceramic on polyethylene (fig 2). The individual and combined scores were similar in all three studies for all bearings.

*Ceramic on ceramic v traditional bearing surfaces*—Five studies compared ceramic on ceramic bearings with ceramic on polyethylene bearing. Four studies had both baseline and follow-up measurements, and one study had measurements only after surgery. Two studies compared ceramic on ceramic with metal on polyethylene bearing and had measurements only after surgery. One of the studies was part of the pre-market

application process and had a randomised and non-randomised arm (and hence is represented twice in table 2).<sup>30-32 37</sup> The Harris hip scores were similar at baseline and follow-up in all groups that compared ceramic on ceramic with other bearing surfaces.

*Ceramic on polyethylene v metal on polyethylene*—Two studies compared ceramic on polyethylene with metal on polyethylene, and one of the studies had only follow-up measurement. There were no differences in Harris hip scores.

*Metal on metal v ceramic on ceramic*—One study compared both alternative bearing surfaces and did not find any difference in scores at follow-up.<sup>24</sup> The study included only 28 hips (patients) and was likely to be underpowered to determine any differences between the groups.

### Patients' outcomes: other functional scores

Five studies compared metal on metal with other bearing surfaces (table 3). Four of these compared metal on metal with metal on polyethylene and one study compared metal on metal with ceramic on ceramic. One study reported substantially better physical functioning after surgery associated with metal on polyethylene compared with metal on metal.<sup>28</sup>

One study compared ceramic on ceramic with ceramic on polyethylene, and there was no difference between the groups.

### Patients' outcomes: revision occurrence after hip replacement

#### *Evidence from comparative studies*

Ten comparative evaluations reported information on revisions, and only seven reported data on dislocations. No study specifically reported aseptic loosening.

*Metal on metal v traditional bearing surfaces*—Two studies compared metal on metal with metal on polyethylene bearing and found no difference in occurrence of revisions. Two other studies compared metal on metal with ceramic on polyethylene bearing, and one of these studies<sup>33</sup> reported substantially higher occurrence of dislocation in the ceramic on polyethylene group. Occurrence of dislocation was not different in the other three studies.

*Ceramic on ceramic v traditional bearing surfaces*—Seven comparative studies reported this information. Two studies compared ceramic on ceramic with metal on polyethylene bearing. As described before, one of the studies (Capello and colleagues) included both randomised and non-randomised arms (represented twice in table 4).<sup>30-32 37</sup> This study reported substantially lower occurrence of revision in the ceramic on ceramic arms compared with metal on polyethylene arm. Five studies compared ceramic on ceramic with ceramic on polyethylene and found no qualitative or quantitative differences between the groups in terms of revision occurrence.

*Ceramic on polyethylene v metal on polyethylene*—One comparative study compared ceramic on polyethylene with metal on polyethylene and found no difference in revision occurrence.

*Metal on metal v ceramic on ceramic*—There were no comparative studies discussing occurrence of revision.

### Emerging evidence on revision from national registries

We reviewed the annual reports of 29 national and regional registries. Five national registries<sup>8-10 41 42</sup> including over 830 000 surgeries reported information on bearing surface, and one US study including over 57 000 surgeries reported data from the

Medicare database.<sup>2</sup> Some of the large national registries, such as the Swedish registry,<sup>43</sup> reported information on every implant used in the country but did not stratify by bearing surface. National registries provided important quantitative data, but they varied in their methods of risk adjustment. While we considered national registry reports as comparative studies, we did not evaluate quality because of lack of information on methods.

### Main findings

Metal on metal was associated with higher occurrence of revision compared with metal on polyethylene in the adjusted analyses of three national registries: Australian, New Zealand, and England and Wales National Registries (including over 720 000 patients).<sup>8-10</sup> Three other smaller registries (including the Center for Medicare and Medicaid Services (CMS) database) did not report such results (fig 3⇓).

Ceramic on ceramic was associated with a higher occurrence of revision than metal on polyethylene in the adjusted analyses of the New Zealand National Registry.<sup>9</sup> Five registries (including the CMS database) did not report any difference.

Ceramic on polyethylene was associated with a higher occurrence of revision than metal on polyethylene in the New Zealand National Registry but a lower occurrence in the England and Wales National Registry. Three other national registries did not report any difference (ceramic on polyethylene implant category was not available in the CMS database).

### Discussion

Functional outcomes, traditionally thought of as primary effectiveness outcomes, were no different among patients receiving hip replacements with various bearings. We conducted this systematic evidence review as part of the new project initiated by the Food and Drug Administration to create a framework for post-market evaluation of orthopaedic implants. This systematic evaluation helps to deal with evidence gaps in comparisons of hip bearing surfaces.

We found some evidence for lower Harris hip scores (by 2.4 points) at two year follow-up for metal on metal bearing compared with metal on polyethylene. There was also evidence from one trial that patients receiving metal on metal implants have a lower quality of life (functional component) than patients receiving metal on polyethylene bearing.<sup>28</sup> Though these differences are small and might not be clinically relevant, they are still valuable findings to be considered for future hypothesis generation.

Evidence on implant revision did not favour metal on metal implants. While one study reported fewer dislocations associated with metal on metal implants, three other trials did not support this finding, and the three largest national registries reported substantially higher occurrence of revision associated with metal on metal implants compared with metal on polyethylene bearing (fig 2⇓).

One medium size comparative clinical trial reported a substantially lower rate of revision associated with ceramic on ceramic implants compared with metal on polyethylene implants.<sup>30</sup> This conflicts with results reported in national registries (fig 2⇓). Five national data sources/registries did not report any difference between the ceramic on ceramic and traditional bearings, and a large registry report from New Zealand favoured ceramic on polyethylene (traditional bearing) compared with ceramic on ceramic surfaces.

### Comparison with other studies

To our knowledge this is the first systematic appraisal of assessment of clinical outcomes after hip replacement with various bearing surfaces. We did not summarise the evidence related to metal sensitivity or toxicity, but our findings related to clinical outcomes of metal on metal implants are augmented by reports that show severe metallosis associated with such implants.<sup>14</sup> Concerns regarding the effects of metal ions have been noted after a landmark paper by MacDonald et al<sup>36</sup> that noted a 5.3-fold increase in erythrocyte cobalt, a 35-fold increase in urinary cobalt, and a 17-fold increase in urinary chromium concentrations in patients who received metal on metal hip replacements compared with those who received metal on polyethylene.

### Strengths and limitations of study

We found 18 comparative studies of bearing surface and six national data source/registry reports and comprehensively summarised the strengths and limitations of current data. The overall quality of reporting in the comparative studies was less than adequate in the trials of bearing surface. Many investigators did not report their power to detect minimally important differences among groups and often omitted standard deviations and range of scores observed. In addition, few studies were adequately powered or reported subgroup analyses by age, sex, underlying aetiology, or femoral head size. Classification of polyethylene inserts used in hip replacement also needs harmonisation; the various terms used to define cross linked polyethylene hinder subgroup analyses. Our review was limited to English language publications, and there is chance of publication bias. There was substantial heterogeneity in registry evidence. Formal data pooling from registries will require harmonisation of data definitions and analytical methods.

### Policy implications

Changing technology, the need to have large numbers of patients, long term follow-up to establish the safety, and short and long term outcomes reported by patients to establish the effectiveness all lead to a lack of strong evidence in orthopaedics. For example, patients who receive hip implants often require revision surgery within 10 years, when the implants fail because of infection, dislocation, wear, instability, loosening, or other mechanical failures. In addition, clinical trials assessing bearing surfaces are conducted in relatively unique environments defined by skilled surgeons operating at high volume centres and often concentrating on clinical and radiological indirect measures of device safety and effectiveness. In general this is a reflection of the current state of research in implantable devices, where comparative trials are rare and often not applicable.<sup>44</sup>

Registries are likely to fill the evidence gap in the immediate future. Large registries or networks of registries capturing various orthopaedic devices are particularly important for evaluation of comparative outcomes and active surveillance. Only large longitudinal multinational registries can provide denominator data for adverse events related to specific implants and allow proper conduct of comparative safety and effectiveness studies, particularly for rare end points. Current registry annual reports that include comparative evaluations, however, are not harmonised in terms of methods and reporting. Accordingly, the findings are tentative. To deal with this limitation the US Food and Drug Administration Center for Devices and Radiologic Health started an important initiative called the International Consortium of Orthopaedic Registries,

which aims to build the foundations for a worldwide research consortium of orthopaedic registries. The consortium represents more than 15 countries that have existing registries with a mission to improve the safety and effectiveness of orthopaedic devices and procedures through collaboration. Currently, these international registries included more than 3 500 000 orthopaedic surgeries capturing all implantable devices on the market.

## Conclusions

There is limited evidence regarding comparative effectiveness of various hip implant bearings, and the results do not indicate any advantage for metal on metal or ceramic on ceramic implants compared with traditional bearings. A large and high quality randomised controlled trial of bearing surfaces in total hip replacement needs to be conducted before any claims of benefit are made. Until then national registries provide important real world data that are critical for the safety and future comparative safety and effectiveness evaluation.

Contributors: AR and DM-D were responsible for study design and concept. AR, DM-D, SD, and SJ acquired the data, which were analysed by AR, SG, S-LTN, SD, and SJ. AR, SG, DM-D, and S-LTN interpreted the results. AR drafted the manuscript and is guarantor. AR and S-LTN provided statistical expertise. AR and DM-D provided administrative, technical, or material support. All authors critically revised the manuscript and revision for important intellectual content.

Funding: The study is funded by the US Food and Drug Administration Center for Devices and Radiological Health. The authors accept full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish. The opinions expressed in this article are those of the authors and not necessarily those of the Food and Drug Administration.

Competing interests: All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

- Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am* 2007;89:780-5.
- Bozic KJ, Ong K, Lau E, Kurtz SM, Vail TP, Rubash HE, et al. Risk of complication and revision total hip arthroplasty among Medicare patients with different bearing surfaces. *Clin Orthop Relat Res* 2010;468:2357-62.
- Hooper GJ, Rothwell AG, Stringer M, Frampton C. Revision following cemented and uncemented primary total hip replacement: a seven-year analysis from the New Zealand Joint Registry. *J Bone Joint Surg Br* 2009;91:451-8.
- Jameson SS, Lees D, James P, Serrano-Pedraza I, Partington PF, Muller SD, et al. Lower rates of dislocation with increased femoral head size after primary total hip replacement: a five-year analysis of NHS patients in England. *J Bone Joint Surg Br* 2011;93:876-80.
- Pedersen AB, Svendsen JE, Johnsen SP, Riis A, Overgaard S. Risk factors for revision due to infection after primary total hip arthroplasty. A population-based study of 80,756 primary procedures in the Danish Hip Arthroplasty Registry. *Acta Orthop* 2010;81:542-7.
- Mullins MM, Norbury W, Dowell JK, Heywood-Waddington M. Thirty-year results of a prospective study of Charnley total hip arthroplasty by the posterior approach. *J Arthroplasty* 2007;22:833-9.
- Soderman P, Malchau H, Herberts P, Zugner R, Regner H, Garellick G. Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the Swedish National Total Hip Arthroplasty Register. *Acta Orthop Scand* 2001;72:113-9.
- Australian National Registry. [www.dmac.adelaide.edu.au/aoanrr/index.jsp](http://www.dmac.adelaide.edu.au/aoanrr/index.jsp).
- New Zealand National Registry. [www.cdhb.govt.nz/njr/](http://www.cdhb.govt.nz/njr/).
- UK National Registry. [www.njrcentre.org.uk/njrcentre/default.aspx](http://www.njrcentre.org.uk/njrcentre/default.aspx).
- Skinner J, Kay P. Commentary: Metal on metal hips. *BMJ* 2011;342:d3009.
- Cuckler JM, Moore KD, Lombardi AV, Jr, McPherson E, Emerson R. Large versus small femoral heads in metal-on-metal total hip arthroplasty. *J Arthroplasty* 2004;19(suppl 3):41-4.

- Delaunay CP, Bonomet F, Clavert P, Laffargue P, Migaud H. THA using metal-on-metal articulation in active patients younger than 50 years. *Clin Orthop Relat Res* 2008;466:340-6.
- Medical Device Alert: All metal-on-metal (MoM) hip replacements (MDA/2010/033). [www.mhra.gov.uk/Publications/Safetywarnings/MedicalDeviceAlerts/CON079157](http://www.mhra.gov.uk/Publications/Safetywarnings/MedicalDeviceAlerts/CON079157).
- J&J unit recalls 93,000 hip implant systems. [http://money.cnn.com/2010/08/26/news/companies/johnson\\_deputy\\_hipsystem\\_recall/index.htm](http://money.cnn.com/2010/08/26/news/companies/johnson_deputy_hipsystem_recall/index.htm).
- Cohen D. Out of joint: the story of the ASR. *BMJ* 2011;342:d2905.
- Meier B. With warning, a hip device is withdrawn. *New York Times* 2001 March 10.
- Godlee F. The trouble with medical devices. *BMJ* 2011;342:d3123.
- Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening of reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370:1453-7.
- Hamilton WG, McAuley JP, Dennis DA, Murphy JA, Blumenfeld TJ, Politi J. THA with Delta ceramic on ceramic: results of a multicenter investigational device exemption trial. *Clin Orthop Relat Res* 2010;468:358-66.
- Lombardi AV Jr, Berend KR, Seng BE, Clarke IC, Adams JB. Delta ceramic-on-alumina ceramic articulation in primary THA: prospective, randomized FDA-IDE study and retrieval analysis. *Clin Orthop Relat Res* 2010;468:367-74.
- Ochs U, Ilchmann T, Ochs BG, Marx J, Brunnhuber K, Luem M, et al. EBRA migration patterns of the Plasmacup with ceramic or polyethylene inserts: a randomised study. *Z Orthop Unfall* 2007;145(suppl 1):S20-4.
- Poggie RA, Turgeon TR, Coutts RD. Failure analysis of a ceramic bearing acetabular component. *J Bone Joint Surg Am* 2007;89:367-75.
- Grubl A, Weissinger M, Brodner W, Gleiss A, Giurea A, Gruber M, et al. Serum aluminium and cobalt levels after ceramic-on-ceramic and metal-on-metal total hip replacement. *J Bone Joint Surg Br* 2006;88:1003-5.
- Kawate K, Ohmura T, Kawahara I, Tamai K, Ueha T, Takemura K. Differences in highly cross-linked polyethylene wear between zirconia and cobalt-chromium femoral heads in Japanese patients: a prospective, randomized study. *J Arthroplasty* 2009;24:1221-4.
- Von Schewelov T, Sanzen L, Onsten I, Carlsson A, Besjakov J. Total hip replacement with a zirconium oxide ceramic femoral head: a randomised roentgen stereophotogrammetric study. *J Bone Joint Surg Br* 2005;87:1631-5.
- Dahlstrand H, Stark A, Anissian L, Hailer NP. Elevated serum concentrations of cobalt, chromium, nickel, and manganese after metal-on-metal alloarthroplasty of the hip: a prospective randomized study. *J Arthroplasty* 2009;24:837-45.
- Engh CA Jr, MacDonald SJ, Sritulanondha S, Thompson A, Naudie D, Engh CA. 2008 John Charnley award: metal ion levels after metal-on-metal total hip arthroplasty: a randomized trial. *Clin Orthop Relat Res* 2009;467:101-11.
- Zijlstra WP, Cheung J, Sietsma MS, van Raay JJ, Deutman R. No superiority of cemented metal-on-metal vs metal-on-polyethylene THA at 5-year follow-up. *Orthopedics* 2009;32:479.
- Capello WN, D'Antonio JA, Feinberg JR, Manley MT, Naughton M. Ceramic-on-ceramic total hip arthroplasty: update. *J Arthroplasty* 2008;23(7 suppl):39-43.
- D'Antonio JA, Capello WN, Manley MT, Naughton M, Sutton K. A titanium-encased alumina ceramic bearing for total hip arthroplasty: 3- to 5-year results. *Clin Orthop Relat Res* 2005;441:151-8.
- D'Antonio J, Capello W, Manley M, Naughton M, Sutton K. Alumina ceramic bearings for total hip arthroplasty: five-year results of a prospective randomized study. *Clin Orthop Relat Res* 2005;436:164-71.
- Clarke MT, Lee PT, Villar RN. Dislocation after total hip replacement in relation to metal-on-metal bearing surfaces. *J Bone Joint Surg Br* 2003;85:650-4.
- Dorr LD, Long WT, Siritanni L, Campana M, Wan Z. The argument for the use of Metasul as an articulation surface in total hip replacement. *Clin Orthop Relat Res* 2004;429:80-5.
- Kim YH, Kim JS, Choi YW, Kwon OR. Intermediate results of simultaneous alumina-on-alumina bearing and alumina-on-highly cross-linked polyethylene bearing total hip arthroplasties. *J Arthroplasty* 2009;24:885-91.
- MacDonald SJ, McCalden RW, Chess DG, Bourne RB, Rorabeck CH, Cleland D, et al. Metal-on-metal versus polyethylene in hip arthroplasty: a randomized clinical trial. *Clin Orthop Relat Res* 2003;406:282-96.
- Capello WN, D'Antonio JA, Feinberg JR, Manley MT. Alternative bearing surfaces: alumina ceramic bearings for total hip arthroplasty. *Instr Course Lect* 2005;54:171-6.
- Bascarevic Z, Vukasinovic Z, Slavkovic N, Dulic B, Trajkovic G, Bascarevic V, et al. Alumina-on-alumina ceramic versus metal-on-highly cross-linked polyethylene bearings in total hip arthroplasty: a comparative study. *Int Orthop* 2010;34:1129-35.
- Lewis PM, Al-Belooshi A, Olsen M, Schemmich EH, Waddell JP. Prospective randomized trial comparing alumina ceramic-on-ceramic with ceramic-on-conventional polyethylene bearings in total hip arthroplasty. *J Arthroplasty* 2010;25:392-7.
- Jacobs M, Gorab R, Mattingly D, Trick L, Southworth C. Three- to six-year results with the Ultima metal-on-metal hip articulation for primary total hip arthroplasty. *J Arthroplasty* 2004;19(suppl 2):48-53.
- Italian Register of Orthopaedic Prosthetic Implants. <https://ripo.cineca.it/>.
- Paxton E, Namba RS, Maletis GB, Khatod M, Yue EJ, Davies M, et al. A prospective study of 80,000 total joint and 5000 anterior cruciate ligament reconstruction procedures in a community-based registry in the United States. *J Bone Joint Surg* 2010;92:117-32.
- Swedish National Registry. [www.jru.orthop.gu.se/](http://www.jru.orthop.gu.se/).
- Sedrakyan A, Marinac-Dabic D, Normand SL, Mushlin A, Gross T. A framework for evidence evaluation and methodological issues in implantable device studies. *Med Care* 2010;48(6 suppl):S121-8.

Accepted: 14 November 2011

Cite this as: *BMJ* 2011;343:d7434

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

**What is already known on this topic**

There have been severe cases of accumulation of metal ions in tissues of patients with metal on metal hip implants  
 Metal on metal and ceramic on ceramic hip implants might not be associated with any advantage compared with traditional bearings such as metal on polyethylene or ceramic on polyethylene

**What this study adds**

Disease specific functional outcomes such as Harris hip score and general quality of life measures were no different between patients with metal on metal or ceramic on ceramic hip implants compared with traditional hip implants  
 There is some evidence of higher rates of revision surgery associated with metal on metal implants compared with metal on polyethylene implants  
 While an investigational device exemption trial reported fewer revisions associated with ceramic on ceramic implants, the emerging evidence from national registries does not support these results

**Tables**

**Table 1 | Baseline data in comparative studies of various bearing surfaces in hip replacement systems**

Study	Design	Study quality	No of hips (patients)	Date of procedure	Site, surgeon, centre	Mean (SD, range) age (years)	% women	% osteo-arthrosis (avascular necrosis)	Mean follow-up (years)	(%) follow-up	Manufacturer
<b>Metal on polyethylene v metal on metal</b>											
Dahlstrand, 2009, Sweden <sup>27</sup>	Randomised	Moderate	54 (54)	Oct 1998-Jan 2001	Single surgeon, single centre	65 (NA, 40-75)	50	100 (0)	24 months (22-27)	100	Zimmer (previously Sulzer)
Engh, 2009, US <sup>28</sup>	Randomised	Moderate	91 (91)	Oct 2003-Oct 2005	NA	NA (40-80)	NA	NA	NA (6 months-2 years)	86	DePuy Orthopaedics
Jacobs, 2004, US <sup>40</sup>	Randomised	Moderate	171 (171)	March 1997-July 2000	Multi-surgeon, multicentre	54.4 (NA)	43	66 (20)	43 months	72	Johnson & Johnson
MacDonald, 2003, US <sup>36</sup>	Randomised	Moderate	41 (41)	March 1998- Nov 1999	Multi-surgeon, multicentre	NA	NA	NA	3.2 (2.2-3.9)	98	BioMet
Zijlstra, 2009, Netherlands <sup>29</sup>	Randomised	Moderate	200 (195)	NA	Multi-surgeon, multicentre	71 (8, NA)	79	NA	5.6 (4.2-7.1)	97	Biomet
<b>Ceramic on polyethylene v metal on metal</b>											
Clarke, 2003, UK <sup>33</sup>	Comparative	Moderate	249 (229)	C-P 1995-8; M-M 1998-2002	Single surgeon, single centre	56.1 (NA, 24-69)	51	100 (0)	3 months (NA)	100	Johnson & Johnson
Dorr, 2004, US <sup>34</sup>	Randomised	Low	615 (561)	Aug 1999 start	Multi-surgeon, multicentre	60 (NA, 27-85)	NA	NA	5 (NA)	NA	Zimmer (previously Sulzer)
<b>Ceramic on ceramic v metal on polyethylene</b>											
Bascarevic, 2010, Serbia <sup>38</sup>	Randomised	Moderate to high	157 (150)	Jan 2003-April 2008	Multi-surgeon, single centre	54.8 (6.8)	26	84 (5)	51 months		Zimmer
Capello, 2008, <sup>30</sup> D'Antonio, 2005, US <sup>31</sup>	Randomised	Moderate to high	289 (278)	Oct 1996-Oct 1998	Multi-surgeon, multicentre	53 (NA)	36	82 (15)	8 (5-NA)	93	Stryker Orthopaedics
Capello "partial," 2008, addition of 4th system, Trident, US <sup>30</sup>	Comparative	Moderate to high	186 (174)	Sep 1999-Sep 2000	Multi-surgeon, multicentre	51.8 (NA)	31	NA	6.5 (5-NA)	NA	Stryker Orthopaedics
<b>Ceramic on ceramic v ceramic on polyethylene</b>											
Hamilton, 2010, US <sup>20</sup>	Randomised	Moderate to high	264 (263)	2003	Multi-surgeon, multicentre	56.7 (NA, 20-75)	48	88 (6)	31 months (21-49 months)	88	CeramTec, DePuy, DePuy-Johnson & Johnson
Kim, 2009, South Korea <sup>35</sup>	Randomised	Moderate to high*	200 (100)	Jan 2000-Jan 2002	Single surgeon, single centre	45.3 (NA, 21-49)	34	10 (63)	5.6 (5-7)	100	DePuy

Table 1 (continued)

Study	Design	Study quality	No of hips (patients)	Date of procedure	Site, surgeon, centre	Mean (SD, range) age (years)	% women	% osteo-arthritis (avascular necrosis)	Mean follow-up (years)	(%) follow-up	Manufacturer
Lewis, 2010, Canada <sup>20</sup>	Randomised	Moderate	56 (55)	Oct 1997-Oct 1999	1-2 surgeons, single centre	42.1 (8.0)	NA	41 (11)	96 months (1-10)	93	Wright Medical
Lombardi, 2010, US <sup>21</sup>	Randomised	Low	89 (80)	June 2000	Single surgeon	58 (NA)	46	84 (9)	6.1 (2.1-9.0)	NA	BioMet
Ochs, 2007 Germany <sup>22</sup>	Randomised	Moderate to high	66 (66)	1997-9	Multi-surgeon, single centre	58.7 (7.8, NA)	42	53 (NA)	8.1 (7.1-9.2)	65	Braun-Aesculap, CeramTec
Poggie, 2007, US <sup>23</sup>	Randomised	Low	472 (429)	Oct 1999-Jan 2003	Multi-surgeon	54 (12, 23-76)	45	77 (17)	2	99	Zimmer/CeramTec
<b>Ceramic on polyethylene v metal on polyethylene</b>											
Kawate, 2009, Japan <sup>25</sup>	Randomised	Low	62 (60)	July 2001-July 2003	Multi-surgeon	61.1 (NA, 30-78)	88	69 (18)	5 (4.8-5.4)	100	Japan Medical Material
von Schewelov, 2005, Sweden <sup>26</sup>	Randomised	Low	114 (114)	1996-8	Multi-surgeon, multicentre	64 (NA, 50-76)	54	NA	5 (NA)	92	DePuy, DePuy-Johnson & Johnson
<b>Ceramic on ceramic v metal on metal</b>											
Grubl, 2006, Austria <sup>24</sup>	Randomised	Low	28 (28)	Nov 2000-Feb 2002	NA	62.2 (NA, 32-81)	64	93 (7)	1 (NA)	NA	Zimmer, CeramTec

NA=not available (not applicable or not provided).

\*In this study each person served as his/her own control (bilateral procedures). Functional outcomes are challenging to determine but potentially possible. Uncertainty about surgeon's involvement in functional outcome measurement leads to less than perfect quality classification.

**Table 2 | Harris hip score in comparative studies of bearing surfaces. Figures are mean scores (higher scores indicating better outcomes) (SD and range when available) for longest period of follow-up reported**

Study	Metal on metal		Ceramic on ceramic		Metal on polyethylene		Ceramic on polyethylene	
	Preoperative	Postoperative	Preoperative	Postoperative	Preoperative	Postoperative	Preoperative	Postoperative
Dahlstrand, 2009*, Sweden <sup>27</sup>	38.1 (10.9)	92.8 (12.9)	NA	NA	38.9 (12.7)	95.9 (4.8)	NA	NA
Engh, 2009, US <sup>28</sup>	Done but NA	93.5 (9)	NA	NA	Done but NA	96 (6)	NA	NA
Jacobs, 2004, US <sup>40</sup>	42	95.4 (65-100)	NA	NA	43	96.1 (65-100)	NA	NA
MacDonald, 2003, US <sup>36</sup>	46.5 (13.4)	91.6 (11.5)	NA	NA	46.6 (12.4)	89.2 (12.5)	NA	NA
Zijlstra, 2009†, Netherlands <sup>29</sup>	48 (15)	90 (7)	NA	NA	46 (13)	87 (13)	NA	NA
Dorr, 2004, US <sup>34</sup>	NA	93.5 (12, 14-100)	NA	NA	NA	NA	NA	93.5 (11, 41-100)
Bascarevic, 2010, Serbia <sup>38</sup>	NA	NA	NA	95.1	NA	93.8	NA	NA
Capello, 2008, <sup>30</sup> D'Antonio, 2005, <sup>31</sup> US	NA	NA	NA	96.35	NA	96.4	NA	NA
Capello, 2008, addition of 4th system, Trident, US <sup>30</sup>	NA	NA	NA	95.7	NA	96.4	NA	NA
Hamilton, 2010, US <sup>20</sup>	NA	NA	50.6	94.4	NA	NA	50.7	93.8
Kim, 2009, South Korea <sup>35</sup>	NA	NA	39 (6-47)	93 (85-100)	NA	NA	41 (9-51)	94 (86-100)
Lombardi, 2010, US <sup>21</sup>	NA	NA	51 (6-68)	90 (50-100)	NA	NA	48 (6-69)	92 (49-100)
Ochs, 2007, Germany <sup>22</sup>	NA	NA	NA	91 (21.6)	NA	NA	NA	89 (12.1)
Poggie, 2007, US <sup>23</sup>	NA	NA	45 (14-91)	92 (36-100)	NA	NA	43 (10-78)	93 (51-100)
Kawate, 2009, Japan <sup>25</sup>	NA	NA	NA	NA	47 (28-77)	91 (60-100)	43 (15-75)	89 (58-100)
von Schewelov, 2005, Sweden <sup>26</sup>	NA	NA	NA	NA	NA	91.7 (11.2)	NA	91.3 (10.5)
Grubl, 2006, Austria <sup>24</sup>	50.4 (15.1-71.4)	85.1 (61.1-100)	58.4 (24.4-82.3)	86.6 (41.5-100)	NA	NA	NA	NA

NA=not available (not applicable or not provided).

\*Significant difference favouring metal on polyethylene compared with metal on metal at 12 months but not at 24 months after surgery.

†Measurements at two and five years. Only score at five years reported here (fig 2 shows both two and five year scores).

**Table 3| Other functional outcomes in comparative studies of various bearing surfaces in hip replacements. Figures are mean (SD) scores on each scale**

Study	Metal on metal		Ceramic on ceramic		Metal on polyethylene		Ceramic on polyethylene	
	Preoperative	Postoperative	Preoperative	Postoperative	Preoperative	Postoperative	Preoperative	Postoperative
Dahlstrand, 2009, Sweden <sup>27</sup>								
SF-36: PF	32.6 (22.6)	67.2 (26.9)*	NA	NA	40 (24.5)	77 (19.3)*	NA	NA
SF-36: BP	33.3 (19.4)	67.9 (31.7)*	NA	NA	39.3 (19.2)	76.7 (27.9)*	NA	NA
Engh, 2009, US <sup>28</sup>								
WOMAC	NA	88 (15)	NA	NA	NA	89 (19)	NA	NA
PCS (SF-12)	NA	46.2 (11.6)†	NA	NA	NA	51 (9)†	NA	NA
MCS (SF-12)	NA	56.8 (7)‡	NA	NA	NA	55 (6)‡	NA	NA
MacDonald, 2003, US <sup>36</sup>								
WOMAC	58.7 (15.3)	17.3 (15.1)	NA	NA	59.1 (14.5)	19.9 (18.9)	NA	NA
PCS SF-12	32.8 (10.2)	47.2 (9.6)	NA	NA	31.1 (4.7)	45.5 (11.7)	NA	NA
MCS SF-12	53.5 (9.6)	55.9 (6.2)	NA	NA	53.6 (11.6)	53.3 (11.7)	NA	NA
Zijlstra, 2009, Netherlands <sup>29</sup>								
Oxford hip score	40 (8)	19 (8)	NA	NA	40 (8)	18 (8)	NA	NA
Lewis, 2010, Canada <sup>39</sup>								
St Michael's score	NA	NA	15.8	22.9	NA	NA	16.2	22.8
Grubl, 2006, Austria <sup>24</sup>								
UCLA activity score	NA	7/10 points§	NA	7/10 points§	NA	NA	NA	NA

UCLA=University of California, Los Angeles; PF=physical function; BP=bodily pain; PCS=physical component score; MCS=mental component score; WOMAC=Western Ontario and McMaster Universities activity score; NA=not available (not applicable or not provided).

\*At 24 months.

†Difference significant (P=0.03). Mean score for metal on metal calculated from two groups (28 mm and 36 mm head sizes).

‡Difference not significant. Mean score for metal on metal calculated from two groups (28 mm and 36 mm head sizes).

§At 12 months.

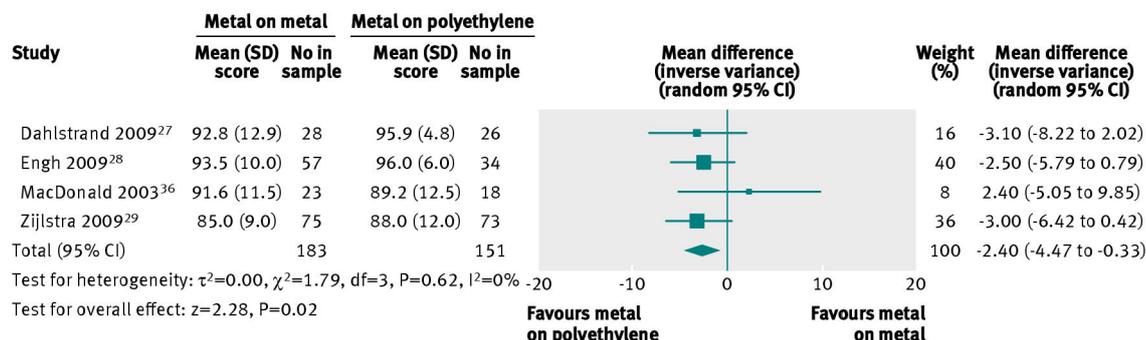
**Table 4** Revision surgery and reported dislocation in comparative studies of various bearing surfaces in hip replacements. Figures are numbers (percentages)

Study	Metal on metal		Ceramic on ceramic		Metal on polyethylene		Ceramic on polyethylene	
	Revision surgery	Reported dislocation	Revision surgery	Reported dislocation	Revision surgery	Reported dislocation	Revision surgery	Reported dislocation
Jacobs, 2004, US <sup>40</sup>	1/95 (1.1)	1/95 (1.1)	NA	NA	1/76 (1.3)	1/76 (1.3)	NA	NA
Zijlstra, 2009, Netherlands <sup>29</sup>	3/98 (2.9)	NA	NA	NA	1/92 (1.1)	NA	NA	NA
Clarke, 2003, UK <sup>33</sup>	NA	1/109 (0.9)*	NA	NA	NA	NA	NA	9/140 (6.4)*
Dorr, 2004, US <sup>34</sup>	10/311 (3.2)	12/304 (3.9)	NA	NA	NA	NA	8/304 (2.6)	12/304 = 3.9)
Bascarevic, 2010, Serbia <sup>38</sup>	NA	NA	2/82 (2.4)	1/82 (1.2)	2/75 (2.7)	2/75 (2.7)	NA	NA
Capello, 2008, <sup>30</sup> D'Antonio, 2005, US <sup>31</sup>	NA	NA	6/349 (1.7)*	10/349 (2.9)	10/165 (6.1)*	7/165 (4.2)	NA	NA
Capello, 2008, addition of 4th system, Trident, US <sup>30</sup>	NA	NA	1/186 (0.5)*	NA	10/165 (6.1)*	7/165 (4.2)	NA	NA
Hamilton, 2010, US <sup>20</sup>	NA	NA	4/177 (2.3)	5/177 (2.8)	NA	NA	2/87 (2.3)	4/87 (4.6)
Kim, 2009, South Korea <sup>35</sup>	NA	NA	NA	1/100 (1)	NA	NA	NA	2/100 ( 2)
Lewis, 2010, Canada <sup>39</sup>	NA	NA	1/30 (3.3)	NA	NA	NA	1/26 (3.8)	NA
Lombardi, 2010, US <sup>21</sup>	NA	NA	3/44 (6.8)	NA	NA	NA	3/45 (6.7)	NA
Poggie, 2007, US <sup>22</sup>	NA	NA	4/315 (1.3)	NA	NA	NA	2/157 (1.3)	NA
von Schewelov, 2005, Sweden <sup>26</sup>	NA	NA	NA	NA	1/58 (1.7)	NA	3/56 (5.3)	NA

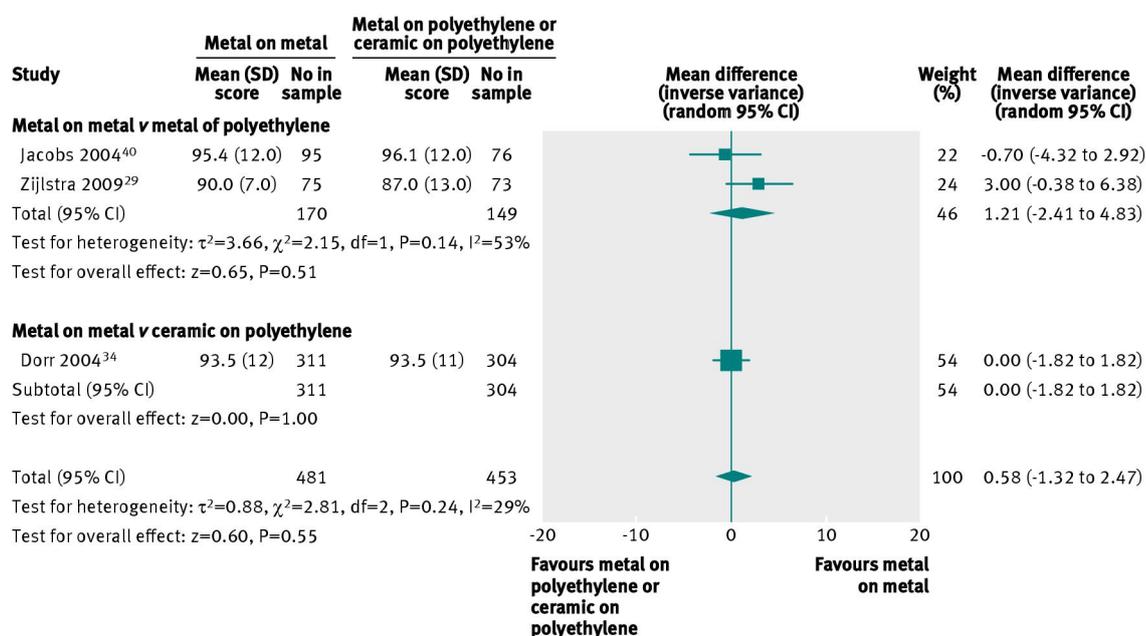
NA= not available (not applicable or not provided).

\*Significant group difference.

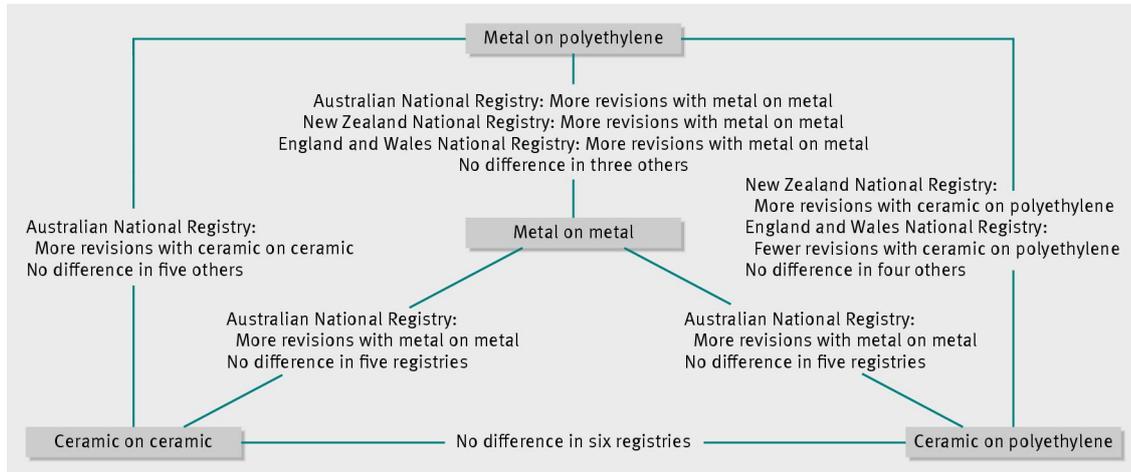
## Figures



**Fig 1** Harris hip scores at two years (top) after hip implants with metal on metal versus metal on polyethylene bearings. Zijlstra et al<sup>29</sup> measured score at both 24 months and 60 months



**Fig 2** Harris hip score beyond two years after hip implants with metal on metal versus metal on polyethylene or ceramic on polyethylene bearings. Jacobs et al<sup>40</sup> assessed score at mean 43 months while Dorr et al<sup>34</sup> assessed score at 60 months. Zijlstra et al<sup>29</sup> measured score at both 24 months and 60 months. SD of mean imputed for Jacobs et al and sensitivity analyses performed; varying estimates of SD did not change overall estimate of effect



**Fig 3** Qualitative evidence reported in six national data sources/registries (Australia, New Zealand, England and Wales, Italy (Emilia-Romagna), Kaiser and CMS databases)