Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/87579

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher’s version. Please see the ‘permanent WRAP URL’ above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
Outcome measures used in arthroplasty trials: Systematic review of the 2008 and 2013 literature

<table>
<thead>
<tr>
<th>Journal:</th>
<th>The Journal of Rheumatology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>2016-1477.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Manuscript</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Richards, Bethan; Institute of Rheumatology and Orthopaedics, Royal Prince Alfred Hospital, Sydney, Australia and Sidney Medical School, University of Sydney, Wall, Peter; Warwick Clinical trials Unit, University of Warwick, Orthopaedics Sprowson, Andrew; Warwick Clinical trials Unit, University of Warwick, Orthopaedics Singh, Jasvinder; University of Alabama Medical Center, Medicine Buchbinder, Rachelle; Cabrini Hospital and Monash University,</td>
</tr>
<tr>
<td>Keywords:</td>
<td>arthroplasty trials, systematic review, outcome measures</td>
</tr>
</tbody>
</table>
Title

Outcome measures used in arthroplasty trials: Systematic review of the 2008 and 2013 literature

Running Title: Outcome measures in arthoplasty RCTs

Authors

Bethan L Richards¹,², FRACP, MClínEpi, MSportsMed
Bethan.Richards@sswahs.nsw.gov.au

Peter DH Wall³ MBChB (Hons), MRCS (Edin), PhD
pdhwall@gmail.com

Andrew Sprowson³* MD, FRCS
A.P.Sprowson@warwick.ac.uk

Jasvinder A Singh⁴,⁵ MD, MPH
jassingh@uab.edu

Rachelle Buchbinder⁶,⁷ FRACP, MD
rachelle.buchbinder@monash.edu

*Dr Sprowson died March 13th, 2015

Affiliations

1. Institute of Rheumatology and Orthopaedics, Royal Prince Alfred Hospital, Sydney, Australia

2. Sydney Medical School, University of Sydney, Sydney, Australia
3. Birmingham Veterans Affairs Warwick Orthopaedics, University of Warwick, Coventry, UK.

4. Medical Center and University of Alabama at Birmingham, Birmingham, Alabama, USA

5. Mayo Clinic College of Medicine, Rochester, MN, USA

6. Department of Epidemiology and Preventive Medicine, School of Public Health & Preventive Medicine, Monash University, VIC Australia

7. Monash Department of Clinical Epidemiology, Cabrini Hospital, Melbourne, VIC, Australia.

**Conflict of interest statement**

Each author certifies that he or she has no commercial associations (e.g., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article

**Word Count:** 3282 (excluding Abstract)

**Corresponding Author:**

Dr Bethan L Richards
Bethan.Richards@sswhs.nsw.gov.au
Abstract

Background/objective: Previously published literature assessing the reporting of outcome measures used in joint replacement randomised controlled trials (RCTs) has revealed disappointing results. It remains unknown as to whether international initiatives have led to any improvement in the quality of reporting and/or a reduction in the heterogeneity of outcome measures used. Our objective was to systematically assess and compare primary outcome measures and risk of bias of joint replacement RCTs published in 2008 and 2013.

Methods: We searched MEDLINE, EMBASE and CENTRAL for RCTs investigating adult patients undergoing joint replacement surgery. Two authors independently identified eligible trials, extracted data and assessed risk of bias using the Cochrane tool.

Results: Seventy RCTs (30 in 2008; 40 in 2013) met the eligibility criteria. There was no significant difference in the number of trials judged to be at low overall risk of bias (N=6, 20%) in 2008 compared with six (15%) in 2013 ($\chi^2=0.302$, P=0.75). Significantly more trials published in 2008 did not specify a primary outcome measure (N=25, 83%) compared with 18 (45%) trials in 2013, $\chi^2= 10.6316$, P=0.001). When specified, there was significant heterogeneity in the measures used to assess primary outcomes.

Conclusion: While less than a quarter of trials published in both 2008 and 2013 were judged to be at low overall risk of bias, significantly more trials published in 2013 specified a primary outcome. Although this might represent a temporal trend towards improvement, the overall frequency of primary outcome reporting and the wide heterogeneity in primary outcomes reported remain suboptimal.
Introduction

With an expanding and aging population, an escalating prevalence of obesity and a rising need for both initial and joint revision surgery, the incidence and associated economic burden of joint replacement surgery has been projected to increase exponentially(1). For many patients, joint replacement surgery is an effective management option to reduce pain, restore function and improve quality of life. However, individuals who undergo joint replacement are also at risk for a variety of adverse events associated with both the anaesthetic and the surgery. With expanding indications for joint replacement and the continuing evolution of surgical techniques and implants, many important research questions need to be answered. To address these issues there is an ongoing need for high quality trials within this field of orthopaedics.

Randomised controlled trials (RCTs) are widely acknowledged to be the best type of trial design to evaluate the effectiveness and safety of health care interventions.(2-5) However, the RCT’s ability to answer important clinical questions will always be limited by its design and the outcome measures used. To draw meaningful conclusions from individual RCTs, relevant, robust and validated outcomes measures are required. In addition, these outcomes should be pre-specified and clearly reported as to whether they are primary or secondary. This enables readers to assess whether the RCT is adequately powered and avoids the perception of selective reporting bias.

In the field of joint replacement surgery, previously published literature assessing the reporting frequency, relevance and homogeneity of outcome measures used has revealed disappointing results(6). Specifically, primary outcomes were often not specified, and when they were, there
was significant heterogeneity in the types of outcomes measures used to assess the same endpoint. To address this on a large scale, several multi-national collaborations and initiatives have been established. For example, following the poor findings in their systematic review, Riddle and colleagues proposed that consensus from an international group of experts involved in the care of these patients was needed(7). In 2008 a working group within Outcome Measures in Rheumatology (OMERACT) and Osteoarthritis Research Society International (OARSI) was established with the aim of improving the reporting of relevant, evidence based health outcome domains within joint replacement trials(8). In addition, to facilitate accurate, complete, and transparent reporting of all clinical trials, in 2008 the EQUATOR (Enhancing the Quality and Transparency of Health Research) Network was launched and in 2010 the Consolidated Standards of Reporting Trials (CONSORT) Statement (first published in 1996) was published to provide researchers with a check list of 25 items to facilitate complete and transparent reporting of trial findings(9).

It remains unknown as to whether these international initiatives have led to any improvement in the quality of joint replacement trial reporting and/or a reduction in the heterogeneity of outcome measures used. To investigate this question and inform the OMERACT 2014 Working Group meeting (which aimed to define an internationally agreed upon core set of domains and outcome measures that should be reported in every joint replacement clinical trial)(10), we performed a systematic review of outcomes that had been reported in joint replacement trials published in 2008 and 2013. This paper reports the risk of bias of included trials and assesses and compares their primary outcomes. A separate paper will report the extent to which all reported outcomes
met the OMERACT criteria of truth, discrimination, and feasibility and map the reported outcomes to OMERACT Filter 2.0(11).

Methods

Search Strategy and Criteria

This systematic review was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) format(12) and the protocol was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO) Registration number: CRD42014009216. This study did not require ethics committee approval, since it does not involve any human or animals and is a systematic review of published articles in the medical literature.

All randomised or quasi-randomised (where allocation not strictly random) controlled trials investigating adult patients undergoing joint replacement surgery (defined as substitution of any joint surface with a prosthesis) were identified. Trials were included if the comparator was another type of joint implant, surgical placebo or sham, usual care, physical therapy or other active treatment and at least one outcome had been reported. Studies were excluded if they evaluated spinal joint replacement surgery, had a primary intervention of interest that was not the insertion of a joint prosthesis (e.g. trials investigating pre-operative education, peri-operative analgesia or post-operative care) or were not published as a full report in English.

An electronic literature search for articles published in 2008 and 2013 was performed in MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL)
using a comprehensive search strategy (Supplementary Table 1). We also performed a hand search of reference lists of relevant articles to identify additional relevant trials.

Two authors (BR and PW) independently screened the titles and abstracts of all studies identified by the search strategy and then independently reviewed the full text of all potentially eligible studies to identify studies that fulfilled inclusion criteria. Any disagreement in study selection was resolved by consensus or by discussion with a third reviewer (RB).

The same two authors independently extracted data from the included using predetermined forms. Differences in data extraction were resolved by referring back to the original articles and establishing a consensus. A third reviewer (RB) was consulted to help resolve differences as necessary. The information extracted included study site, funding, enrolment date, size, design, population, interventions, and outcome measures (and whether they were pre-specified). Each outcome was recorded as either primary or secondary. An outcome was recorded as primary if it was reported as a “primary outcome” in the manuscript or registered protocol or was used to calculate the sample size. More than one primary outcome could be recorded provided these criteria were met.

Two authors (BR and PW) independently assessed risk of bias for all included studies using methods recommended by the Cochrane Collaboration,(13) which assess the following key criteria: random sequence generation, allocation concealment, blinding of participants, care provider and outcome assessor for each outcome measure, incomplete outcome data, selective outcome reporting and other sources of bias. Other sources of bias that were considered included
whether or not co-interventions and adherence to treatment (e.g. for analgesics and physical therapy programs) were assessed and reported to be equal between groups, and sources of funding. Each criterion was rated as low risk of bias, high risk of bias, or unclear risk (indicating either lack of information or uncertainty over the potential for bias). Information to inform the risk of bias rating was sourced from searching trial registries and the published papers.

An assessment of overall risk of bias was made where low overall risk of bias indicated plausible bias unlikely to seriously alter the results (low risk of bias for all key domains), unclear overall risk of bias indicated plausible bias that raises some doubt about the results (unclear risk of bias for one or more key domains) and high overall risk of bias indicating plausible bias seriously weakens confidence in the results (high risk of bias for one or more key domains)(13). A third reviewer (RB) was consulted to resolve differences as necessary.

**Data Analysis**

A descriptive analysis of the primary outcomes was performed which compared frequency and description of primary outcome measures by site of joint replacement and year. Comparison of 2008 and 2013 results for overall risk of bias and number of trials reporting primary outcomes was made using Pearson chi-squared tests. We also used the Pearson chi-squared test to determine whether an association might exist between overall risk of bias and reporting of primary outcomes.

**Results**

The search strategy identified 1635 studies. Of these 70 randomised controlled trials (30 with 2,789 participants published in 2008 and 40 with 4253 participants published in 2013), met the
eligibility criteria and were included in the review. A PRISMA flow diagram of the searches through to the final inclusion is shown in Figure 1. The summary characteristics of the included trials according to their year of publication are shown in Table 2.

**Study characteristics of trials published in 2008**

Eighteen (60%) of the 30 trials published in 2008 were conducted in Europe (14-26), six (20%) in the USA(27-32), two (7%) each in Asia(33, 34) and Canada (35, 36) and one (3%) each in Australia(37) and New Zealand(38). Twenty (67%) trials were independently funded(14-17, 19, 20, 22-24, 26-28, 30-32, 34, 35, 37-39) , 9 (30%) had industry funding(17, 18, 21, 29, 31, 36, 40-42) and in one trial the source of funding was unclear(25). The year of first recruitment ranged from 1994 to 2007 (median, 2001). The most common joint evaluated was the knee  (N= 19, 63%)(17-24, 26, 29, 30, 33-37, 39, 42, 43), followed by the hip  (N=10, 33%)(14-16, 27, 28, 31, 32, 38, 40, 41), and one trial (3%) evaluated the shoulder(25). There were no trials evaluating joint replacement surgery of the elbow, wrist, hand, ankle or foot. Fifteen (50%) trials evaluated two or more different prostheses or components(14, 16-19, 25, 26, 30, 31, 38-43), 13 (43%) evaluated the same prosthesis but used differing surgical techniques(20-24, 27-29, 33-37) and two (1%) evaluated joint replacement versus other joint surgeries(15, 32).

**Study Characteristics of trials published in 2013**

Twenty (50%) of the 40 trials published in 2013 were conducted in Europe(44-63), nine (23%) in Asia(64-71), five (13%) in the USA(72-76), four (10%) in Canada(77-80), and one (3%) each in Australia(81) and the Middle East(82). Twelve (30%) trials were industry funded(45, 52, 54-56, 59, 62, 63, 70, 73, 77, 80), 15 (38%) had independent funding(44, 46, 48, 50, 51, 53, 57, 60,
64, 65, 69, 72, 78, 81, 82) and in 13 (33%) trials the source of funding was not specified(47, 49, 58, 61, 66-68, 70, 71, 74-76, 79). The year of first recruitment ranged from 1996 to 2011 (median, 2007). The knee (n= 20, 50%)(45, 48, 49, 54, 55, 57, 61, 64-66, 68-71, 73-76, 81) and hip (N=17, 43%)(44, 46, 47, 51-53, 56, 58-60, 62, 63, 67, 72, 77, 79, 80) were again the most commonly studied joints, with two trials (5%) studying the shoulder(78, 82) and one (3%) studying the wrist(50). There were no trials evaluating joint replacement surgery of the elbow, hand, ankle or foot. Eighteen (43%) trials evaluated two or more different prostheses or components(44, 45, 47-50, 52, 54, 56, 57, 61-64, 67, 73, 79, 80), 20 (50%) evaluated the same prosthesis but used differing surgical techniques(53, 55, 58-60, 65, 66, 68-72, 74-78, 81, 82) and two (5%) evaluated joint replacement versus other joint surgeries(46, 51).

**Risk of Bias**

Of the 30 trials published in 2008, 6 (20%) were judged to be at low overall risk of bias, and the remainder were all judged to be at high or unclear overall risk of bias (Figure 2). Four (21%) of the 19 knee trials(36, 39, 42, 43), and 2 (20%) of the 10 hip trials were judged to be low risk of bias(40, 41). In the one shoulder joint trial(25), the risk of bias was deemed unclear.

Of the 40 trials published in 2013, only six (15%) were judged to be at low overall risk of bias (Figure 3). Three (15%) of the 20 knee trials were judged to be at low overall risk of bias(56, 58, 77), 11 (55%) were deemed unclear(48, 57, 64, 66, 68-70, 73, 76, 81) and 6 (30%) were judged to be at high risk of bias(49, 54, 61, 71, 74, 75). Three (21%) of the 17 hip trials were judged to be at low overall risk of bias(56, 58, 77), six (43%) were deemed unclear(46, 47, 59, 60, 62, 80) and eight (57%) were judged to be at high risk of bias(44, 51-53, 63, 67, 72, 79). One shoulder
trial was deemed unclear(82) and one was judged to be at high overall risk of bias(78), while the single wrist trial was also judged to be at high overall risk of bias(50).

Figure 4 compares the number of trials with low risk of bias for each risk of bias domain and overall low risk of bias according to publication year. There was no difference between years in number of trials judged to be at overall low risk of bias (6/30 in 2008 compared with 6/40 in 2013, \( \chi^2 = 0.302, P=0.75 \)). The method used to generate the random sequence was adequately reported in 63% of the 2008 trials and 53% of the 2013 trials, however details of allocation concealment were reported in only 10 (33%) trials in 2008 and 12 (30%) trials in 2013. More trials reported blinding of patients (13, 33%) and outcomes (N=14, 35%) in 2013 in comparison with 2008 (N=17 (23%) and 4 (13%) respectively). Twenty-six (87%) trials in 2008 reported detailed baseline characteristics however this only applied to 22 (55%) of the 2013 trials. Few trials pre-specified or reported the use of relevant co-interventions (N=13, 43%) in 2008; N=9, 22% in 2013), or described how incomplete data were addressed (N=4, 13% in 2008; N=14, 35% in 2013). Several trials also had evidence of selective outcome reporting (N=11, 37% in 2008; N=9, 22% in 2013). These issues may have influenced outcomes.

**Primary Outcomes**

A summary of the primary outcomes reported in the 2008 and 2013 trials is shown in Table 3. Compared with trials published in 2008, more trials published in 2013 reported a primary outcome (N=22/40 (55%) compared with N=5/30 (17%), \( \chi^2 = 10.6316, P=0.001 \)). Four out of 19 (21%) knee trials specified a primary outcome in 2008 compared with 11/20 (55%) knee trials in 2013. In these 15 knee trials, 17 different primary outcomes were specified despite evaluating
similar questions. Within each time period only two trials reported the same primary outcome (knee range of motion was specified in two trials in 2008 and difference in mechanical axis deviation measured in degrees in the coronal plane on x-ray in two trials in 2013). No primary outcome was used in both time periods. The majority of primary outcomes reported evaluated technical aspects of the procedures (N=12/17, 71%), rather than patient centred outcomes.

Only one out of 10 (10%) hip trials reported a primary outcome in 2008 compared with 8/17 (47%) hip trials in 2013. Similar to the knee trials, the primary outcomes varied widely and were focussed on technical outcomes of the procedure. Primary outcomes used in more than one hip trial included component migration (RSA) (N=3), deviation ≥5 degrees of planned stem shaft angle (x-ray) (N=2), computerized gait assessment (mean gait velocity, stride length) (N=2) and revision rates (N=2). Assessment of component migration was the only primary outcome measure utilised in both time periods.

Two of the three shoulder trials reported a primary outcome. Only the two 2013 trials evaluating shoulder joint replacement surgery reported a primary outcome, and each trial used a different measures to assess “improvement” (post-operative pain on a VAS 0-10mm and healing rate of the subscapularis tendon visualised on MRI). The single wrist trial reported a primary outcome and used RSA to measure component migration in mm.

**Discussion**

We observed a significant difference in the frequency of reporting of primary outcomes in joint replacement trials in 2013 compared with 2008. Only 17% (21% knee and 10% hip) of RCTs
published in 2008 reported a primary outcome measure in comparison with 55% (55% knee and 47% hip) in 2013. Without knowledge of the frequency of reporting before 2008 and between 2008 and 2013 it is not possible to know with certainty whether or not the improved reporting in 2013 reflects a real improvement over time. Nevertheless almost half of all joint replacement trials continue to fail to specify a primary outcome despite widely accepted CONSORT recommendations(9, 83). Similar inadequate reporting of primary outcomes have been shown in other surgical fields including ophthalmic surgery(84), solid organ transplantation(85), plastic surgery(86), urology(87), trauma surgery(88) and neurosurgery(89).

In addition, we found that among trials that did specify one or more primary outcomes, these varied widely despite the trials addressing similar research questions. In both years (2008 and 2013), no primary outcome measure was used in more than two trials despite similar research questions. This heterogeneity in primary outcome reporting is consistent with results from a previous systematic review(6). Heterogeneity in outcome measurement hampers our ability to combine, contrast and accurately interpret the results from multiple RCTs answering the same (and sometimes, similar) research questions. To improve the quality of information available for patients undergoing joint replacement surgery, RCTs evaluating the same clinical questions need to utilise a homogenous set of outcome measures. Further efforts are required to achieve this(10).

Furthermore, the majority of primary outcomes reported were predominantly focussed upon technical aspects of the surgery. Hence, despite the significant investment of time, money and resources in evaluating these important research questions, we found that the majority of trials in
this systematic review were not designed nor powered to evaluate other important core domains of health for both the patient and society.

Few trials published in both 2008 and 2013 were judged to be at low overall risk of bias (20% in 2008 and 15% in 2013). Not surprisingly, the trials at less potential for bias were more likely to report a primary outcome measure. In addition almost a quarter of the trials we included trials (20/70, 23%) were judged to be at unclear risk of bias as they reported insufficient information. While it is often not possible in surgical RCTs to blind the investigators to the group assignments, or standardize surgical techniques, it should be possible to minimise other potential sources of bias. Common areas of potential bias occurring in more than 50% of the studies included failing to describe allocation concealment, participant blinding, how incomplete data were addressed, and selective reporting of outcomes. Lack of adequate reporting of details of randomisation, allocation concealment, blinding, co-intervention use and outcomes is not limited to joint replacement trials or orthopaedic surgery(84, 87-93).

Our study has several limitations. Firstly, while we used a comprehensive systematic search strategy to identify all relevant studies, we excluded foreign language publications. Given the high proportion of papers published in English language journals (80–90%) this is unlikely to affect generalisability(94). Secondly, the majority of joint replacement trials in this review involved hip and knee surgery. There were limited trials evaluating the shoulder, wrist and hand, and no trials evaluating elbow or ankle joint replacement. Therefore our results may or may not be generalisable to joint replacement trials of other joints. Thirdly, in selecting two publication years, there is a possibility that this literature may not have been truly representative of periods
just before, between and after these dates. The five year gap between study years may not have been long enough to capture meaningful change; however we chose the start year as the year of publication of the study that showed poor quality of arthroplasty trials (6). More studies may be needed in the future with a longer interval to look for improvements using the same quality criteria. Our results are however consistent with previous reviews. Finally, we judged risk of bias and specification of primary outcomes on the basis of the published paper. It may be that we overestimated potential for bias and underestimated frequency of primary outcome specification due to poor reporting practices rather than suboptimal trial methodology. However we tried to limit this effect by also searching the trial registries for protocols.

In conclusion, despite an observed increase in frequency of reporting of primary outcome measures in joint replacement trials in 2013 compared with 2008, almost 50% of trials published in 2013 did not report their primary outcomes. In addition among trials that did report primary outcomes, these were heterogeneous, frequently measured technical aspects of surgery rather than patient important endpoints and few trials used the same primary outcome even for similar research questions. In addition the majority of trials published in both years were at high or unclear overall risk of bias and reflect a lack of implementation of quality improvement initiatives such as the CONSORT guidelines (or similar). Further efforts are needed to improve the quality of joint replacement trials and ensure primary outcomes are reported. A standardised, universally accepted core set of outcomes to be used in all joint replacement trials, based upon their clinical relevance would enhance this field.

Acknowledgement
Co-author Dr. Andrew P. Sprowson died tragically on March 13, 2015. The authors remember him as an orthopaedic surgeon with immense enthusiasm for research and for robust clinical evidence in the field of joint replacement surgery. He was an integral part of the Working group for joint arthroplasty within OMERACT, and a great friend.
References


Title

Outcome measures used in arthroplasty trials: Systematic review of the 2008 and 2013 literature

Running Title: Outcome measures in arthroplasty RCTs

Authors

Bethan L Richards1,2, FRACP, M ClinEpi, M SportsMed
Bethan.Richards@sswahs.nsw.gov.au

Peter DH Wall3 MB ChB (Hons), MRCS (Edin), PhD
pdhwall@gmail.com

Andrew Sprowson3* MD, FRCS
A.P.Sprowson@warwick.ac.uk

Jasvinder A Singh4,5 MD, MPH
jassingh@uab.edu

Rachelle Buchbinder6,7 FRACP, MD
rachelle.buchbinder@monash.edu

*Dr Sprowson died March 13th, 2015

Affiliations

1. Institute of Rheumatology and Orthopaedics, Royal Prince Alfred Hospital, Sydney, Australia

2. Sydney Medical School, University of Sydney, Sydney, Australia
3. Birmingham Veterans Affairs Warwick Orthopaedics, University of Warwick, Coventry, UK.

4. Medical Center and University of Alabama at Birmingham, Birmingham, Alabama, USA

5. Mayo Clinic College of Medicine, Rochester, MN, USA

6. Department of Epidemiology and Preventive Medicine, School of Public Health & Preventive Medicine, Monash University, VIC Australia

7. Monash Department of Clinical Epidemiology, Cabrini Hospital, Melbourne, VIC, Australia.

Conflicts of interest statement

JAS has received research grants from Takeda and Savient and consultant fees from Savient, Takeda, Regeneron, Merz, Iroko, Bioiberica, Crealta/Horizon and Allergan pharmaceuticals, WebMD, UBM LLC and the American College of Rheumatology. JAS serves as the principal investigator for an investigator-initiated study funded by Horizon pharmaceuticals through a grant to DINORA, Inc., a 501 (c)(3) entity. JAS is a member of the executive of OMERACT, an organization that develops outcome measures in rheumatology and receives arms-length funding from 36 companies; a member of the American College of Rheumatology's (ACR) Annual Meeting Planning Committee (AMPC); Chair of the ACR Meet-the-Professor, Workshop and Study Group Subcommittee; and a member of the Veterans Affairs Rheumatology Field Advisory Committee. JAS is the editor and Director of the UAB Cochrane Musculoskeletal Group Satellite Center on Network Meta-analysis.

All other authors certify that he or she has no commercial associations (e.g., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article.
**Word Count:** 3282 (excluding Abstract)

**Corresponding Author:**

Dr Bethan L Richards
Bethan.Richards@sswahs.nsw.gov.au
Abstract

**Background/objective:** Previously published literature assessing the reporting of outcome measures used in joint replacement randomised controlled trials (RCTs) has revealed disappointing results. It remains unknown as to whether international initiatives have led to any improvement in the quality of reporting and/or a reduction in the heterogeneity of outcome measures used. Our objective was to systematically assess and compare primary outcome measures and risk of bias of joint replacement RCTs published in 2008 and 2013.

**Methods:** We searched MEDLINE, EMBASE and CENTRAL for RCTs investigating adult patients undergoing joint replacement surgery. Two authors independently identified eligible trials, extracted data and assessed risk of bias using the Cochrane tool.

**Results:** Seventy RCTs (30 in 2008; 40 in 2013) met the eligibility criteria. There was no significant difference in the number of trials judged to be at low overall risk of bias (N=6, 20%) in 2008 compared with six (15%) in 2013 ($\chi^2=0.302$, $P=0.75$). Significantly more trials published in 2008 did not specify a primary outcome measure (N=25, 83%) compared with 18 (45%) trials in 2013, $\chi^2=10.6316$, $P=0.001$). When specified, there was significant heterogeneity in the measures used to assess primary outcomes.

**Conclusion:** While less than a quarter of trials published in both 2008 and 2013 were judged to be at low overall risk of bias, significantly more trials published in 2013 specified a primary outcome. Although this might represent a temporal trend towards improvement, the overall frequency of primary outcome reporting and the wide heterogeneity in primary outcomes reported remain suboptimal.
Introduction

With an expanding and aging population, an escalating prevalence of obesity and a rising need for both initial and joint revision surgery, the incidence and associated economic burden of joint replacement surgery has been projected to increase exponentially(1). For many patients, joint replacement surgery is an effective management option to reduce pain, restore function and improve quality of life. However, individuals who undergo joint replacement are also at risk for a variety of adverse events associated with both the anaesthetic and the surgery. With expanding indications for joint replacement and the continuing evolution of surgical techniques and implants, many important research questions need to be answered. To address these issues there is an ongoing need for high quality trials within this field of orthopaedics.

Randomised controlled trials (RCTs) are widely acknowledged to be the best type of trial design to evaluate the effectiveness and safety of health care interventions.(2-5) However, the RCT’s ability to answer important clinical questions will always be limited by its design and the outcome measures used. To draw meaningful conclusions from individual RCTs, relevant, robust and validated outcomes measures are required. In addition, these outcomes should be pre-specified and clearly reported as to whether they are primary or secondary. This enables readers to assess whether the RCT is adequately powered and avoids the perception of selective reporting bias.

In the field of joint replacement surgery, previously published literature assessing the reporting frequency, relevance and homogeneity of outcome measures used has revealed disappointing results(6). Specifically, primary outcomes were often not specified, and when they were, there
was significant heterogeneity in the types of outcomes measures used to assess the same endpoint. To address this on a large scale, several multi-national collaborations and initiatives have been established. For example, following the poor findings in their systematic review, Riddle and colleagues proposed that consensus from an international group of experts involved in the care of these patients was needed(7). In 2008 a working group within Outcome Measures in Rheumatology (OMERACT) and Osteoarthritis Research Society International (OARSI) was established with the aim of improving the reporting of relevant, evidence based health outcome domains within joint replacement trials(8). In addition, to facilitate accurate, complete, and transparent reporting of all clinical trials, in 2008 the EQUATOR (Enhancing the Quality and Transparency of Health Research) Network was launched and in 2010 the Consolidated Standards of Reporting Trials (CONSORT) Statement (first published in 1996) was published to provide researchers with a check list of 25 items to facilitate complete and transparent reporting of trial findings(9).

It remains unknown as to whether these international initiatives have led to any improvement in the quality of joint replacement trial reporting and/or a reduction in the heterogeneity of outcome measures used. To investigate this question and inform the OMERACT 2014 Working Group meeting (which aimed to define an internationally agreed upon core set of domains and outcome measures that should be reported in every joint replacement clinical trial)(10), we performed a systematic review of outcomes that had been reported in joint replacement trials published in 2008 and 2013. This paper reports the risk of bias of included trials and assesses and compares their primary outcomes. A separate paper will report the extent to which all reported outcomes
met the OMERACT criteria of truth, discrimination, and feasibility and map the reported outcomes to OMERACT Filter 2.0(11).

Methods

Search Strategy and Criteria

This systematic review was performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) format(12) and the protocol was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO) Registration number: CRD42014009216. This study did not require ethics committee approval, since it does not involve any human or animals and is a systematic review of published articles in the medical literature.

All randomised or quasi-randomised (where allocation not strictly random) controlled trials investigating adult patients undergoing joint replacement surgery (defined as substitution of any joint surface with a prosthesis) were identified. Trials were included if the comparator was another type of joint implant, surgical placebo or sham, usual care, physical therapy or other active treatment and at least one outcome had been reported. Studies were excluded if they evaluated spinal joint replacement surgery, had a primary intervention of interest that was not the insertion of a joint prosthesis (e.g. trials investigating pre-operative education, peri-operative analgesia or post-operative care) or were not published as a full report in English.

An electronic literature search for articles published in 2008 and 2013 was performed in MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL)
using a comprehensive search strategy (Table 1). We also performed a hand search of reference lists of relevant articles to identify additional relevant trials.

Two authors (BR and PW) independently screened the titles and abstracts of all studies identified by the search strategy and then independently reviewed the full text of all potentially eligible studies to identify studies that fulfilled inclusion criteria. Any disagreement in study selection was resolved by consensus or by discussion with a third reviewer (RB).

The same two authors independently extracted data from the included using predetermined forms. Differences in data extraction were resolved by referring back to the original articles and establishing a consensus. A third reviewer (RB) was consulted to help resolve differences as necessary. The information extracted included study site, funding, enrolment date, size, design, population, interventions, and outcome measures (and whether they were pre-specified). Each outcome was recorded as either primary or secondary. An outcome was recorded as primary if it was reported as a “primary outcome” in the manuscript or registered protocol or was used to calculate the sample size. More than one primary outcome could be recorded provided these criteria were met.

Two authors (BR and PW) independently assessed risk of bias for all included studies using methods recommended by the Cochrane Collaboration,(13) which assess the following key criteria: random sequence generation, allocation concealment, blinding of participants, care provider and outcome assessor for each outcome measure, incomplete outcome data, selective outcome reporting and other sources of bias. Other sources of bias that were considered included
whether or not co-interventions and adherence to treatment (e.g. for analgesics and physical therapy programs) were assessed and reported to be equal between groups, and sources of funding. Each criterion was rated as low risk of bias, high risk of bias, or unclear risk (indicating either lack of information or uncertainty over the potential for bias). Information to inform the risk of bias rating was sourced from searching trial registries and the published papers.

An assessment of overall risk of bias was made where low overall risk of bias indicated plausible bias unlikely to seriously alter the results (low risk of bias for all key domains), unclear overall risk of bias indicated plausible bias that raises some doubt about the results (unclear risk of bias for one or more key domains) and high overall risk of bias indicating plausible bias seriously weakens confidence in the results (high risk of bias for one or more key domains)(13). A third reviewer (RB) was consulted to resolve differences as necessary.

Data Analysis
A descriptive analysis of the primary outcomes was performed which compared frequency and description of primary outcome measures by site of joint replacement and year. Comparison of 2008 and 2013 results for overall risk of bias and number of trials reporting primary outcomes was made using Pearson chi-squared tests. We also used the Pearson chi-squared test to determine whether an association might exist between overall risk of bias and reporting of primary outcomes.

Results
The search strategy identified 1635 studies. Of these 70 randomised controlled trials (30 with 2,789 participants published in 2008 and 40 with 4253 participants published in 2013), met the
eligibility criteria and were included in the review. A PRISMA flow diagram of the searches through to the final inclusion is shown in Figure 1. The summary characteristics of the included trials according to their year of publication are shown in Table 2.

**Study characteristics of trials published in 2008**

Eighteen (60%) of the 30 trials published in 2008 were conducted in Europe (14-26), six (20%) in the USA(27-32), two (7%) each in Asia(33, 34) and Canada (35, 36) and one (3%) each in Australia(37) and New Zealand(38). Twenty (67%) trials were independently funded(14-17, 19, 20, 22-24, 26-28, 30-32, 34, 35, 37-39), 9 (30%) had industry funding(17, 18, 21, 29, 31, 36, 40-42) and in one trial the source of funding was unclear(25). The year of first recruitment ranged from 1994 to 2007 (median, 2001). The most common joint evaluated was the knee (N= 19, 63%(17-24, 26, 29, 30, 33-37, 39, 42, 43), followed by the hip (N=10, 33%)(14-16, 27, 28, 31, 32, 38, 40, 41), and one trial (3%) evaluated the shoulder(25). There were no trials evaluating joint replacement surgery of the elbow, wrist, hand, ankle or foot. Fifteen (50%) trials evaluated two or more different prostheses or components(14, 16-19, 25, 26, 30, 31, 38-43), 13 (43%) evaluated the same prosthesis but used differing surgical techniques(20-24, 27-29, 33-37) and two (1%) evaluated joint replacement versus other joint surgeries(15, 32).

**Study Characteristics of trials published in 2013**

Twenty (50%) of the 40 trials published in 2013 were conducted in Europe(44-63), nine (23%) in Asia(64-71), five (13%) in the USA(72-76), four (10%) in Canada(77-80), and one (3%) each in Australia(81) and the Middle East(82). Twelve (30%) trials were industry funded(45, 52, 54-56, 59, 62, 63, 70, 73, 77, 80), 15 (38%) had independent funding(44, 46, 48, 50, 51, 53, 57, 60,
64, 65, 69, 72, 78, 81, 82 and in 13 (33%) trials the source of funding was not specified(47, 49, 58, 61, 66-68, 70, 71, 74-76, 79). The year of first recruitment ranged from 1996 to 2011 (median, 2007). The knee (n= 20, 50%)(45, 48, 49, 54, 55, 57, 61, 64-66, 68-71, 73-76, 81) and hip (N=17, 43%)(44, 46, 47, 51-53, 56, 58-60, 62, 63, 67, 72, 77, 79, 80) were again the most commonly studied joints, with two trials (5%) studying the shoulder(78, 82) and one (3%) studying the wrist(50). There were no trials evaluating joint replacement surgery of the elbow, hand, ankle or foot. Eighteen (43%) trials evaluated two or more different prostheses or components(44, 45, 47-50, 52, 54, 56, 57, 61-64, 67, 73, 79, 80), 20 (50%) evaluated the same prosthesis but used differing surgical techniques(53, 55, 58-60, 65, 66, 68-72, 74-78, 81, 82) and two (5%) evaluated joint replacement versus other joint surgeries(46, 51).

**Risk of Bias**

Of the 30 trials published in 2008, 6 (20%) were judged to be at low overall risk of bias, and the remainder were all judged to be at high or unclear overall risk of bias (Figure 2). Four (21%) of the 19 knee trials(36, 39, 42, 43), and 2 (20%) of the 10 hip trials were judged to be low risk of bias(40, 41). In the one shoulder joint trial(25), the risk of bias was deemed unclear.

Of the 40 trials published in 2013, only six (15%) were judged to be at low overall risk of bias (Figure 3). Three (15%) of the 20 knee trials were judged to be at low overall risk of bias(56, 58, 77), 11 (55%) were deemed unclear(48, 57, 64, 66, 68-70, 73, 76, 81) and 6 (30%) were judged to be at high risk of bias(49, 54, 61, 71, 74, 75). Three (21%) of the 17 hip trials were judged to be at low overall risk of bias(56, 58, 77), six (43%) were deemed unclear(46, 47, 59, 60, 62, 80) and eight (57%) were judged to be at high risk of bias(44, 51-53, 63, 67, 72, 79). One shoulder
trial was deemed unclear (82) and one was judged to be at high overall risk of bias (78), while the single wrist trial was also judged to be at high overall risk of bias (50).

Figure 4 compares the number of trials with low risk of bias for each risk of bias domain and overall low risk of bias according to publication year. There was no difference between years in number of trials judged to be at overall low risk of bias (6/30 in 2008 compared with 6/40 in 2013, $\chi^2 = 0.302$, $P=0.75$). The method used to generate the random sequence was adequately reported in 63% of the 2008 trials and 53% of the 2013 trials, however details of allocation concealment were reported in only 10 (33%) trials in 2008 and 12 (30%) trials in 2013. More trials reported blinding of patients (13, 33%) and outcomes (N=14, 35%) in 2013 in comparison with 2008 (N=17 (23%) and 4 (13%) respectively). Twenty-six (87%) trials in 2008 reported detailed baseline characteristics however this only applied to 22 (55%) of the 2013 trials. Few trials pre-specified or reported the use of relevant co-interventions (N=13, 43%) in 2008; N=9, 22% in 2013), or described how incomplete data were addressed (N=4, 13% in 2008; N=14, 35% in 2013). Several trials also had evidence of selective outcome reporting (N=11, 37% in 2008; N=9, 22% in 2013). These issues may have influenced outcomes.

Primary Outcomes

A summary of the primary outcomes reported in the 2008 and 2013 trials is shown in Table 3. Compared with trials published in 2008, more trials published in 2013 reported a primary outcome (N=22/40 (55%) compared with N=5/30 (17%), $\chi^2 = 10.6316$, $P=0.001$). Four out of 19 (21%) knee trials specified a primary outcome in 2008 compared with 11/20 (55%) knee trials in 2013. In these 15 knee trials, 17 different primary outcomes were specified despite evaluating
similar questions. Within each time period only two trials reported the same primary outcome (knee range of motion was specified in two trials in 2008 and difference in mechanical axis deviation measured in degrees in the coronal plane on x-ray in two trials in 2013). No primary outcome was used in both time periods. The majority of primary outcomes reported evaluated technical aspects of the procedures (N=12/17, 71%), rather than patient centred outcomes.

Only one out of 10 (10%) hip trials reported a primary outcome in 2008 compared with 8/17 (47%) hip trials in 2013. Similar to the knee trials, the primary outcomes varied widely and were focussed on technical outcomes of the procedure. Primary outcomes used in more than one hip trial included component migration (RSA) (N=3), deviation ≥5 degrees of planned stem shaft angle (x-ray) (N=2), computerized gait assessment (mean gait velocity, stride length) (N=2) and revision rates (N=2). Assessment of component migration was the only primary outcome measure utilised in both time periods.

Two of the three shoulder trials reported a primary outcome. Only the two 2013 trials evaluating shoulder joint replacement surgery reported a primary outcome, and each trial used a different measures to assess “improvement” (post-operative pain on a VAS 0-10mm and healing rate of the subscapularis tendon visualised on MRI). The single wrist trial reported a primary outcome and used RSA to measure component migration in mm.

**Discussion**

We observed a significant difference in the frequency of reporting of primary outcomes in joint replacement trials in 2013 compared with 2008. Only 17% (21% knee and 10% hip) of RCTs
published in 2008 reported a primary outcome measure in comparison with 55% (55% knee and 47% hip) in 2013. Without knowledge of the frequency of reporting before 2008 and between 2008 and 2013 it is not possible to know with certainty whether or not the improved reporting in 2013 reflects a real improvement over time. Nevertheless almost half of all joint replacement trials continue to fail to specify a primary outcome despite widely accepted CONSORT recommendations(9, 83). Similar inadequate reporting of primary outcomes have been shown in other surgical fields including ophthalmic surgery(84), solid organ transplantation(85), plastic surgery(86), urology(87), trauma surgery(88) and neurosurgery(89).

In addition, we found that among trials that did specify one or more primary outcomes, these varied widely despite the trials addressing similar research questions. In both years (2008 and 2013), no primary outcome measure was used in more than two trials despite similar research questions. This heterogeneity in primary outcome reporting is consistent with results from a previous systematic review(6). Heterogeneity in outcome measurement hampers our ability to combine, contrast and accurately interpret the results from multiple RCTs answering the same (and sometimes, similar) research questions. To improve the quality of information available for patients undergoing joint replacement surgery, RCTs evaluating the same clinical questions need to utilise a homogenous set of outcome measures. Further efforts are required to achieve this(10).

Furthermore, the majority of primary outcomes reported were predominantly focussed upon technical aspects of the surgery. Hence, despite the significant investment of time, money and resources in evaluating these important research questions, we found that the majority of trials in
this systematic review were not designed nor powered to evaluate other important core domains of health for both the patient and society.

Few trials published in both 2008 and 2013 were judged to be at low overall risk of bias (20% in 2008 and 15% in 2013). Not surprisingly, the trials at less potential for bias were more likely to report a primary outcome measure. In addition almost a quarter of the trials we included trials (20/70, 23%) were judged to be at unclear risk of bias as they reported insufficient information. While it is often not possible in surgical RCTs to blind the investigators to the group assignments, or standardize surgical techniques, it should be possible to minimise other potential sources of bias. Common areas of potential bias occurring in more than 50% of the studies included failing to describe allocation concealment, participant blinding, how incomplete data were addressed, and selective reporting of outcomes. Lack of adequate reporting of details of randomisation, allocation concealment, blinding, co-intervention use and outcomes is not limited to joint replacement trials or orthopaedic surgery(84, 87-93).

Our study has several limitations. Firstly, while we used a comprehensive systematic search strategy to identify all relevant studies, we excluded foreign language publications. Given the high proportion of papers published in English language journals (80–90%) this is unlikely to affect generalisability(94). Secondly, the majority of joint replacement trials in this review involved hip and knee surgery. There were limited trials evaluating the shoulder, wrist and hand, and no trials evaluating elbow or ankle joint replacement. Therefore our results may or may not be generalisable to joint replacement trials of other joints. Thirdly, in selecting two publication years, there is a possibility that this literature may not have been truly representative of periods
just before, between and after these dates. The five year gap between study years may not have been long enough to capture meaningful change; however we chose the start year as the year of publication of the study that showed poor quality of arthroplasty trials (6). More studies may be needed in the future with a longer interval to look for improvements using the same quality criteria. Our results are however consistent with previous reviews. Finally, we judged risk of bias and specification of primary outcomes on the basis of the published paper. It may be that we overestimated potential for bias and underestimated frequency of primary outcome specification due to poor reporting practices rather than suboptimal trial methodology. However we tried to limit this effect by also searching the trial registries for protocols.

In conclusion, despite an observed increase in frequency of reporting of primary outcome measures in joint replacement trials in 2013 compared with 2008, almost 50% of trials published in 2013 did not report their primary outcomes. In addition among trials that did report primary outcomes, these were heterogeneous, frequently measured technical aspects of surgery rather than patient important endpoints and few trials used the same primary outcome even for similar research questions. In addition the majority of trials published in both years were at high or unclear overall risk of bias and reflect a lack of implementation of quality improvement initiatives such as the CONSORT guidelines (or similar). Further efforts are needed to improve the quality of joint replacement trials and ensure primary outcomes are reported. A standardised, universally accepted core set of outcomes to be used in all joint replacement trials, based upon their clinical relevance would enhance this field.

**Acknowledgement**
Co-author Dr. Andrew P. Sprowson died tragically on March 13, 2015. The authors remember him as an orthopaedic surgeon with immense enthusiasm for research and for robust clinical evidence in the field of joint replacement surgery. He was an integral part of the Working group for joint arthroplasty within OMERACT, and a great friend.


Table 1: Search Strategy for Medline and EMBASE

<table>
<thead>
<tr>
<th>Medline</th>
<th>Embase</th>
</tr>
</thead>
</table>
| 1. exp Arthroplasty/                                                   | #1 arthroplasty/exp OR 'arthroplasty'
| 2. arthroplast$.tw.                                                    | #2 total AND knee AND replacement                                      |
| 3. exp Joint Prosthesis/                                               | #3 joint AND replacement                                               |
| 4. (knee and (replace$ OR arthroplast$ OR prosthe$ OR endoprosthe$ OR implant)).tw. | #4 joint AND prosthesis                                               |
| 5. (hip and (replace$ OR arthroplast$ OR prosthe$ OR endoprosthe$ OR implant)).tw. | #5 replacement* OR arthroplast* OR prosthe* OR implant OR endoprosthe* AND knee |
| 6. (elbow and (replace$ OR arthroplast$ OR prosthe$ OR endoprosthe$ OR implant)).tw. | #6 replacement* OR arthroplast* OR prosthe* OR implant OR endoprosthe* AND hip |
| 7. (shoulder and (replace$ OR arthroplast$ OR prosthe$ OR endoprosthe$ OR implant)).tw. | #7 replacement* OR arthroplast* OR prosthe* OR implant OR endoprosthe* AND elbow |
| 8. (joint and (replace$ OR arthroplast$ OR prosthe$ OR endoprosthe$ OR implant)).tw. | #8 replacement* OR arthroplast* OR prosthe* OR implant OR endoprosthe* AND shoulder |
| 9. (tka or tkr or total knee or total hip).tw.                         | #9 tka                                                                 |
| 10. or/1-9                                                             | #10 tha                                                                 |
| 11. randomized controlled trial.pt.                                    | #11 total AND hip AND replacement                                       |
| 12. controlled clinical trial.pt.                                      | #12 knee AND prosthesis                                               |
| 13. randomized.ab.                                                     | #13 hip AND prosthesis                                                |
| 14. placebo.ab.                                                        | #14 #1 OR #2 OR #3 OR #4 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 |
| 15. drug therapy.fs.                                                   | #15 randomized AND controlled AND trial                                |
| 16. randomly.ab.                                                       | #16 randomization                                                      |
| 17. trial.ab.                                                          | #17 double AND blind AND procedure                                     |
| 18. groups.ab.                                                         | #18 single AND blind AND procedure                                     |
| 19. or/11-18                                                           | #19 #15 OR #16 OR #17 OR #18                                           |
| 20. exp animals/ not humans.sh.                                       | #20 animal NOT human                                                   |
| 21. 19 not 20                                                          | #21 #19 NOT #20                                                        |
| 22. 10 and 21                                                          | #22 #14 AND #19 AND #21                                                |
| 23. limit 22 to "all adult (19 plus years)"                            | #23 #22 AND [(adult)/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim) AND 2008:py |
| 24. limit 23 to english                                                | #24 #22 AND [(adult)/lim OR [aged]/lim OR [middle aged]/lim OR [very elderly]/lim) AND 2013:py |
| 25. limit 24 to yr="2008"                                              | #25 english:la                                                         |
Table 2: Characteristics of joint replacement randomised controlled trials published in 2008 and 2013

<table>
<thead>
<tr>
<th></th>
<th>2008 (N=30)</th>
<th>2013 (N=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of participants</strong></td>
<td>(median, range) 85 (14-284)</td>
<td>81 (28-539)</td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>N (%): 18 (60)</td>
<td>20 (50)</td>
</tr>
<tr>
<td>Asia</td>
<td>2 (7)</td>
<td>9 (23)</td>
</tr>
<tr>
<td>USA</td>
<td>6 (20)</td>
<td>5 (13)</td>
</tr>
<tr>
<td>Canada</td>
<td>2 (7)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Middle East</td>
<td>0 (0)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Australia</td>
<td>2 (7)</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Funding</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent</td>
<td>20 (67)</td>
<td>15 (38)</td>
</tr>
<tr>
<td>Industry funded</td>
<td>9 (30)</td>
<td>12 (30)</td>
</tr>
<tr>
<td>Not specified</td>
<td>1 (3)</td>
<td>13 (33)</td>
</tr>
</tbody>
</table>
Figure 1. PRISMA diagram showing selection of included trials

2008
CCTR, Medline, EMBASE,
n=749

Excluded n=710
• No outcome of interest
• No control group
• Non-English language

For detailed review
n=39

Excluded n=9
• Non-arthroplasty intervention
• Non-English language

Included n=30

2013
CCTR, Medline, EMBASE,
n=886

Excluded n=842
• No outcome of interest
• No control group
• Non-English language

For detailed review
n=44

Excluded n=4
• Non-arthroplasty intervention

Included n=40
Figure 2. The Risk of Bias summary of included joint replacement trials in 2008
Figure 3. The Risk of Bias summary of included joint replacement trials in 2013
Figure 4: Number of trials with low risk of bias for different criteria in the Cochrane Risk of Bias tool comparing 2008 and 2013 joint replacement randomised controlled trials