

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association of patient characteristics and surgical variables on symptoms of pain and function over 5-years following primary hip replacement surgery: prospective cohort study
AUTHORS	Arden, Nigel; Judge, Andrew; Batra, Rajbir; Thomas, Geraint; Beard, David; Javaid, Kassim; Cooper, Cyrus; Murray, David

VERSION 1 - REVIEW

REVIEWER	Jeffrey N. Katz, MD, MSc Brigham and Women's Hospital Harvard Medical School Boston, MA USA I have no competing interests
REVIEW RETURNED	21-Dec-2012

THE STUDY	<p>The authors describe this as a population based study but to be eligible the subjects had to receive a Howmedica stem. Unless this was the only stem used in the geographic area, the study was not population based and should not be described as such.</p> <p>In analyses of predictors of change in score, why did you only include variables that were significant predictors of attained score? Isn't it possible that some variables are associated with change and not with attained score?</p> <p>The offset is essentially the distance between the center of the femoral head and the center of the shaft. You equate it with femoral component size and that is not quite correct. Surgical members of your team can clean this up.</p>
RESULTS & CONCLUSIONS	The authors should note that we do not know which of the two outcomes they use -- attained status or change in status -- is most appropriate for judging the value of THR.
GENERAL COMMENTS	I would be happy to write an editorial expanding on the issue of whether the final score or the change in score is the most appropriate metric for assessing surgical outcome. This research question interest me greatly. It ultimately hinges upon the priorities and values of the society that sponsors the procedures.

REVIEWER	Gillian Hawker MD MSc FRCPC Professor of Medicine/Rheumatology and Health Policy, Management and Evaluation Physician-in-Chief of Medicine, Women's College Hospital, University of Toronto, Toronto, Canada I have no competing interests to report.
-----------------	--

REVIEW RETURNED	08-Jan-2013
------------------------	-------------

THE STUDY	<p>The rationale for selection of variables for inclusion in the analyses is not clear - for example, were all possible and available surgical / intra-operative variables included or were there specific hypotheses being tested? No explicit hypotheses are presented.</p> <p>There are some relatively minor concerns regarding the methods: 1. DVT, PE and UTI are listed as co-existent conditions that are included in a comorbidity count for cohort participants - however, these are time limited conditions and would be highly unlikely to be present among individuals undergoing an elective surgical procedure. It is unclear why these are included. 2. 56 individuals who underwent bilateral hip replacement are included - the analysis has been performed at an implant level - still, there is likely to be correlation between hips within these folks - it is unclear how this was dealt with in the analyses. 3. with respect to the prosthesis/surgical covariates, there is no comment on the extent to which they were correlated - was collinearity assessed for? Might the surgeon make multiple decisions about approach based on the same one or two patient specific factors, such as age or sex? if so, how might this have influenced findings?</p>
RESULTS & CONCLUSIONS	<p>The overall surgical outcomes described are excellent (<2% experienced no improvement or worsening in their OA related symptoms) - while it is acknowledged that hip replacement recipients generally have better outcomes than knee replacement recipients, it is still unclear to me whether the cohort examined in this study has better than expected outcomes - for example, the figures indicate that a large proportion had achieved the highest (best) possible score for the OHS post operatively. Comparing this cohort's overall results with other large THR cohorts, such as the EuroHip cohort, would be helpful. Clarification of how OHS questionnaire was administered (by whom?) pre and post operatively would also help to address possible social desirability bias in responders.</p> <p>While the authors acknowledge that a number of previously identified predictors of joint replacement outcome were not available to them, greater discussion of how these variables might have impacted their findings would be useful. For example, there is no mention of the presence of comorbid knee/foot/ankle arthritis or low back pain among the cohort -- absence of comorbid MSK issues might help explain the high proportion of excellent outcomes, and in turn might reduce the generalizability of the study findings. Other predictors that are not mentioned include measures of coping/pain catastrophizing and social support.</p> <p>The authors recommend that greater attention be paid to choice of offset in women based on their findings, yet the variability in patient-reported outcome related to this factor is small. I would suggest that the next step would be to confirm these findings in other, large cohorts</p>
GENERAL COMMENTS	<p>Table 1 - it is unclear what values are provided in the table - e.g., under column 'Baseline' for OHS, are the values provided in the rows the mean OHS scores and ranges? median and IQR?</p> <p>Figure 1 and 2 - find presentation of the data in this format very informative!</p> <p>Supplementary table - says "OKS" in title - presumably a cut and</p>

VERSION 1 – AUTHOR RESPONSE

Response to reviewers

1) General

We are grateful to the reviewers for their helpful comments, and have amended the manuscript accordingly, we are confident that it has been improved as a result.

2) Specific

Reviewer: Jeffrey N. Katz, MD, MSc

- The authors describe this as a population based study but to be eligible the subjects had to receive a Howmedica stem. Unless this was the only stem used in the geographic area, the study was not population based and should not be described as such.

Thank you. We have amended this so the study is described as a prospective cohort study, and the words population based have been removed from the manuscript.

- In analyses of predictors of change in score, why did you only include variables that were significant predictors of attained score? Isn't it possible that some variables are associated with change and not with attained score?

For this paper our primary analysis was to identify predictors of attained score. Having identified predictors of attained score, our secondary question was then to describe how symptoms of pain and function changed over time between baseline and follow up assessments, stratified according to each predictor. The intention of this was to highlight that although there are statistically significant differences in attained score between for example, obese and non obese groups, these differences are greatly outweighed by the improvement (change) in symptoms achieved by patients in each of these groups.

We agree with the reviewer that predictors of attained score may not necessarily be the same as predictors of improvement (change) in symptoms. Both are different research questions and it depends whether the interest is to identify predictors of the difference in attained symptoms, versus the difference in slopes of how much symptoms change between pre and post-operative assessments. Our intention for this paper was to focus on predictors of attained score. In the introduction section of the paper we have made it clear that the focus was on predictors of attained score, and in the discussion section explained that determinants of attained score may not necessarily be the same as for change.

As part of the NIHR funded COASt (Clinical Outcomes in Arthroplasty Study) we have been taking this work forward to develop a clinical predictive tool, combining data from this cohort (EPOS) with additional data from the EUROHIP study providing us with additional information on radiographic variables. Within this ongoing work we have explored the development of a prognostic tool for both sets of outcomes (a) attained score and (b) change in score.

- The offset is essentially the distance between the center of the femoral head and the center of the shaft. You equate it with femoral component size and that is not quite correct. Surgical members of your team can clean this up.

We thank the reviewer for highlighting this. This has now been corrected in the paper and we have made it clear that it is component offset we are looking at.

- The authors should note that we do not know which of the two outcomes they use -- attained status or change in status -- is most appropriate for judging the value of THR.

We have taken this comment on board and amended the discussion section accordingly.

Reviewer: Gillian Hawker MD MSc FRCPC

- The rationale for selection of variables for inclusion in the analyses is not clear - for example, were all possible and available surgical / intra-operative variables included or were there specific hypotheses being tested? No explicit hypotheses are presented.

We thank the reviewer for highlighting that clarity is required over this point. Within the EPOS study there is an extensive number of patient and intra-operative surgical factors recorded in the study questionnaires. A-priori, we went through the variables with clinical co-authors, particularly focusing on the intra-operative variables on which little is known in the existing literature on their relationship with outcomes, and asked them to list those they considered relevant to include in an analysis of predictors of patient reported outcomes. The final list of variables considered as potential predictors for analysis are those listed in the methods section of the paper. This has now been made clear in the paper.

As discussed in the paper, given the extensive list of patient and surgical variables considered for inclusion in the model, we wanted to ensure that we minimised the possibility of making a type 1 error – e.g. the chance that a variable identified as being ‘significant’ in this dataset may not be replicated in other samples of patients. The use of bootstrap backward stepwise selection helped ensure this was minimised by only retaining variables in the final model that were consistently selected across the re-samples at least 70% of the time.

- There are some relatively minor concerns regarding the methods:

- 1. DVT, PE and UTI are listed as co-existent conditions that are included in a comorbidity count for cohort participants - however, these are time limited conditions and would be highly unlikely to be present among individuals undergoing an elective surgical procedure. It is unclear why these are included.

Thank you for pointing this out – the reason this is unclear is due to additional information not being given in the methods section. This has now been addressed. Prior to surgery the following yes/no questions were asked about coexisting diseases:

- (a) Has the patient ever had a Deep Venous Thrombosis (DVT)
- (b) Has the patient ever had a Pulmonary Embolism (PE)
- (c) Is there any evidence of UTI in the 4-weeks prior to surgery
- (d) Does the patient have any other Musculoskeletal disease
- (e) Does the patient suffer from any Neurological disease
- (f) Does the patient suffer from any Respiratory disease
- (g) Does the patient suffer from Cardiovascular disease
- (h) Does the patient suffer from significant Renal disease
- (i) Does the patient suffer from significant Hepatic disease
- (j) Is the patient currently receiving treatment for any other medical conditions

We have amended the methods section for clarity of the wording of data collected on co-existent diseases, and to highlight that DVT and PE are according to whether a patient ever had the condition,

and UTI was in the 4 weeks prior to surgery.

- 2. 56 individuals who underwent bilateral hip replacement are included - the analysis has been performed at an implant level - still, there is likely to be correlation between hips within these folks - it is unclear how this was dealt with in the analyses.

In the dataset a small number of 56 patients out of the total of 1375 had bilateral procedures. Patients with bilateral operation filled in a separate questionnaire for each hip. These patients were included in the analysis where the unit of analysis was the hip rather than the implant. The reviewer is correct in that there is likely to be correlation between hips for these patients. The statistical methods used already include a repeated measures analysis (using Generalised Estimating Equations) to account for clustering within the data due to the repeated measures of outcomes (e.g. to account for correlation between outcomes measures at the 5 follow up time points). The internal validation using a combination of multiple imputation and bootstrapping provides a further level of robustness to the analysis. Given the small number of bilateral patients, and the fact that the analyses already incorporate measures to adjust the standard errors and allow for clustering in the data, we did not consider this likely to affect the analysis.

- 3. with respect to the prosthesis/surgical covariates, there is no comment on the extent to which they were correlated - was collinearity assessed for? Might the surgeon make multiple decisions about approach based on the same one or two patient specific factors, such as age or sex? if so, how might this have influenced findings?

When evaluating the diagnostics of the regression model we assessed for evidence of collinearity. Collinearity implies that two variables are near perfect linear combinations of one another. As the degree of multicollinearity increases, the regression model estimates of the coefficients can become unstable and the standard errors for the coefficients inflated. We used variance inflation factors to check for multicollinearity between variables in the dataset and there was no evidence any merited investigation.

- The overall surgical outcomes described are excellent (<2% experienced no improvement or worsening in their OA related symptoms) - while it is acknowledged that hip replacement recipients generally have better outcomes than knee replacement recipients, it is still unclear to me whether the cohort examined in this study has better than expected outcomes - for example, the figures indicate that a large proportion had achieved the highest (best) possible score for the OHS post operatively. Comparing this cohort's overall results with other large THR cohorts, such as the EuroHip cohort, would be helpful.

We agree with the reviewer that there is a possibility that the outcomes in this study may be better than expected. The data from the EPOS study come from seven high volume centres with skilled surgeons across England and Scotland. In response to the reviewer we have made this clearer in the paper that the excellent outcomes are observed in the context of data from high volume centres, and in terms of generalizability, may not reflect expected outcomes in less skilled surgeons and lower volume centres.

We have previously analysed data from the EUROHIP cohort in relation to patient reported outcomes of primary THR. The EUROHIP study consists of patients coming to primary THR for OA across 20 European orthopaedic centres in 12 countries and the WOMAC score used to measure patients reported outcomes of surgery. In this study based on change in total WOMAC score between baseline and 12-month assessment, 58 (6.9%) of 845 patients had no change or worsening of symptoms – however this varied by region and country. Specifically, in patients from the UK, only 4 (3.5%) of 111 patients had no change or got worse. In other studies we have previously explored, in

the Elective Orthopaedic Centre (EOC) cohort 88 (5.1%) of 1711 patients symptoms got worse between baseline and 6-months based on the OHS, and in the St. Helier cohort 14 (2.3%) of 619 patients had no change or worsening in OHS at 12-months.

Hence, whilst it is plausible that outcomes in the EPOS study may be better than expected, the findings are consistent with other cohorts. We thank the reviewer for raising this issue, and have now expanded on this in the discussion section.

- Clarification of how OHS questionnaire was administered (by whom?) pre and post operatively would also help to address possible social desirability bias in responders.

Pre- and post-operative scores were completed independently by the patient just prior to clinical examination and this has now been clarified in the methods section.

- While the authors acknowledge that a number of previously identified predictors of joint replacement outcome were not available to them, greater discussion of how these variables might have impacted their findings would be useful. For example, there is no mention of the presence of comorbid knee/foot/ankle arthritis or low back pain among the cohort -- absence of comorbid MSK issues might help explain the high proportion of excellent outcomes, and in turn might reduce the generalizability of the study findings. Other key predictors that are not mentioned include measures of coping/pain catastrophizing and social support.

We agree that the paper would benefit from greater discussion of how predictive variables not available to us in this study could impact the findings – we have now added to this in the discussion section.

- The authors recommend that greater attention be paid to choice of offset in women based on their findings, yet the variability in patient-reported outcome related to this factor is small. I would suggest that the next step would be to confirm these findings in other, large cohorts

We have amended the paper to state that this finding requires confirmation in other large cohorts.

- Table 1 - it is unclear what values are provided in the table - e.g., under column 'Baseline' for OHS, are the values provided in the rows the mean OHS scores and ranges? median and IQR?

This has been made clear in the footnote of the table.

- Figure 1 and 2 - find presentation of the data in this format very informative!

Thank you.

- Supplementary table - says "OKS" in title - presumably a cut and paste error?

This has now been corrected.