Author's response to reviews

Title: A comparison of Bipolar hemiarthroplasty and total hip arthroplasty for displaced femoral neck fractures in the healthy elderly: A Meta analysis

Authors:

Feng Wang (wangfengdr_cmu@163.com)
Haifei Zhang (actor_1999@126.com)
Zhiyu Zhang (zhangzhiyudr_cmu4h@163.com)
Chengbin Ma (18900912433@163.com)
Ximin Feng (Fxm1972@126.com)

Version: 4 Date: 26 June 2015

Author's response to reviews: see over
Dear Editor:

We appreciate the comments of the two reviewers. The reply to them will be described as the following. Moreover, we voluntarily re-identified all of the RCTs from the sources we described in our manuscript and found that there was no other RCTs, so we updated the date of searching to May 2015.

Comments of reviewer (Gerard Slobogean)

#1: Line 167, 169, 171: “heterogeneity was obvious” it is not clear as to what it refers to as a stand alone sentence, particularly since the sentence is repeated several times within the paragraph and results sections.

We rewrote this section with more clarity and easy scientific language followed your suggestion. Please see details in the results section for HHS:

Line 167 $P = 0.06$), and heterogeneity across the studies was obvious ($I^2 = 60\%$).
Line 169 but heterogeneity across the studies was obvious ($I^2 = 88\%$).
Line 172 and heterogeneity across the studies was also obvious ($I^2 = 90\%$).

#2: Please clarify in the text if the point estimates throughout the results are comparing BHA to THA, or vice versa? Even though most of the comparisons are non-significant, it is useful to know which treatment the point estimate favors. I realize this information is in the forest plots.

The aim of this study is to evaluate the clinical outcomes, comparing BHA with THA. So we revised the manuscript following with the reviewers’ comments. And we re-conduct this meta analysis, comparing BHA to THA, and rewrote the text to clarify this point with revised data and uploaded the revised pictures.

Please see details in the text (some of them are listed):

Line 87: study is to evaluate the clinical outcomes comparing BHA with THA
Line 166: versus 240 with THA
Line 168 :BHA versus 81 with THA
Line 170 :BHA versus 81 with
Line 177 :BHA versus 260 participants with THA
Line 179 :BHA versus 111 with
Line 182 :BHA versus 224 with THA
Line 184 :BHA versus 204 with THA
Line 190 :BHA versus 184 with THA
Line 197 :BHA versus 220 with THA
Line 207-208: BHA versus 56 with THA
Line 210: BHA versus 98 with THA
Line 217 :BHA versus 297 with THA
Line 250 :BHA versus 491

#3: I am confused why the authors have reported both pooled results and individual study results within the secondary outcomes sections. I understand if they are unable to pool results for an outcome, but I do not understand the rationale for reporting both.

Actually, we did not want to report individual study results in section “Results”, but only one study included in such subgroup analysis. We rewrote this section with more clarity language followed your suggestion.
Line 251: the THA tended to be higher, however the WMD was often very small and likely not clinically significant. Please comment on the clinical significance of your results.

In the discussion section, we, in statistics, discussed the possible reason that caused HHS tended to be higher in THA, but did not comment the clinical significance of the result. According to the reviewer’s opinion, we carefully reviewed reference 31 and relevant literatures. We agree with the author’s point of view of ref.31: even THA may lead to better outcomes, inappropriate implants can cause adverse clinical outcomes. And in paragraph 3 of discussion section, we do the following comments:

1) The results suggested that proper implants had great influence on the patients, and hence, might influence' was written to 'The results might influence'
2) while the HHS favored THA after 2 years without statistical difference' was written to while the HHS favored THA after 2 years without statistical difference. It was suggested that, even though THA might lead to better clinical outcomes, proper implants were of great importance for the patients.

We comment on the potential causes of the moderate-high heterogeneity in section “Discuss”, please see details in LINE 312-317.

We pooled most data of comparable parameters using subgroup analysis with different follow-up periods, which may reduce the bias as much as possible. Actually, when we conduct the pooled analysis of HHS for all time points combined as your suggestion, the results show the point estimates marginally favored THA without statistical difference (MD-4.12; 95% CI-10.14,1.91; $I^2$ =85), which was similar to the previous subgroup analysis results. So, we still believe that the data could not be directly pooled considering different follow-up durations in these studies.

Comment of reviewer (Yoram Weil)

1) The Harris hip score is indeed an important outcome measured but its reliability and validity had been severely criticised due to a "ceiling effect". It is of rather low specificity. Therefore the main "surprise" factor of this analysis is not such a great surprise and I think it is extremely overstated in the conclusions, abstract and results section. If there are other outcome scores in the quoted studies they should be brought as well. The fact that there are no differences in a very heterogeneous group of patients is the least surprising thing. Pain scores, mobility indices and implant survival, as well as reoperations are far more important. Therefore the whole conclusions of this study should be reorganised.

We carefully reviewed the related literatures and we accepted the reviewer’s opinion that it is not such a great surprise and it is extremely overstated for the HHS results of comparing BHA and THA. Therefore we deleted all of the inappropriate adverb “surprisingly” in the conclusions, abstract and results sections and modify the conclusions, in order to accurately and faithfully summarize our results: For the healthy elderly with displaced FNFs, treatment of BHA led to better outcomes in dislocation rate while THA in acetabular erosion rate and reoperation rate. And comparing BHA with THA, there were no
significant differences in some other important outcomes, such as HHS, infection rate, general complication and one year mortality. It was suggested that further high-quality RCTs were needed to provide robust evidence and evaluate the treatments for the patients.

#2: Comparing a 65 years old active patient with a 85 years old household ambulator with severe comorbidities is comparing apples with oranges. This is the most important major limitation of the study. The main interest in THA for FNF is due to the former group of patients. Without even trying to stratify the study for the more young, active patient group is a futile attempt. It should be at least tried by all means to create a sort of a stratification of the patient group according to age, comorbidities and preoperative ambulatory status these are the main concerns that should be addressed prior to consideration for publication.

This issue is of great importance for our analysis, which we have seriously reviewed and discussed before we began with our manuscript. We all agreed that it might lead to better results and conclusions, if we subgrouped the pooled data either according to age, comorbidities and preoperative ambulatory status of the patients. However, in the present analysis, only 8 studies were included and each study has its patient age distribution, such as Dorr’s study (range, 41-89 years), but could not be pooled and subgrouped. Moreover, because of data deficiency of comorbidities and preoperative ambulatory status in the studies, we could not conduct the subgroup analysis according to age, comorbidities and preoperative ambulatory status and so on. And as we have commented in the discussion section, the outcomes of these studies were presented with different follow-up periods, which could be pooled and stratified, we conducted this analysis according to the follow-up durations, which might also reduce the bias in part. And we also thought that, when further RCTs on such subject were designed, clinical outcomes presented according to age, comorbidities or preoperative ambulatory status should be taken into more consideration.

Finally, we appreciate and thank the reviewers and editors very much for their time and job, and we are looking forward to their further comments on our analysis.

BEST REGARD

Zhiyu Zhang