Author's response to reviews

Title: Fulfilment of patients' and surgeons' expectations in total hip arthroplasty: a 1 year follow up study

Authors:

Clémence Palazzo MD (clemence.palazzo@cch.aphp.fr)
Claire Jourdan MD (claire.jourdan@rpc.aphp.fr)
Stéphane Descamps MD (s.descamps@chu-clermontferrand.fr)
Remi Nizard MD PHD (remi.nizard@lrb.aphp.fr)
Moussa Hamadouche MD PHD (moussa.hamadouche@cch.aphp.fr)
Philippe Anract MD PHD (philippe.anract@cch.aphp.fr)
Stéphane Boisgard MD PHD (sboisgard@chu-clermontferrand.fr)
Myriam Galvin (mgalvin@chu-clermontferrand.fr)
Philippe Ravaud MD PHD (philippe.ravaud@htd.aphp.fr)
Serge Poiraudeau MD PHD (serge.poiraudeau@cch.aphp.fr)

Version: 3 Date: 7 October 2013

Author's response to reviews: see over
Sir,

Please find enclosed the manuscript entitled “Fulfilment of patients' and surgeons' expectations in total hip arthroplasty: a 1 year follow up study” to be assessed for possible publication in BMC Musculoskeletal Disorders as a research article.

In a previous work (Jourdan et al, Comparison of patient and surgeon expectations of total hip arthroplasty. *PLoS One.* 2012), we have compared preoperative patients' and surgeons' expectations of total hip arthroplasty (THA), and found that they do not agree on what to expect. The present manuscript reports the results of the postoperative assessment. We believe these data may have a great interest for a general medical audience, for rheumatologist and orthopedic surgeons. These results may help to better identify patients at risk of dissatisfaction after THA, to better manage them before surgery, and so to improve the quality of care.

This article was previously submitted to Arthritis Research & Therapy that was not able to accept the paper. It was said that you would be happy to consider our manuscript for publication if the peer reviewer comments were addressed and the appropriate revisions were made. You will find enclosed the revised manuscript and the answers to the reviewers.

**Response to Reviewer 1:**

*Abstract*

*Study aim is not consistent with objectives in introduction. It is unclear what “domains” refer to in the first paragraph. A better organization of results of abstract is sought. Last sentence in results paragraph: what does “its” refer to? I disagree on the conclusion that surgeons have reliable expectations of postoperative satisfaction. Are the results of 16 surgeons generalizable? It is not possible from the study to tell if the surgeons expectations were conveyed to the patient and how patients were informed about expected outcomes. Thus, it is not possible to draw the conclusion that surgeons could improve preoperative information on expected outcomes.*

We modified the abstract taking into account these comments. Study aim is now in accordance with the objectives in introduction. The word "Domains" has been removed. The results have been reorganised. In the conclusion, we removed as you suggested that surgeons’
expectations were conveyed to the patient. However, we think that it is justified to conclude that surgeons could improve information on postoperative outcome.

Introduction

- The authors state that the prevalence of OA is increasing but the reference is too old to support an ongoing increase.


Why wouldn’t further technical progress improve patient-reported outcomes following THA?

We didn't want to say that further technical progress wouldn't improve PRO, but that a lot of progress have been made in this area, and that we may interest in other aspect of THA management to improve patients' satisfaction, such as a better understanding of patient's expectations and their fulfilment. We replaced "might" by "may" in the following sentence: "Important technical progresses have already been made in THA, so future progress in this field may not significantly impact patients' satisfaction."

- I disagree with the statement that there are few predictors of satisfaction identified. There are a number of risk factor for dissatisfaction identified, e g higher age, female sex, diagnosis, more co-morbidities, other conditions impairing with walking capacity, mental distress, higher preoperative pain, and low socioeconomic status, low expectations (Hussain et al, Rolfson et al 2009 and 2011).

We took this comment into account and modified the text: "A number of preoperative risk factors of dissatisfaction have been identified: higher age, female gender, more co-morbidities, other conditions impairing with walking capacity, mental distress, higher preoperative pain, and low socioeconomic status [9, 14, 15]; but there is little evidence for a strong effect for any of these factors."

- Second paragraph “joint injury” – do they mean “joint disease”? Yes. "Joint injury" was changed in "joint disease".

- Objectives are not clear. The third objective seems to be incorporated in the second.

We rephrased the objectives as you suggested, writing:
"The objectives of this study were to determine factors associated with patient dissatisfaction one year after THA, and to better address the role of patients’ and surgeons' expectations in postoperative outcome."

instead of

"The objectives of this study were: (i) to describe in which domain patients' and surgeons’ expectations were fulfilled or not, (ii) to determine factors associated with patient satisfaction one year after THA, including patients' and surgeons’ expectations and their fulfilment, (iii) to determine factors associated with the fulfilment of patients' expectations”.

Patients and methods
- It is not clear if the complete survey was done via telephone. If so, are all instruments included in the study protocol validated for telephone use in French?

Yes, the complete survey was done via telephone. We explained it more clearly in the text: "The present work is a longitudinal telephone-based follow up of a sample of patients who participated in a previous study". As most of instruments, it was not specifically validated for telephone use (whatever the language).

- The authors describe that fulfillment of patients’ expectations were measured “with the same tool” referring to the HSS THR expectation survey. As far as I know, there is no follow-up version to score the actual patient-reported outcome in the different domains of the expectation survey. Mancuso et al (2009) describe a telephone based follow-up protocol to investigate if and to what extent the expectation was fulfilled similarly to the description in the present study.

The HSS THR expectation survey has not been validated for assessment of expectations’ fulfilment, but no score has been validated in this situation. However, it was used by Mancuso and Scott, and might be a more valuable and standardised method than other instruments/questions used in previous studies to assess expectations' fulfilment (see the last paragraph of the discussion).

- It is not clear from the methods if the patients were informed about their preoperative expectation when they were asked about the extent of fulfillment. I believe there is a problem
when the patient (or the surgeon) preoperatively didn’t have any expectations if you pose the question “to what extent have you obtained relief or improvement...” and ask them to grade it from “not at all” to “completely” or “does not apply”. The summary score is problematic because the preoperative and postoperative values cannot be compared because they don’t measure the same thing. Further, the number of not applicable items is not shown and I believe that will significantly affect the interpretation of the summary score.

We added to the methods: "The patients were not informed of their preoperative answers. If the item was not applicable preoperatively, we considered that it was also not applicable postoperatively."

The non applicable items were not taken into account in the global sum. The method of calculation is not ours, but the method proposed originally during the development of the HSS THR expectation survey.

The number of applicable items is shown under brackets for each item on figure 1.

- The validity and of the HSS THR expectation survey has been questioned and as far as I understand the documentation is flawed (Suk et al, AO Handbook Musculoskeletal outcome measures and instruments, Vol , 2nd ed 2009). The fulfillment calculation also seems to be problematic. Suppose a patient have no expectations assessed as equal to the preoperative rating (no expectations). I apologize if I have misunderstood the method but if this criticism must be addressed properly.

Even if the HSS THR expectation survey has been questioned, no other standardised assessment of expectations has been proposed. Qualitative studies have shown interesting results, and may result in preference-based questionnaires, but none of this questionnaire have been yet developed.

If the patient had no expectations, the item was considered as not applicable. When the preoperative and the postoperative rating were equal, the score for the item was 0.

- The satisfaction question is also questionable. Are the hip symptoms over the last 24 hours representative for the satisfaction of the intervention?

To our knowledge, the only one scale that has been validated to evaluate the satisfaction of the patients after arthroplasty is the visual analogue scale (Brokeman, Eur Orthop Traumatol 2012); but we could not use it 1) because we had already finished this study in june 2012; 2) the visual analogue scale is not adapted for telephone survey. Consequently, even if the question “If you were to spend the rest of your life with your hip symptoms just the way they
have been in the last twenty-four hours, how would you feel?’’ may be questionable because the symptoms of the last 24 hours may be not representative of the postoperative symptoms, it may be more concrete for patients, and this question has been validated for general well-being (Mancuso 2009, Andrews 1976), and already used in post-THA satisfaction assessment (Mancuso 2009).

- SF-12 does not measure quality of life, it measures health-related quality of life and those are not interchangeable.
The modification was done.

- Pain and trust questions using visual scales are insufficiently described.

- Statistics methods: were any other sociodemographic variables but age and sex collected?
Yes. We added these informations in the last paragraph of evaluation: "Demographic characteristics included gender, professional category, and marital status. Health status evaluation included age, Body Mass Index (inferred from patients’ reports of height and weight). Co-morbidities were measured using the Charlson Comorbidity Index (0: no relevant comorbidity versus one or more co-morbidities). A history of ipsilateral hip arthroplasty was also recorded".

- The authors mention they built two models – what does this refer to?
We explained the 2 models more clearly:
" All variables which were associated with satisfaction in univariate analysis at the 0.2 level were initially included; the best model was selected using second-order Akaike information criterion. We built two different models, in order firstly to assess pre-operative predictive factors of satisfaction (model 1), and secondly to assess pre- and post-operative factors associated with satisfaction (model 2)."

- How and what were complications recorded?
They were recorded according to patients' report and medical record as free text.
- How were co-morbidities recorded?
Co-morbidities were measured using the Charlson Comorbidity Index (0: no relevant comorbidity versus one or more co-morbidities).

- Results
  Move preoperative characteristics to methods section?
It is an eventuality. However, we think that it is better to leave the characteristics of the population in the results, and more readable to have pre- and post-operative characteristics in the same table (that is table 1).

- Skip evaluative words in results section such as “Complications were rare”.
This was done " We observed 9 complications (7.3%): 4 dislocations, 4 fractures (3 during surgery and 1 through a fall the day after surgery), 1 severe cutaneous reaction to the bandage. Twenty-one patients (17%) reported a residual limp."

- It is not clear how the two multivariate regression models were built and what decides which variables to include.
This is explained in the "statistic analysis": "All variables which were associated with satisfaction in univariate analysis at the 0.2 level...(that are described in table 2)...were initially included; the best model was selected using second-order Akaike information criterion."

- Discussion
  Indistinct wordings: “These finding...”? What findings? “They suggested...”? Who or what suggested? “...proposed if necessary.”
We modified the discussion to make it understandable.

- Nothing about limitations/weakness of questionnaires.
We added a part about the limitations of questionnaires. " Another weakness of this study is that no approach has been validated to assess the fulfilment of expectations. Several authors used a non standardised retrospective assessment, with an important risk of recall bias [11, 16]. Mancuso et al [13], using the same questionnaire as we did, defined the fulfilment of expectations as the percentage of patients whose expectations were fulfilled completely in
each domain. Our approach is close to the method recently described by Scott et al [22], although it is unclear how unapplicable items were taken into account in their study."

- I don’t agree that the predictions of satisfaction has not been reported before.
We modified the discussion to take into account this comment.

- Conclusions: Make summary/conclusions more distinct and make sure it follows objectives.
We took it into account and modified the conclusion.
"The main determinant of dissatisfaction 1 year after THA is the absence of fulfilment of patients’ expectations, independently of their preoperative level. An older age, a worse mental wellbeing and a higher disability are associated with a poorer expectations’ fulfilment after surgery, and should be specifically managed before surgery. After surgery, expectations fulfilment is mainly determined by postoperative function and the absence of residual pain. Surgeons should better inform their patients of the expected outcomes to avoid disappointment after surgery, particularly for relieving night pain and removing the need of a stick. This study highlights the urgent need to develop a valid tool to standardize the assessment of expectations and their fulfilment."
instead of
"Surgeons could have reliable expectations to predict satisfaction after THA, and should inform their patients better on the expected outcomes, particularly for relieving night pain, or removing the need of a stick. Their expectations are still too optimistic for cutting toenails, putting on shoes, improving sexual activity, sport and professional activity; they need to be aware of such limits of THA, in order to inform their patients more precisely before surgery. The fulfilment of patients’ expectations, independently of their preoperative level, seems to be the main determinant of postoperative satisfaction. An older age, a worse mental wellbeing and a higher disability are associated with a poorer expectations’ fulfilment after surgery, and should be specifically managed before surgery. After surgery, expectations fulfilment is mainly determined by postoperative function and the absence of residual pain."

- References: Not completely up to date.
We added new references and took these new data into account in the text (particularly in the introduction and the discussion). See references: 1, 17, 33, 34 and 38.

Response to Reviewer 2:
1. The follow-up duration is less than the typical 2 years required for arthroplasty outcome assessment. In addition, according to the information provided regarding inter-quartile range, it appears that 25% of patients had less than 10 months of follow-up. The authors need to address this issue and why they think their patients reached their final postoperative state in such a short time.

We disagree with the "typical 2 years required"; most of the studies interesting in satisfaction and fulfillment of expectations after THA have a shorter follow up (from 6 months to 1 year) (See for example: Scott CE, J Bone Joint Surg Br 2012; Meijerink, Arch Orthop Trauma Surg 2009; Anakwe, J Arthroplasty 2011; Mahomed J Rheumatol 2002; Fortin Arthritis Rheum 1999;...). Moreover, it has been shown that the largest gains is seen within the first 3 to 6 months after THA (Jones, Rheum Dis Clin N Am 2007).

Even if 25% of patients had less than 10 months follow up, the rate of satisfaction was very good and similar to previous studies; the rate of dissatisfaction was not more important in this group, and the time after surgery did not significantly impact on postoperative satisfaction or expectations' fulfilment.

2. One of the benefits of the “Delighted – Terrible” scale is gradations of satisfaction. Dichotomizing the group into satisfied or dissatisfied dilutes the value of this scale. Obviously the authors did so because of the small sample size. Nonetheless, the authors at least should report fulfillment scores for different categories in the text.

We took into account this comment and modified the text (see the results): "Overall, 91.9% (n=113) of patients were satisfied after THA (52 were delighted, 39 very satisfied and 22 mostly satisfied); so 8.1% (n=10) were dissatisfied (5 had mixed feelings, 4 were unhappy and 1 felt terrible).

3. Presenting data according to center is not useful unless the authors describe what is different among these centers that may impact THA expectations or outcomes. Unless there are tangible differences, comparing by site is not instructive and is distracting and could imply favoritism.

We suspected a clustering effect, and so used a mixed effect model. Consequently, it seems necessary to present results according to centre. We anonymized centres (using "centre 1", "centre 2", "centre 3" instead of the centres name) to avoid the feeling of favoritism.

4. The authors should also present WOMAC and SF-12 data in terms of change in pre- to postoperative scores.
We think that it would be confusing and not informative to present preoperative Womac and SF12, postoperative Womac and SF12, and the change in these scores. Thus we only presented pre- and postoperative scores that are used as variables in the regression models.

5. In the discussion the authors need to comment on why they think preoperative WOMAC scores did not remain a variable in the model, as other studies have demonstrated that patients who are the most seriously impaired preoperatively are less likely to have good outcomes.

The modification was done: "Although other studies have demonstrated that patients who have the worse preoperative function are less likely to have good outcomes [35-37], the preoperative Womac did not remain a variable in the final regression model, probably because we adjusted on surgeons' expectations (what others authors did not); as it was previously discussed, surgeons mainly rated their expectations according to preoperative function."

Sincerely,

Clémence Palazzo MD