Author’s response to reviews

Title: Histopathological characterization of corrosion product associated adverse local tissue reaction in hip implants: a study of 285 cases.

Authors:

Benjamin Ricciardi (bfr2001@nyp.org)
Allina Nocon (nocona@hss.edu)
Seth Jerabek (jerabeks@hss.edu)
Gabi Wilner (wilnerg@hss.edu)
Eliana Kaplowitz (kaplowitzel@hss.edu)
Stephen Goldring (goldrings@hss.edu)
P Purdue (purduee@hss.edu)
Giorgio Perino (perinog@hss.edu)

Version: 1 Date: 12 Dec 2015

Author’s response to reviews:

Reviewer reports:

Reviewer #1: The authors present a well-structured paper where background, methods, results and conclusion are fully described. Although, I do have some recommendations regarding the research work presented in the paper.

The first issue regards the patient's physical activity. At no point along the paper, do the authors mention this issue. Another biomedical aspect that I consider very critical for discussion is the issue regarding the bone density, before implantation of the implants and after when performing
the revision. It would be very interesting to observe these results. Did the patients gain or loss bone density along the process?

Thank you for these interesting comments. The issue of physical activity could be important in correlation to the amount of implant wear, but of less relevance to the host immunological response to the wear particles, unless it could be demonstrated that different levels of physical activity influence the type of wear generated at the metal-on-metal surfaces. There is also no general consensus on which activities would be indicated or not and their respective level. Moreover, our retrieval and clinical registries do not record in any standardized form the patients’ pre-operative and post-operative physical activity levels as a standard of care. Previous studies by other authors have shown no correlation between physical activity and raise in cobalt and chromium ion serum levels [Heisel C et al. JBJS Am (2005) 87A: 781-787; Khan et al JBJS Br (2008)] 90B: 1152-1157]; b. The bone density during implantation time could affect the rate of macrophagic infiltration in the cases of osteolysis and/or implant loosening but would not influence the biological reaction in the periprosthetic soft tissues, which is the core observation of our paper, unless it could be proven that early contact of particle laden macrophages with hematopoietic marrow elements would stimulate macrophagic recruiting from myeloid cells. Moreover, the standard of care of patients at our institution consists of plain radiographs before and after the operation with or without MRI examination to assess adverse tissue reaction at the surgeons’ discretion. Proper assessment of bone density requires a more sophisticated method, such as high-spatial-resolution bone densitometry with dual-energy X-ray absorptiometric region-free analysis [Morris et al. Radiology (2015)274: 532-539], which is not currently standard of care in these patients and was not performed. This could be an interesting issue for future prospective evaluation of these patients, although the time zero at implantation for comparison would be missing and also depending on the occurrence of the primary and revision surgery at the same institution. This has been added to the manuscript on page 26, line 592-601: “The third is the absence of the following sets of clinical data: a. physical activity pre and post-operative, although it has shown a weak correlation to elevated serum metal ion levels, suggesting that activity-related bearing surface wear plays only a minor role in elevated serum cobalt or chromium levels [65,66]; b. pre and post-operative bone density, which may influence the occurrence/rate of implant mechanical loosening/osteolysis especially in the female population which requires a sophisticated method for proper assessment, such as high-spatial-resolution bone densitometry with dual-energy X-ray absorptiometric region-free analysis [67], which is not currently performed as standard of care at our institution;

The second issue regards a more mechanical observation. Did the authors observe the surface roughness after the revision of the implant? I would like to know if there is a correlation (direct or indirect) between the several medical cases and the surface roughness of the implants.
Thank you for this interesting question. Previous studies reporting medical complications related to hip prosthesis have shown elevated serum cobalt and chromium ion levels in association with these systemic symptoms [Bradberry SM et al. Clin Toxicol (2014) 52(8): 837-8477]. The limited medical cases we had in our series also displayed significantly elevated serum ion levels and in particular cobalt when recorded, suggesting that the type/amount of wear plays some role in systemic toxicity. Our study did not focus on correlation between biomechanical analysis and histological findings, and in some cases the acetabular component was retained at revision making this analysis not possible in all cases. Our study focused on the histological analysis of the reaction to the particulate wear material in three classes of implants and on the comparison between two classes of MoM bearing surface implants and one class of Non-MoM bearing surface with a CoCr dual modular neck to highlight the importance of corrosion products in the occurrence of the adverse reactions. However, we highlighted in figure 5 the change in elastohydrodynamic lubrication regime secondary to the change in thickness of the film due to cell debris and corrosion products [Gao L et al. J Biomed Mater Res B (2015)DOI: 10.1002/jbm.b.33568]. Retrieval analysis and blood metal measurements contribution to the understanding of ALTR has been previously addressed in a comprehensive review and no clear dose-response relationship between wear and ALTR could be established [Campbell P et al. Clin Orthop Relat Res (2014) 472: 3718-3727] and could also contribute to the partial or total loss of protective coating of the bearing surface as recently reported [de Villiers D et al. Proc Inst Mech Eng H (2015) 229(11): 804-811]. The modification of the lubrication film during implantation is a topic of great interest of further study in vivo. This has been added to the manuscript on page 26, line 601-606: “The third [limitation] is the absence of the following sets of clinical data: c. wear analysis by biomechanics examination of the metal-on-metal implants for surface roughness, although retrieval analysis and blood metal measurements contribution to the understanding of ALTR has been previously addressed in a comprehensive review and no clear dose-response relationship between wear and ALTR could be established [68]. “

Reviewer #2:

Thank you for the opportunity to review this interesting manuscript.

Given the nature of the study, it should be reported in accordance with the STROBE guidelines and a flow chart should be included that should commence with all revision THAs performed in the institution in the inclusion period (June 2011 to December 2014) and demonstrate each step of the screening and recruitment to the study cohort, concluding with the final numbers available for analysis.
Thank you for the comment. We provided the flowchart as suggested in Figure 1.

For the exclusions due to infection, the number of samples taken intraoperatively and the number of positive samples required for a diagnosis of infection to be made should be stated.

The diagnosis of infection is established at our institution in compliance with the criteria of periprosthetic joint infection (PJI) as reported by the International Consensus Meeting on PJI and accepted by the Centers for Disease Control (CDC). These details have been added to the manuscript on page 7, line 159-162: “Exclusion criteria included infection diagnosed in compliance with the criteria reported by the International Consensus Meeting on periprosthetic joint infection and accepted by the Centers for Disease Control [22] [N=3] with 5 out of 5, 6 out of 6, and 5 out of 5 intraoperative positive cultures…”

The exclusion of the periprosthetic fracture and recurrently dislocating case is not justified given both of these pathologies may occur in association with ALTR.

We agree that both pathologies can occur in association with ALTR. The total number of patients excluded was 2 and occurred in the Non-MoM DMNTHA group. The revision occurred at less than 6 months after implantation without evidence of ALTR at histological examination of the retrieved periprosthetic soft tissue. This has been added to the manuscript on page 7, line 165-166: “…two cases for non-ALTR related post-operative complications with histological examination: periprosthetic fracture [N=1] and recurrent dislocation [N=1].”

What number of cases of metal-on-polyethylene THAs with non-modular necks with head sizes 28mm and above did the authors observe in their cohort? The number of cases reported in this cohort is increasing (http://www.ncbi.nlm.nih.gov/pubmed/23289127; http://www.ncbi.nlm.nih.gov/pubmed/26224816).

Thank you for this interesting question. We agree that this is an issue that is receiving more attention and is reported more frequently. We recorded 18 cases during the time period extracted from the department of pathology data base. These patients were excluded from our analysis because they initially presented with unexplained pain or other symptoms with a conventional MoP implant without CoCr dual modular neck. Therefore, the diagnosis of ALTR was
unexpected preoperatively (infection and unexplained pain were most common preoperative diagnoses) and the patients were not identified, consented, and enrolled into our registry. Our findings are similar to those recently reported [Whitehouse MR et al. Adverse local tissue reactions in metal-on-polyethylene total hip arthroplasty due to trunnion corrosion. Bone Joint J (2015) 97-B: 1024-1030]. We agree that a separate, retrospective study of this cohort in correlation with biomechanical retrieval analysis of the implants would be warranted. The occurrence of these cases has been addressed in the section Methods, paragraph Patients, page 7, line 167-171: “In addition, 18 cases were identified with a post-operative unexpected diagnosis of ALTR in conventional MoP implants without dual modular neck that were not consented for inclusion in our registry prospectively and were not eligible for enrollment in the current study. “

The extensive sampling strategy is described with the comment that the strategy was at the discretion of the operating surgeon, it needs to be made clear how many cases such samples were actually taken in and hence available for analysis by each cohort as this may influence the reported results.

We stated that the bone samples and not the soft tissue samples were at the discretion of the operating surgeon. The discretionary sampling of the acetabular and/or femoral bone can provide additional and valuable information, although only when feasible without compromising the success of the revision procedure and/or implant stability. We also added the average number of tissue blocks per group which contained one or two tissue sections per block according to the thickness of the pseudocapsule/neo-synovial membrane, section Methods, paragraph Tissue collection and sampling, page 9, line 211-215: “The mean number of individual surgical specimens between the groups was not different [DMN cohort was 4.3 (SD 1.5), for the MoM THA cohort was 3.5 (1.6), and for the resurfacing cohort 3.5 (1.3); p>0.05]. Extensive samples between 5 and 15 tissue blocks containing one or two histological sections were taken depending on the available tissue for each specimen. “

Given two observers performed the histological examination, how were the results correlated and how was disagreement between observers handled? I assume all the cases were examined by both observers?

All cases were examined by both observers. Disagreement was handled by consensus between the two observers. This method of grading and assessment has been reported and validated in
previous publications of MRI/pathology correlation, histological analysis, and also for intraobserver variability [Nawabi DH et al. Clin Orthop Relat Res. 2014; 472(2):471-81]. This comment has been addressed in the section Methods, paragraph Patients, page10, line 223-226: “All cases were examined by both observers. Disagreement was handled by consensus between the two observers. This method of grading and assessment has been reported in previous publication [17] and also validated for intraobserver variability [28].”

No power calculation has been performed yet the authors quote highly significant p values. When groups are compared either between groups or in the same group at different time points, this should be done. Given the nature of the study, a post hoc analysis should be performed.

This is a valid point; however, the purpose of this study was to describe the histological patterns of soft tissue failure ALTR. An a priori power calculation will be done for future studies that will further evaluate comparisons between the specific cohorts explored in this study. We did not specify a primary end-point or primary hypothesis for this study as our intent was simply to illustrate the findings in all available patients at our institution. The purpose of our univariate comparisons was to serve as preliminary data to help determine if there were any indications for additional research among these groups. We have revised the manuscript to clarify this intent. This was discussed with our statistician (Allina Nocon) who is an author on the paper. We agree that a post hoc analysis would be beneficial and this was added to the throughout the results section. In addition, Table 6 has been modified to show the p-values based on post-hoc pairwise comparison corrected for multiple comparisons.

How was accurately defined the onset of symptoms and what symptoms specifically are the authors referring to?

The onset of symptoms was assessed retrospectively at the time of revision surgery based on patient questionnaire and verified for possible discrepancies with clinical charts at follow-up visits. Symptoms included increasing pain around the hip and mechanical symptoms such as “grinding “sensation. Other symptoms as discomfort around the hip, although frequent in the non-MoM group were not considered positive unless progression to pain was recorded before revision. This issue has been addressed in the section Methods, paragraph Patients, page 8, line 183-188: “The onset of symptoms was assessed via questionnaire at the time of revision surgery. Symptoms included increasing pain around the hip and mechanical symptoms such as “grinding
sensation”. Other symptoms such as discomfort around the hip, although frequent in the Non-MoM DMNTHA group were not considered positive unless progression to pain was recorded before revision.”

The comparison between the three cohorts in terms of the histological patterns observed is interesting but currently lacks context, a comparison against a cohort of patients without a diagnosis of ALTR at the time of revision surgery (e.g. aseptic loosening of MoP THA) would be of great benefit.

We have added a group of patients from the registry from the same study time period who failed due to aseptic loosening/osteolysis of MoP THAs without dual modular neck for the histological analysis. The issue has been addressed in the section Methods, paragraph Histological Analysis, page 10, line 233-237: “All patients enrolled in our registry over the same time period with a diagnosis of aseptic loosening due to osteolysis with conventional MoP implants without dual modular neck were subject to the same tissue collection, histological analysis, and scored to serve as non-ALTR controls for pathological data [N=31].”

It is difficult to accept the authors' assertion that the absence of particle laden macrophages in some reported cases may be due to tissue sampling bias as surely all of the samples and conclusions drawn from this analysis are subject to the same possible bias?

Thank you for this question. In our examination of more than 250 cases of ALTR either as primary occurrence or as consultation cases we did not observe any case with pure lymphocytic pattern, even with less than 6 months of implantation. Moreover, no cases with this pattern have been reported before in the literature, besides in the recent report by Berstock JR et al. Hip Int (2014) 24(3):243-248. Our observation is logical because the macrophages are present in the neo-synovial superficial cell layer and are the first cells to respond to the particle wear. This finding can be obscured by tissue necrosis where only deep seated perivascular lymphocytic infiltrate can be viable and the macrophagic infiltrate can be focal and subject to sampling error. However we used “may” and not “can” to leave open the possibility of a pure lymphocytic infiltrate in exceptional cases which did not come to our extensive observation. The report of a distinct subtype gives the impression that it is common and we just underlined that if it is present, it must be very infrequent.
The authors state that the histological analysis of periprosthetic tissues can provide useful information for the longitudinal monitoring of implants but this has not been shown by any primary evidence in the study and mistakes the associations seen here with causation. This should be removed.

We agree that this part of the discussion was not clear and we revised it addressing the reviewer’s criticism. However, we disagree that we mistake associations with causation. We still addressed the issue of ALTR and corrosion in the title as association, even if the role of corrosion products as causative agent of the adverse reaction is currently well established by many multidisciplinary studies, although the precise physical and chemical properties of the particulate material of each class of implants and its possible association with protein component(s) has not been demonstrated yet with a complete chemical and physical analysis.

This part of the discussion has been revised and also made more concise in the section Discussion, paragraph Public Health Implications, page 24, line 550-569: “Our study suggests that the histological analysis of periprosthetic tissue in cases of ALTR can provide information that may be useful for longitudinal monitoring of implants. For example, we found that mixed lymphocytic and granulomatous subtypes were associated with shorter durations of implantation and were more common in the MoM LHTHA and Non-MoM DMNTHA with a known occurrence of taper corrosion [5, 7-9, 27, 61, 62]. In contrast, the predominantly macrophagic pattern is more common in the MoM HRA group which generates nano-size corrosion/conventional metallic debris particles only at bearing surface. The association between histological classification and time to revision may have clinical implications because implants with high number of patients with mixed macrophagic/lymphocytic pattern may fail earlier due to formation of pseudotumors with soft tissue necrosis, and this has resulted in implant recalls, such as the Stryker Rejuvenate and ABGII models. Implants with predominant macrophagic pattern, may fail at medium-long implantation time at an undetermined rate due to changes in the tribological lubrication process and/or macrophagic driven osteolysis. This unpredictable risk at the present time would call for a follow-up program with a frequency and modalities to be determined coupled with studies aiming at identifying biological and cellular factors associated with this type of adverse reaction [53, 63].”

The case has not convincingly been made in this study for the routine inclusion of histology reports in all registries. There are substantial logistic and cost implications in such a statement and therefore it should not be made until a clear health economic analysis has been performed to demonstrate a significant benefit to the inclusion of such data. The variability of such reports
from different observers and institutions is likely to be very large and render comparison of such reports of little value.

Thank you for this interesting comment. We revised the sentence on page 26, lines 571-581: “Our analysis showed that similar patterns of ALTR were present in implant classes of similar configuration and material composition independent of the manufacturer. This suggests the need for prompt observation and monitoring of any class of implants exhibiting a pattern of early failure with immediate reporting of sentinel cases to regulatory agencies/implant registries with the aim of avoiding high rates of complications for a large number of patients. Additionally, our results have made a case for the inclusion of the pathology report of revision cases in hospital based, regional, and national implant registries as an important and valuable tool in assessing modalities of implant failures along with the implementation of an international consensus classification, as the one recently reported for the periprosthetic soft tissue [64].”

Below are some of the reasons why believe that our histological analysis has demonstrated the utility of the inclusion of the pathology report in the registries:

1. The pathology report at least in the USA but also in some European countries is part of the patient’s standard of care and therefore would be available to the registries at no additional cost. Moreover, the cost of the histological examination has been calculated as less than 1% of the cost of the primary hip replacement or revision procedure and calculated at an average of 102.59 US dollars reimbursed for histological examination per primary arthroplasty in 1999 [Lawrence T et al. J Bone Joint Surg Am, 1999; 81 (7): 926 -31] with the cost of the hip prosthesis only ranging from 2392 to 12651 as calculated in 2012 [Robinson J et al. J Bone Joint Surg Am, 2012; 94 (18): 1693 -1698].

2. The pathological diagnosis is more accurate than the clinical diagnosis and provides clues of the natural history and biological response, as example the clinical diagnosis of aseptic loosening under which are grouped all patients with non-infectious loosenings, among them all cases of ALTR in conventional implants which the reviewer states that they are increasing only because of more consideration of the histological examination and all cases of osteolysis secondary to particle laden macrophagic infiltration of the bone. I would also like to emphasize that the first report alerting of the occurrence of ALVAL, a new type of adverse reaction, was from observation from pathologists with extensive experience in periprosthetic tissue retrieval.
3. The inclusion of the pathology report in registries with modalities to be agreed by a multidisciplinary expert panel could help to identify sentinel cases of failing implants as it has in the past for occupationally-linked tumors such as mesothelioma for asbestos exposure and angiosarcomas for vinyl chloride exposure and lead to increased scrutiny and surveillance of specific implants or classes of implants, potentially sparing thousands of patients worldwide from post-operative complications and allow substantial savings for national health care systems and mounting costs of litigation worldwide.

4. Variability from different observers and institutions could be minimized without substantial difficulty through education and standardization of diagnosis as it is done for neoplastic and non-neoplastic diseases with ad hoc panels and international classifications. One example is the recently reported international consensus classification which we mentioned at the end of the relevant paragraph [Krenn V et al. Pathol Res Pract. 2014; 210 (12):779-86)].