

Nonsensus in the Treatment of Proximal Humerus Fractures? An Uncontrolled, Blinded, Comparative Behavioural Analysis Between Homo Chirurgicus Accidentus and Macaca Sylvanus

Response to the editor and reviewers:

Dear BMJ,

Firstly, we would like to thank you very much for considering our submission. We thank the reviewers for their insightful comments and extremely thorough review of our manuscript, giving us the opportunity to improve. We have revised our manuscript and organized our responses in the table below for clarity. Furthermore, after a critical review of the manuscript, the following changes were made:

- various stylistic and punctuation mistakes have been corrected (page 3, line 101; page 4, line 131; page 8, line 249, lines 256-257, and lines 269; page 13, lines 417-418)

All changes in the manuscript have been marked with a word processing program. A copy of the manuscript with changes highlighted has been uploaded as a supplemental file with file designation 'Revised Manuscript Marked copy'.

We sincerely hope that the reviewers and editor find our revisions satisfactory.

I confirm that the revised manuscript has been read and approved by all authors and confirm that the work has not been submitted or published elsewhere in whole or in part. Each author confirms that the manuscript represents honest work.

If you have any questions at all, please do not hesitate to contact me.

Sincerely,

Sam Razaiean, M.D.

Corresponding author
Hannover Medical School
Trauma Department
Hannover, Germany

Editor and Reviewer Remarks	Authors' Responses	Text Changes
<p>Editor:</p> <p>1. Can you consider a less text heavy way to present the results?</p> <p>2. Please consider how to improve the presentation of the results including addition of the raw agreements for each case against the actual treatment and observed outcome.</p>	<p>First of all, we would like to thank you very much for considering our submission and giving us the opportunity to improve.</p> <p>Ad 1) We have recognized, that the results are presented in a very text heavy way. Therefore, we have revised this part, deleted some duplicated information and replaced the big text block about experts' qualifications by a clear table (Table 2).</p> <p>Ad 2) We deleted formerly table 6 and created table 7-9 including patient demographics as also required by reviewer #1 and raw agreements for each case against the actual treatment and observed outcome. Icons were created representing the two groups and their coloured digits in order to keep these raw data enjoyable for the reader, if desired. Furthermore, we would like to provide radiographs and CT scans of each case as required by reviewer #1 in the way they</p>	<p>Ad 1) Formerly lines 338-346 has been deleted and replaced by lines 346-347 and table 2.</p> <p>Duplicated information in lines 354-356 and lines 364-367 has been deleted.</p> <p>Line 359: A comprehensive sentence has been added.</p> <p>Line 369: A uninformative sentence has been deleted.</p> <p>Lines 373-375: A sentence referred to new tables 7-9 has been added.</p> <p>Ad 2) See tables 7-9 and supplementary files, if these are desired.</p>

<p>3. We felt that the methods could be much better explained. Further it isn't clear how the macaques could perform better than the surgeons</p>	<p>were presented in the web-based survey for the experts as well as printed as a poster for the macaques as supplementary file. The images could be linked with the tables, if desired. In addition, we added the radiographs after one year of nonoperative treatment for each case as supplementary file, if desired.</p> <p>Ad 3) We have realized, that there are some uncertainties due to an insufficiently explained method section. We have revised this section, provided more details, and edited the figure legend no. 2 in order to address these concerns. Furthermore, it seems that some unduly weighted, satirical suggestions gave the misleading impression that the macaques performed better than the surgeons. In lines 379-381 we have stated that the macaques' reliability is inferior and in some cases similar. Only regarding outcome prediction there was a trend that they predicted a bit more accurately. Nevertheless, we deemed it necessary to "tone down" some suggestions and conclusions as also required by reviewer #1 in</p>	<p>Ad 3) Page 8, lines 228-232: The case presentations included radiographs and a reconstructed 3D-CT image as well as patient demographics, information about secondary illnesses, and general health state before the injury given in the form of the 3-level version of the EuroQoL 5-dimensional instrument (EQ-5D-3L) (11) (Supplementary data).</p> <p>Pages 8 and 9, lines 256-257: [...] during the winter season in January 2020 under the exclusion of the general public in order to guarantee the anonymity of participating macaques.</p>
---	---	--

	<p>order to prevent false and misleading impressions (see ad 4+5).</p>	<p>Page 9, lines 263-271:</p> <p>Each kidney dish functioned as one of the aforementioned response options. An equally dosed mixture [...] and were placed into the kidney dishes. The first grasp into a kidney dish was defined as a treatment or outcome selection, and this behaviour was noted. With regard to question number two, any nonresponding among the macaques was defined as the response option “something else”. Apart from that [...].</p> <p>Figure legend 2:</p> <p>A two-pieced rating scale in analogous fashion for question number 1 and its two response options (nonoperative or operative) is not shown as it could not be secured in intact condition out of the macaques` hands and was lost to follow-up.</p>
--	--	--

<p>4. In some cases the satire might cause problems, such as the humorous suggestion that macaques should be considered to help with decision making--the dry humor might be interpreted poorly in some cultures</p>	<p>Ad 4) We agree, that the satirical conclusion that the macaques should be considered to be involved in decision-making could be misunderstood in some cultures as well as the hypothesis in our introduction that this specie could serve as a more worthwhile and reliable aid or the suggestion in our discussion that some findings would confirm this specie as a serious alternative. We have recognized that these statement could weaken our serious key message. Therefore, we have decided to delete these statements and to edit the conclusion.</p>	<p>Ad 4+5) Page 4, lines 135-139 and Page 15, lines 457-462 deleted and replaced by lines 135-139 and 452-456: Consensus on treatment and expected outcomes of PHFs is lacking even beyond the boundaries of the human species. Although Barbary macaques tend to predict the clinical outcome more accurately, their reliability to assist surgeons in making a consistent decision is limited. Future high-quality research is needed to guide surgeon decision-making on the optimal treatment of this common injury.</p>
<p>5. If we read this correctly and the point of this is illustrating a lack of data and evidence based information to assist surgeons in making a consistent decision, then it is worth it to consider how to drive home this point. In fact we found this humorous method of approaching this one of the more positive aspects of the study.</p>	<p>Ad 5) This is indeed true. The continuing controversy and lack of any consensus on the optimal treatment of this common injury was our original motivation to carry out this study. The satirical approach and method involving Barbary macaques should just highlight the key problem for the reader in an enjoyable way. We have seen, that it is necessary to emphasize more this key point in our conclusion as a take-home message than the unduly weighted satirical post-hoc</p>	<p>Page 5, lines 167-174 deleted and replaced by:</p> <ul style="list-style-type: none"> • Consensus on treatment and expected outcomes of PHFs is lacking even beyond the boundaries of the human species. • Future high-quality research is needed to guide surgeon decision-making on the optimal treatment of this common injury.

	<p>analysis suggestion that the apes should be considered in future decision-making process as also stated by reviewer #1. Therefore, we have revised the hypothesis, conclusion and summary box (section no. 2).</p>	<p>Page 7, lines 217-219 deleted and replaced by: [...] to determine the extent of consensus on treatment of this common injury.</p> <p>Page 13, formerly lines 383-385 deleted and replaced by lines 385-387: These findings highlight the continuing controversy and lack of expert consensus on the optimal treatment of these fractures even beyond the boundaries of the human species (4, 6, 16).</p>
<p>Reviewer #1:</p>	<p>Authors' Responses</p>	<p>Text Changes</p>
<p>I congratulate the authors on a thought-provoking and entertaining study.</p> <p>A few minor reporting issues which they may wish to address:</p> <p>1. A table with the demographics and radiographs of each of the nine cases would be helpful - I believe that many of us surgeons would wish to compare our own predictions with those of the apes, although with no great anticipation that we would fare any better than the experts involved in this study.</p>	<p>Thank you very much for your positive comment and your advice.</p> <p>Ad 1) Thank you very much for this hint. We deleted formerly table 6 and created table 7-9 including patient demographics as required. Furthermore, we provided radiographs and CT scans of each case in the way they were presented in the web-based survey for the experts as well as printed as a poster for the macaques as supplementary file.</p>	<p>Ad 1) See tables 7-9 and supplementary files, if these are desired.</p>

<p>2. The conclusions seem unduly weighted towards the sub-group analysis of older patients. This appears to have been a post-hoc analysis and could be construed as a bias in the reporting. The authors may wish to tone down these conclusions as they detract from an otherwise methodologically sound investigation.</p>	<p>The images could be linked with the tables, if desired. In addition, we added the radiographs after one year of nonoperative treatment for each case as supplementary file, if these are desired.</p> <p>Ad 2) It is true. It was indeed a post-hoc analysis and we did not expect this result before designing this study. We have recognized, that these unduly weighted post-hoc conclusions could detract from our key take-home message. Therefore, we have entitled these finding in our results section as post-hoc analysis findings, and we edited these suggestions and conclusions in our discussion and abstract.</p>	<p>Ad 2) Page 11, line 357: In a post-hoc subgroup analysis [...]</p> <p>Page 13, formerly lines 383-385 deleted and replaced by lines 385-387: These findings highlight the continuing controversy and lack of expert consensus on the optimal treatment of these fractures even beyond the boundaries of the human species (4, 6, 16).</p> <p>Page 15, lines 457-462 deleted and replaced by lines 452-456: Consensus on treatment and expected outcomes of PHFs is lacking even beyond the boundaries of the human species. Although Barbary macaques tend to predict the clinical outcome more accurately, their reliability to assist surgeons in making a consistent decision is limited.</p>
---	--	---

<p>3. The point about conflict of interest in the ape group is well-made, but perhaps a single image is sufficient to make this case.</p>	<p>Ad 3) Thank you for that hint. We have deleted the image series (formerly figure 4) and left only a single image (figure 3).</p>	<p>Future high-quality research is needed to guide surgeon decision-making on the optimal treatment of this common injury.</p>
<p>Reviewer #2:</p>	<p>Authors' Responses</p>	<p>Text Changes</p>
<p>I congratulate the authors on an engaging read and more importantly an irrelevant but important insight into the unjustifiable and diverse decision-making in the management of proximal humerus fractures. This is highlighted in the methodology of the two trials in this field - Profher 1 (completed) and Profher 2.</p> <p>The agreement tests are appropriate for handling of the data. Clearly the experimental protocol itself with the macaques is intended to be amusing and seems justifiable in this setting.</p> <p>Overall a cheerful spotlight on an importantly and rapidly growing area of clinical medicine where we as a community seem to be unclear about the way forwards.</p>	<p>Thank you very much for your benevolent comment.</p>	

Reviewer #3:	Authors' Responses	Text Changes
<p>The article is enjoyable to read and I think it may fit to the Christmas edition of the Journal. However, there are some issues especially in material and methods section, which needs addressing.</p> <p>1) The low number of group 2 (Barbary macaques) may lead to type II statistical error, which can be added to the limitation section.</p> <p>2) Secondly, they need to define the second group's behaviour clearer for example, what behaviour was considered as conservative management. There is not enough explanation regarding the behaviour of Barbary macaques to conclude on their response to the treats. A couple of references would be useful on behaviour of Barbary macaques.</p>	<p>Thank you very much for your positive comment and you hint.</p> <p>Ad 1) The statistical analysis includes Fleiss' kappa as a measure of agreement, no hypothesis testing was performed. Therefore, we think, that type II error is not likely to occur in this type of analysis. Nevertheless, we have seen the lower number of only five macaques compared to ten experts as a limitation when interpreting overall interrater agreement of the two species. This important aspect has been added for the reader.</p> <p>Ad 2) We have recognized that our methods are not explained sufficiently. This is why, there are some uncertainties, especially regarding the macaques' behaviour, the definition of a behaviour, and question no. 1 concerning recommended treatment (nonoperative vs. operative). In order to clarify these aspects, we have revised the methods section, provided</p>	<p>Ad 1) Page 14, lines 434-435: Furthermore, the lower number of only five macaques compared to ten experts should be considered as a limitation when interpreting overall interrater agreement of the two species.</p> <p>Ad 2) Page 9, lines 263-271: Each kidney dish functioned as one of the aforementioned response options. An equally dosed mixture [...] and were placed into the kidney dishes. The first grasp into a kidney dish was defined as a treatment or outcome selection, and this behaviour was noted.</p>

<p>3) Was there any strategy to avoid bias related to Barbary macaques?</p>	<p>more details, edited figure legend no. 2, and created table 7-9 including the raw selections of the macaques to conclude better on their responses.</p> <p>3) As this behavioural analysis was to be carried out on a voluntary basis by the macaques in their familiar enclosure under uncontrolled conditions, any attempt to prevent or minimise bias was omitted. However, we have added this important issue in an enjoyable way for the reader in our discussion.</p>	<p>With regard to question number two, any nonresponding among the macaques was defined as the response option “something else”. Apart from that [...].</p> <p>Figure legend 2: A two-pieced rating scale in analogous fashion for question number 1 and its two response options (nonoperative or operative) is not shown as it could not be secured in intact condition out of the macaques` hands and was lost to follow-up.</p> <p>Ad 3) Page 14, lines 420-426: As this behavioural analysis was to be carried out on a voluntary basis by the macaques in their familiar enclosure under uncontrolled conditions, any attempt to prevent or minimise this occurrence was omitted. The authors chose the winter season for this analysis in order to avoid general public access and to guarantee the anonymity of participating macaques; however, this choice may</p>
---	--	--

		have been poor, as the authors did not know that conflicts of interest among Barbary macaques are a seasonal affair beginning in November and lasting until March (20).
Reviewer #4:	Authors' Responses	Text Changes
This is a well written manuscript that approaches an actual subject of controversy (management of proximal humeral fractures) in a humorous fashion. It highlights the fact that there is still quite a lot of uncertainty in the prediction of treatment and outcomes. It appears to be well conducted, given the limitations of the study design. It gives pause for thought about the actual implications for clinical care and the need for further research in the subject. I think it would be a reasonable paper to consider for the Christmas BMJ.	Thank you very much for the positive comment.	