Reviewer’s report

Title: "Principal Component Analysis of the Relationship Between Pelvic Inclination and Lumbar Lordosis"

Version: 0 Date: 24 May 2018

Reviewer: Tomaz Vrtovec

Reviewer's report:

The authors present a study in which they measured the geometrical parameters of the spine and pelvis via surface tomography, and then performed principal component analysis (PCA) on the obtained data to observe the dependency among the measured parameters as well as trends in the parameters that contribute to their variability. Their results show that there is no relationship between pelvic inclination and lumbar lordosis. Nevertheless, the authors need to elaborate more their approach, especially from the point of view of data acquisition (accuracy, reliability and repeatability of surface topography measurements) as well as principal component analysis (from the perspective of the stopping rule that defines the degree of the variance in the retained principal components). Detailed comments are provided below.

Comment #1: The title of the manuscript is in my opinion too general, as the conclusions to which the authors came need to be further validated, which was acknowledged also by the authors. I would propose that they somehow include in the title the fact that this findings are "based on surface topography measurements" and "principal component analysis".

Comment #2: Abstract appears to be too long, especially the Background and Methods sections. The Background part is somehow a speculation, except for the last sentence, which describes the aim of the study. The abbreviation PI is commonly used for pelvic incidence, and not for pelvic inclination. Expressions like "Golden Calf" are not adequate for scientific publications. In the Methods part, the expression "to measure the spine and pelvis" is ambiguous - please be more precise, as the spine and pelvis are not being measured, but specific morphological parameters are obtained through surface topography measurements. The exact model and manufacturer of the scanner is usually not reported in the abstract. Please state your hypothesis clearly in the Background part, and refer to it clearly in the Conclusion part. Please use abbreviations consistently once defined (e.g. LL was defined, but "lumbar lordosis" still appears in the Results and Conclusions part).

Comment #3: In the first paragraph of section Introduction, the authors state that the morphological parameter named "pelvic inclination" was defined by Legaye et al. (1998, 2011). However, no such term can be found in the cited publications. The authors should be more precise and follow the established terms: do they mean pelvic tilt (or tilting) - PT, pelvic incidence - PI, or a different one?
Comment #4: The whole data analysis that leads to the obtained conclusions is based on the accuracy of the surface topography scanner in performing the described measurements of pelvic and spinal parameters. Could the authors perform experiments to assess the reproducibility, reliability and accuracy? This needs to be presented, as for example the variability in reproducibility may show or clear the correlation among some of the observed parameters, which was not or was, respectively, observed by the authors. On the other hand, from the cited work of Knott et al. (2016), it can be observed that while there is a strong correlation between lumbar lordosis measured by surface topography and X-rays (Pearson correlation of 0.867), the average difference in measurements was ±9.7 degrees, which is quite large. On the other hand, weak correlation was observed for pelvic obliquity (0.223) with a relatively low average difference of ±2.3 degrees. How do these results influence the perspective and conclusions of the present study?

Comment #5: The resolution of all figures is too low to observe the depicted elements. I do not know if this is the authors' responsibility or the submission system, but it appears that the resolution of figures is 72 dpi, which is ridiculous. For a proper assessment, they should be of 600 dpi (or at least of 300 dpi). Moreover, in Figure 1 the morphological parameters have to be marked in the figure as well, and not only reported in its caption.

Comment #6: The subfigures in Figure 2 should be presented in a more proper way. If the data are binned, the abscissa should show the values that occurred for each measured parameter, and not the label "Population" (this label actually wrongly refers to the actual population of observed subjects).

Comment #7: The authors mention that they conducted the Grubb’s test for outliers, however, no results are reported. How many outliers were there, why did the outliers occur, etc.? Or it reported in section Results - Distribution of the data, where they report that "no statistically significant outliers“? If so, then they have to report in more in detail by referring to the above mentioned test.

Comment #8: In the caption of Table 1, the authors report "107 normal subjects". It was however understood that the subjects underwent treatment because of pelvic misalignment. Is it still correct to label such subjects as "normal"? Moreover, the values in Table 1 are reported without units - please correct.

Comment #9: The reference to the Supplemental Figure 1 occurs so many times that it is weird that the authors did not incorporate it into a regular figure. For example, each subfigure of Figure 3 could be divided into top and bottom section, one representing the current Figure 3 and the other Supplemental Figure 1.
Comment #10: The authors have to provide the characteristics of the studied population in terms of gender and age. Also, the distribution according to gender and analysis of statistical differences between genders should be provided, as they may bias the resulting findings.

Comment #11: Why was the D'Agostino-Pearson test used for testing the normal distribution of data, and not some other test, for example, the Shapiro-Wilk test, which is more established in such studies? Please elaborate.

Comment #12: In section Results - Principal component analysis, the authors mention the "null hypothesis", which should be defined already in the Introduction.

Comment #13: In the same section as above, they describe the isolation of principal components PC1 and PC2 as the two components that contribute the most to the variability of the data. However, when analyzing data with PCA, each principal component can be described by how it affects the data. The authors should provide such descriptions as well, meaning how PC1 and PC2 affect the data, since both account for almost 50% of the variability.

Comment #14: In relation to the previous comment, the authors chose to retain just the first two principal components (PC), which represent 43.5% and 47.2% of the total variability in the data before and after treatment. However, the retained variability is relatively low, as it does not even reach 50%. This is basically a problem of the "stopping rule" that questions how many PCs should be retained so that they are still representative for the underlying data. While a rule of thumb is at least 80% or even 90%, the others retain those with corresponding eigenvalues larger than 1. The authors should also study "Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches" by D.A. Jackson (Ecology, 74:2204-2214, 1993) and "How many principal components? stopping rules for determining the number of non-trivial axes revisited" by Peres-Neto et al. (Computational Statistics & Data Analysis, 49:974-997, 2005) for a deeper insight of this problem. Nevertheless, retaining below 50% of variance does not seem convincing, and the authors should at least prove that the conclusions and findings they report remain the same even when they include more PCs (and therefore retain a larger variance) into the analysis.

Comment #15: In relation to the previous comment, it is assumed that the authors used linear PCA, therefore inferring a linear relationship among the studied variables. Is this assumption correct? Please elaborate.

Comment #16: While PCA analysis is an interesting approach, the authors should also report the Pearson correlation among the measured variables, as such statistics is established in similar studies and would also allow to compare their findings to other studies.
Comment #17: As noted in a previous comment, the quality (resolution) of Figure 3 has to be improved.

Comment #18: The authors correctly conclude that this topic needs more discussion and investigation, and more properly devised studies in the future. As a result, I would suggest that the authors address the community with a milder tone, especially in the first paragraph of section Discussion and in the Background part of the Abstract.

Comment #19: On several occasions, the authors use expressions that are too colloquial and not appropriate for a scientific publication. Also, please use abbreviations consistently after they are defined, e.g. PCA, ST, etc. On the other hand, some abbreviations are not defined, e.g. AIS. Some parts of the manuscript should be revised (e.g. the first sentence in section Results - Population).

**Level of interest**
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

**Quality of written English**
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal