Author’s response to reviews

Title: With axial loading during MRI diurnal T2-value changes in lumbar discs are neglectable: a cross sectional study

Authors:

Leif Torén (leif.toren@vgregion.se)
Hanna Hebelka (hanna.hebelka@vgregion.se)
Izabela Kasperska (izabela.kasperska@vgregion.se)
Helena Brisby (helena.brisby@vgregion.se)
Kerstin Lagerstrand (kerstin.lagerstrand@vgregion.se)

Version: 2 Date: 21 Dec 2017

Author’s response to reviews:

Dear Editor and reviewers;

Thank you for your valuable and constructive comments and suggestions to improve our manuscript entitled “With axial loading during MRI diurnal T2 changes in lumbar discs are unneglectable: a cross sectional study”. We have responded to your requests and questions. See enclosed below the reviewer comments (in black) and our subsequent answers/clarifications (in blue). In the manuscript, alterations/clarifications have been highlighted in grey color.

(Please see this response also as attached file, in which colorsetting is possible that will make it easier to follow raised questions and corresponding answers)

Reviewer reports:

Alessandra Splendiani (Reviewer 1): None.

Tue Secher Jensen (Reviewer 2): The authors present results from a study aiming to investigate the diurnal variation of T2-values in lumbar discs in six asymptomatic volunteers when doing unloaded and loaded supine MRI.

This is a scientifically sound research question and, as the authors state, may be of importance for the planning of clinical and research MRI.
The manuscript is well-written and concise. However, there are a number of significant methodological/study design and interpretive limitations that should be considered prior to publication as outlined below:

Major issues:

1. There is a low number of participants (discs) included in this study and therefore a risk of underpowered analyses and type 2 error. Visually (Figs 2 & 3) there is a decrease in T2 values over time, especially in the central disc (ROI3). However, it is likely that the reason for not reaching statistically significance is due to the small sample size.

Please provide sample size calculations for the study to ensure that the study is sufficiently powered for the required analyses.

Answer: Thank you for valuable comments. We agree that the sample size is small in regard to the measured effect size and, hence, the study most probably is underpowered for determination of statistical significance. When we designed the study, the sample size was chosen ‘big enough’ to statistically verify significance for an expected effect size of clinical significance. Since there is a lack of similar studies, in which quantitative T2-mapping measurements are performed with axial loading during MRI, estimation of the sample size was based on reported diurnal IVD changes in the T2-value over the day for unloaded MRI, i.e. with a regional diurnal IVD change reported to range from 5 to 16%. A diurnal effect of 10% was considered clinical relevant, as such a sample size of 5 individuals with 25 discs should be sufficient at 80% power with 95% significance and two sided test. One extra healthy volunteer was included in the study to ensure power. With axial loading during MRI (alMRI), however, the effect size was found to be much smaller and, hence, no statistical significance could be verified. Requested reference has also been added.

Clarification has been added in the manuscript, please see limitation section.

2. As I read the manuscript, the number of discs included in the pair-wise analyses in the analysis for reproducibility is five (one volunteer). This seems to be a very low number for an analysis of reproducibility. For a study of this kind, it is important that the precision of the method used is available to the reader.

Please provide a sample size calculation [Watson, 2010] and redo the analysis if necessary. It may be sufficient to do use the existing material (six volunteers = 30 discs) for this purpose. Also, please include 95% confidence intervals for the ICCs.

Answer: Thank you for valuable comments. We agree with you and have added ICC and 95% CI for measurements performed on 36 IVDs in 5 LBP patients. This was actually performed before initiating this study, in order to evaluate the reliability of the measures. We have now added the data requested to the method section, please see Table 1 and corresponding text highlighted in grey.

Minor issues:

1. Please add reference for this sentence: "Quantitative T2-values are known to correlate to hydration grade with an inverse correlation to degeneration grade of the IVD.", Page 3, para 2, line 7-8. Reference #13 may be useful.

Answer: Accordingly, 2 relevant references have been added. Please see manuscript.

2. Results, Table 1: Please provide means and SD for all ROIs/timepoints in Table 1.

Answer: Accordingly data has been added upon request. Since Table 1 has been added, former Table 1 is now Table 2, please see Table 2 with added means and SD at all time instances for all ROIs

3. Results, Figure 2-4: Adding error bars (SD) for each data point would increase the reader's ability to assess the variability in the dataset visually. This is probably not possible for Fig. 4 due to the number of data points.

Answer: Error bars (SD) for each data point has been added upon request, however only to Figure 2 & 3 since corresponding error bars in Figure 4 did not increase reader’s ability to assess the variability visually due to too many data points. Please see revised Figure 2 & 3

4. Discussion, p. 8, para 2: The divergent results are most likely not due to the small sample size and under-powered analysis.

Answer: These divergent results could partly be caused by small sample size, however these divergent results are more likely explained by a higher mean age (38 years) and increased degeneration grade in the current cohort. Clarification has been added, please see discussion section

5. Discussion, Limitations: Could the use of both unloaded an axial loading in the same day have influenced the results? Would the authors expect different results if loaded and unloaded scans were performed on different days?

Answer: Yes, hypothetically this could have an impact. However if so, the effect would have been greater over the day regarding the unloaded MRI measurements and in such case we would have a greater decline at the first measurement and not so great at the second measurement. In our results there is no indication that alMRI accelerates this effect. In addition the load the alMRI exerts is equivalent to the load in upright position why the effect of alMRI likely not is greater than if the individual stands up/walk, i.e. simulating a clinical situation.