Author’s response to reviews

Title: Are two readers more reliable than one? A study of upper neck ligament scoring on magnetic resonance images

Authors:

Ansgar Espeland (ansgar.espeland@helse-bergen.no)
Nils Vetti (nils.vetti@helse-bergen.no)
Jostein Kråkenes (jostein.kraakenes@helse-bergen.no)

Version: 3 Date: 3 December 2012

Author’s response to reviews: see over
Regarding the manuscript “Are two readers more reliable than one? A study of upper neck ligament grading on magnetic resonance images”, manuscript number 1115251142624900

The comments from the reviewers were very useful. Below please find a point by point response to their suggestions for revision of the manuscript.

Changes are marked yellow in the manuscript; however, should the manuscript be accepted, we hope that these yellow marks will be removed prior to publication of the provisional PDF.

Reviewer: Pedro Machado

Reviewer’s comment: Sum scores are often used in MRI scoring systems to assess status and change in clinical trials and observational studies. As an additional methodological exercise I suggest complementing the analyses with the calculation of an upper neck MRI sum score (0-12) and comparison of the sum scores (reader A vs. reader B vs. A and B combined vs. average of A and B) regarding:

a) Intra-reader reliability using intraclass correlation coefficients (ICCs)
b) Smallest detectable change and Bland and Altman plots with 95% levels of agreement

Response: We have performed all of these analyses, as now reported in Methods (page 6, paragraph 3), Results (page 8, paragraph 2) and Discussion (page 8, paragraph 2 and page 11, paragraph 2). Bland and Altman plots are found in Figure 2 and the corresponding figure legend has been added on page 18 and 19. A new reference has also been added (current reference 18). We have adjusted the Abstract slightly to incorporate some of the additional analyses without exceeding the word limit.

Reviewer: Ulrich Weber

Reviewer’s comment: I have just one suggestion for the section "discretionary revisions". This study as a first step towards a "2 readers approach for MRI studies" has no clinical "anchor" or gold standard (page 9 2nd para: no difference between groups; no predictive utility) and resolves discrepancies by consensus. Major unresolved clinical questions for imaging study designs are whether a "2 readers approach" has better agreement with a clinical gold standard than a readout with one single reader, whether a multiple reader approach (>2 readers) with an analysis by the majority of readers may outperform a "2 readers approach" regarding a clinical anchor and whether an assessment of discrepant reads per se may provide clinically more meaningful information than consensus reads in case of discrepancies.

Response: We have incorporated these important unresolved clinical questions for imaging study designs into the Discussion (page 10, last paragraph and page 11, first paragraph).

We hope the revised manuscript can be accepted for publication.

Best regards

Ansgar Espeland