Reviewer's report

**Title:** Quantification of Cervical Spine Muscle Fat: A Comparison between T1-weighted and Multi-Echo Gradient Echo Imaging using a Variable Projection Algorithm (VARPRO)

**Version:** 1  **Date:** 29 January 2013

**Reviewer:** Arne Fischmann

**Reviewer's report:**

The authors compare two methods to quantify muscular fatty infiltration in healthy patients in order to evaluate whether an MRI sequence with shorter scanning time (3D GRE) can replace one standard technique using a T1-weighted sequence. Quantitative MRI of muscle fat infiltration is increasingly used in patients with neuromuscular, traumatologic and orthopaedic disorders. However up to now, no standardized measurement are established and comparison of different protocols are an important step in this process.

While the questions asked in this paper are important, there are still some shortcomings, especially the small sample size as well as the limited number of reviewers. As the authors state in their conclusion, further studies including histologic validation are undertaken in the near future, these shortcomings are acceptable and should not prevent publication.

Several specific points might be addressed before publication:

1. **Major Compulsory Revisions**

1.1 From table 4 it is apparent that MFI using 3d GRE is about 2% lower than T1w on all levels except at the level of C7, where it is 7% higher. Although most of these changes are not significant, the obvious differences between these levels should be addressed. Especially as for C7 there is a significant inter-reader difference as visible on the bland altman-plots (figure 2e). Might it be possible, that one reader included fatty tissue adjacent to the muscles at this level? e.g. due to inclusion of trapezius etc.? Another possibility might be the inclusion of muscles with larger amounts of intramuscular fat (see also comment 3.4)

1.2 The sample included only 5 patients, therefore the Kolmogorov-Smirnov test should not be used as it has only limited power to reject the null-hypothesis (of the normal distribution) in a small sample size Either a test that could be used for smaller samples (e.g. the Shapiro-Wilk test) should be used to test for normality, or non-parametric statistical methods should be used to evaluate the significance of differences.

See also: http://instatmy.org.my/downloads/e-jurnal%202/3.pdf
2. Minor Essential Revisions

2.1 On Page 3 line 88: 3D GRE is introduced without explanation, leaving the casual reader confused.

2.2 Line 214-216 should be moved to the discussion.

2.3 Line 226-228 is questionable. The size of each group (experienced vs. non-experienced readers) is only 1 and therefore differences can easily be attributed to different strategies to include marginal pixels. This cannot be generalized to a statement on the influence of experience, which would require a larger number of readers (preferably > 5 in each group)

2.4 It is not clearly stated (but implied), that figure 1 and 2 are based on data acquired with the 3D GRE sequence.

2.5 If data in table 2 are means of both readers, this should be explained in the table description.

2.6 On page 7, line 180 is stated that 2 readers established repeatability. However, there is no mentioning of any repeated measurements or evaluation. I suspect that inter-reader differences were assessed, and this should be corrected in this section.

3. Discretionary Revisions

3.1 This study is based on a limited number of participants, comprising only 5 volunteers, therefore the results might be skewed and no normative data can be derived from this trial. This might be discussed and expanded upon.

3.2 Table 1 consists mainly of repeat data from the methods section (study population) and might be integrated in the text of the results section.

3.3 As most measurements included not only intramuscular fat, but also fat between different fibres of the muscles, it would be helpful to have exemplary ROIs by both readers- e.g. at the levels C7 and C6 (as these are levels with significant differences between the readers.

3.4 The T1 sequence only measures fat fraction using the number of pixels with fat-equivalent signal intensity. As most muscles also have minimal amounts of fat within pixels with normal appearance on T1 weighted images, this fat is usually lost to the standard T1 weighted sequence. In healthy volunteers, the amount of fat in the legs is usually about 2%, which might explain the small difference between T1w images and 3D GRE images found at most levels (see Table 2)

3.5 The authors use both terms: muscle fat infiltrates (MFI) and muscle fat percentage (MFP) in the text interchangeably. As this is confusing, they should preferably stick to one abbreviation, especially as MFP is not properly introduced but only explained in the abbreviations section.

**Level of interest:** An article whose findings are important to those with closely related research interests
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests