

Reviewer 1:

The readability of this paper should be further improved. Abstract: the sentence 'to retrospectively and prospectively compare diffusion-weighted (DW) images in the abdomen in 1.5T and 3.0T systems' is confusing.

We agree. The sentence has been modified.

Abstract: Results: Subjectively, both dpTX 3.0T and 1.5T systems provided higher image quality... (higher than?)

We agree. Changes have been made.

All MR scanners were equipped with the same gradient systems' is this true?

Yes, this is true.

It will be interesting to include ADC value analysis.

We have evaluated ADC values in another publication (Rao et al., PLoS One, 2012)

Please provide information on inter-reader agreement (senior radiologist and junior radiologist)

We have now described the inter-reader agreement in the Results: Subjective Assessment section.

Discussion can be further expanded regarding the technical explanation which are still too simplistic currently.

We agree. Additions have been made with regards to technical explanation.

Reviewer 2:

Several points need to be further addressed in order to be published. - Although the content of the manuscript is comprehensive, it seems not focused and sometimes it is difficult to follow, such as in Results and Table 3. - Is it really necessary to include eight anatomical distributions in SNR measurements?

We would like to maintain these 8 anatomical distributions as they represent regions chosen due to clinical significance in clinical practice. Nevertheless if you still prefer to omit some of them and this is a premise for a positive vote, than we could also leave them out.

It would be better to reduce some content to make the manuscript more concise and clear. - The authors measured signal intensities across all b values. But the clinical relevance of doing this or the significance of measuring all b values are not mentioned.

In our introduction, we have mentioned, “Our aim was to make a comparative analysis of images produced by the three scanners in various abdominal regions across a set of b-values commonly used clinically.” As above, we feel the values chosen are important for clinical practice, but if this would result in a positive vote, we can limit our paper to only two selected values.

In addition, why ADC values were not included in the measurement?

We have evaluated ADC values in another publication (Rao et al., PLoS One, 2012)

Minor comments Abstract: Retrospective and prospective study populations should be directly defined as 150 patients and 10 volunteers respectively.

We agree. Changes have been made as recommended.

Image analysis: It is said that “This process was repeated with all 150 of the hospital patients, and in all 30 of the studies from the ten volunteers”. What do you mean here “in all 30 of the studies from the ten volunteers”?

We apologize for this confusing sentence. We have modified the sentence.

Results: what is the statistical significance comparing the results between two readers?

We have now described the inter-reader agreement in the Results: Subjective Assessment section.

Figure 1: Fig 1a and 1b are not marked.

We have corrected this.