We want to thank the reviewers for taking the time to review our manuscript. We appreciate all the helpful comments.

Below is an itemized list of all the reviewer comments. We have made changes to the manuscript as described below.

Thank you

Janio Szklaruk, PhD MD

Professor Diagnostic Radiology

University of Texas at MD Anderson Cancer Center

Houston TX

77025.

I have the following specific comments:

1) It is not clear what exactly caused the indicated artifact in the FB image in Fig 1a. Were similar artifacts observed in other patients?
   a. The artifact in Figure 1a refers to the relative signal loss on the liver. It appears darker than the rest of the liver.
   b. The following phrase was added.
      i. *, seen as signal loss on the left liver compared to the right liver,*

2) How frequent did artifacts appear in the FB sequence compared to the other 2 sequences?
   a. The overall score of image quality that included artifacts ranged from 1-5. With 5=best image quality.
   b. The image quality included artifacts but also included other factor such as signal to noise and blurriness. This was a qualitative score.
   c. We did not quantify the number of artifacts.

3) Did all the sequences use the same readout bandwidth (that is related to magnetic susceptibility artifacts)?
   a. Yes, three sequences used the same readout bandwidth (2.44 kHz/pixel) for 20 patients. This information was added to the method:
   b. The following sentence was added
      i. *"The readout bandwidth was 2.44 kHz/pixel for 3 DWI sequences*

4) In the beginning of the Discussion, the paper attributes the poorer image quality and more artifacts in the FB sequence to longer scan time and lack of respiration triggering. However, other factors might also be at play. For example, the longer TE (67ms) used in the FB sequence than in the SMS and PACE sequences (56ms) would have led to lower signal.
   a. The following text was added:

> *i. We also note that TE for SMS-DWI and PACE-DWI is 11ms shorter than that of FB-DWI. The decrease in TE was enabled by the higher overall acceleration of SMS than parallel imaging alone and was likely helpful in improving the image SNR.*

5) The paper does not explain why longer TE was used in the FB sequence.
   a. For clarity the following text was added,
      > *i. The decrease in TE was enabled by the higher overall acceleration of SMS than parallel imaging alone and was likely helpful in improving the image SNR.*

6) Furthermore, later in the Discussion, it is mentioned that "FB-DWI showed fewer artifacts", which seems to contradict what was said earlier.
   a. For clarity, this phrase was removed
      > *i. FB-DWI showed fewer artifacts*

7) On page 9, it is stated "The histogram with the most usable pixels was considered the superior one". How was "usable" defined and determined?
   a. The histogram is a representation of the color coded pixels of the region of interest in the MR image. The red pixels correspond the red bars on the histogram. The histogram is generated automatically by translating the ROI to pixel by pixel 2D graph. The x- and y- coordinates are

defined as x- (ADC value) and y – (number of pixels with that value). The frequency y-axis is generated automatically depending on the number of pixels with that ADC value. For the example on Figure 2, the ADC histogram was considered superior to PACE DWI due to the higher number of pixels for the whole image. The ROI was the same for all 3 ROI.

b. The following was added for clarification.

  i. *In general, the histogram with more pixels per ADC value was considered superior quality*.

8) How were the upper and lower limits of the frequency scale in Fig. 2 determined?

a. The frequency y-axis is generated automatically depending on the number of pixels with that ADC value.

9) Please provide more description for the PACE DWI technique and explain why it did not significantly increase the scan time over the FB DWI sequence (4min 58s vs 4min 44s) in this study.

a. We added the following text to the manuscript.

  i. *"In particular, prospective acquisition correction (PACE) is an internal navigator-based technique for respiratory signal monitoring. The navigator is typically placed at the diaphragm, then the series of measured diaphragm position is passed to a*

respiratory trigger algorithm to acquire images only during the end of expiration phases. The PACE DWI allows the synchronization of the DWI acquisition with respiratory cycles without the need to place external monitoring devices on the patient [18]."

ii. *The average scan times for FB-DWI, SMS-DWI, and PACE-DWI were 4 minutes 56 seconds (4 minutes 44 seconds ~ 6 minutes), 3 minutes 8 seconds (3 minutes 4 seconds ~ 3 minutes 38 seconds), and 5 minutes 40 seconds (3 minutes 40 seconds ~ 12 minutes)*

9. The paper claims that "an advantage of our study is that we were able to compare ADC values and evaluate the precision of the ADC calculations". However, given that "all lesions were detected in all sequences" as stated in the Discussion, it is not clear whether there are any significant benefits of the "better image quality and less ADC variability" to the diagnosis of liver lesions.

   a. We agree that in terms of lesion detection, in our study population, there was no clear benefit.
   b. However, the decreased in overall image quality suggest an improvement over free-breathing DWI.

c. In addition, changes in ADC histograms are currently used to assess tumor treatment response.
d. We proposed that the improvement in image quality and improved histogram makes the PACE and SMS superior to FB-DWI.

10. As mentioned in the paper, a limitation of this study may be that most lesions were malignant and therefore the usefulness of ADC values for lesion characterization could not be determined. Furthermore, without a gold standard for comparison, how could one be sure that the larger ADC variation in the FB sequence was entirely due to poorer image quality and not reflecting (at least partially) the real condition of the liver lesions?
    a. We agree that the variation in ADC values can be due to heterogeneity of the liver lesions.
    b. We added the following phrase,
        i. *The variation in ADC values may also be due to the inherent heterogeneity of the liver lesions.*

11. Minor comments: 1) In the caption of Fig. 2, please change "The bottom image …" to "The diagram on the right …"
    a. This was changed.

12. Title: "Comparison of three different DWI sequences for liver imaging quality" seems better than the original one.
    a. Title was modified
        i. Comparison of free breathing and respiratory triggered DWI sequences for liver imaging

13. Abstract: Aim: "To qualitative and quantitatively compare "should be "To qualitatively and------".
    a. The following was modified.
        i. *To qualitatively and quantitatively compare*

14. Methods: Here, the authors used a lot of abbreviations without giving the full phrase like FB-DWI, SMS-DWI and PACE-DWI.
    a. In the introduction we defined the terms
        i. *Simultaneous multi-slice (SMS)*
        ii. *Prospective acquisition correction (PACE)*
        iii. *Free-breathing*

15. When using an abbreviation, the full phrase should be given at the first time of use. Later, you can always use the abbreviation without mentioning the full phrase. This is the rule. Please check the whole article and give the full phrase of abbreviations at the first time of use.
    a. *We reviewed and revised the abbreviations.*

16. Results: "The mean qualitative image quality score of PACE-DWI (4.48)" should be "The mean qualitative image quality score of PACE-DWI was 4.48."
    a. The following was revised.
        i. *The mean qualitative image quality score of PACE-DWI (4.48)*

17. Core tip: here, "we compared three diffusion weighted for liver imaging" should be "we compared three diffusion weighted sequences for liver imaging".
    a. The following was revised
        i. *We compared 3 DWI techniques for liver imaging:*

18. Introduction: Here, the authors mentioned "Respiratory-triggered acquisition---". Does Respiratory-triggered acquisition need breath-hold technique?
    a. The following was added
        i. *Respiratory triggering, that are not breath hood techniques, increases the total scan time by as much as 3-fold,*

19. The authors mentioned free breathing but without mentioning breath-holding technique.
    a. Breath holds were not employed in any of the three DWI techniques.
    b. The following sentence was added,

i. *Breath holds were not employed in any of the three DWI techniques.*

20. Do the three sequences the authors used for comparison here integrate breath-hold technique?
    a. Breath holds were not employed in any of the three DWI techniques.
    b. The following sentence was added,
        i. *Breath holds were not employed in any of the three DWI techniques.*

21. What is the difference between the three comparisons and the breath-holding technique? Please specify this and also discuss it in the discussion section.
    a. Breath holds were not employed in any of the three DWI techniques.

22. 5. Qualitative analysis: In this section, there is a grammar mistake in the sentence: " Two readers, independently, qualitatively compared the ADC histograms from all 3 DWI sequences side-by-side on the basis of the ADC histogram distribution reflected the tumor heterogeneity ". Here, reflected should be reflecting.
    a. The text was modified
        i. *ADC histograms from all 3 DWI sequences side-by-side on the basis of the ADC histogram distribution that reflected the tumor*

23. DISCUSSION: Here, you mentioned breath-hold DWI, and you should compare this with the three DWI investigated in your article.
    a. For clarity, we removed BH-DWI from the manuscript.
    b. BH-DWI is not part of our project and it was never part of our standard of care imaging.

24. In Table 2, please add a row with the name of total and add all the number in this row.
    a. This was added.

25. 8. Table 4: What do you mean by "Sequence A" and "Sequence B"? Are they related to the MRI scanning technique? This table seems useless and belongs only to the statistical presentation, which does not show many useful information. If possible, just describe the outcomes in the text.
    a. Table 4 was removed.