AUTHORS RESPONSE TO REVIEWER

Reviewer #1:

NAFLD is a chronic disease worldwide, and the grading and quantification of liver fat is dependent on histological examination. Due to the invasive attributes and potential risks, liver biopsy is not easy to for broad generalization, and the development of credible imaging techniques is crucial. The authors used a retrospective single center cross-sectional study to compare the diagnostic performance of two point Dixon and Six point Dixon MR in detecting liver steatosis, including quantification and grading. The study was approved by the institutional ethics and review committee, and the requirement for informed patient consent was waived. Considering the quality control of MR and a detailed data analysis plan, it is evident that Six-point Dixon MR has unparalleled advantages. However, the following issues need to be considered: 1. The number of enrolled patients in the study is relatively small, although statistical differences were obtained, it cannot represent the clinical significance. 2. It is a single center study, and there are issues with the representativeness of the study. Lack of necessary histological control studies. 3. Comparison between diffuse and focal lesions is needed; 4. The parameters of instruments from different providers should be considered.

Firstly, our authors would like to thank the reviewer for kindly reviewing our manuscript. We are very appreciative of the reviewer's comments and encouraged that the reviewer has highlighted the detailed data analysis plan that we have conducted and the evident advantages of MRI including six-point Dixon. The response to each of the reviewer's points is included below.

1. *''The number of enrolled patients in the study is relatively small, although statistical differences were obtained, it cannot represent the clinical significance''*

While our study population of 62 patients is modest, this is still a reasonable number of patients for a clinical study, and it both meets and exceeds the minimum number (52) required as per the a-prior power analysis that we performed.

We found not just statistical significance but **very strong and compelling** statistical significance (**P < 0.001**) with ROC analysis showing an excellent accuracy for detecting steatosis on both two-point Dixon [AUROC = 0.96, P < 0.001] and six-point Dixon [AUROC = 0.95, P < 0.001].

As such, the study findings are of **potential** clinical interest, and there is a need for the findings to be confirmed or refuted for clinical signficance in larger prospective multi institution studies. We respectfully disagree with the reviewer, to categorically state that the findings cannot represent clinical significance without the opportunity for further testing in larger studies is premature, and does not conform to established scientific principles including due process and the testing-retesting cycle.

2. *''It is a single center study, and there are issues with the representativeness of the study. Lack of necessary histological control studies''*

We acknowledge the limitation that this a single center study with issues of representativeness. This limitation is included in the Discussion section of the manuscript, and we have stated that our findings need to be evaluated by future multi-institution studies to determine their generalizability. **The following appear in the manuscript text:**

''This was a single-center retrospective analysis performed on a relatively modest number of patients. Notwithstanding, the study was adequately powered to detect statistically significant differences. Future multi-institution studies are necessary to determine the generalizability of our findings.''

AND

''a similar study in a large patient population with various grades of hepatic steatosis would be helpful to establish the clinical translatability of the findings. ''

2. *''It is a single center study, and there are issues with the representativeness of the study. Lack of necessary histological control studies''*

Regarding the lack of histology, we would argue that this is not essential. The gold standard used in our study was magnetic resonance spectroscopy (MRS). It is widely regarded as the non-invasive reference standard for liver fat quantification, is validated and strongly correlated with histopathology, and has been touted in the medical literature as a replacement for biopsy itself (1-3).

Accordingly, we like to point out that histology by liver biopsy is not without its own limitations. It is invasive, and associated with complications such as pain, bleeding and even death. Biopsy only samples 1:50000th of the liver volume, far less than MRI (4). As steatosis often affects the

liver non-uniformly, biopsy can itself be poorly representative. Finally, the histopathology assessment is a qualitative process prone to intra- and inter-reader variability (5).

References

1. Roldan-Valadez E, Favila R, Martínez-López M, Uribe M, Ríos C, Méndez-Sánchez N. In vivo 3T spectroscopic quantification of liver fat content in nonalcoholic fatty liver disease: Correlation with biochemical method and morphometry. Journal of hepatology. 2010;53(4):732-737. [PMID: 20594607 DOI: 10.1016/j.jhep.2010.04.018]
2. van Werven JR, Schreuder TC, Aarts EO, et al. Hepatic steatosis in morbidly obese patients undergoing gastric bypass surgery: assessment with open-system 1H-MR spectroscopy. AJR American journal of roentgenology. 2011;196(6):W736-742. [PMID: 21606262 DOI: 10.2214/AJR.10.5215]
3. McPherson S, Jonsson JR, Cowin GJ, et al. Magnetic resonance imaging and spectroscopy accurately estimate the severity of steatosis provided the stage of fibrosis is considered. Journal of hepatology. 2009;51(2):389-397. [PMID: 19505740 DOI: 10.1016/j.jhep.2009.04.012]
4. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128(7):1898-1906. [PMID: 15940625 DOI: 10.1053/j.gastro.2005.03.084
5. Bravo AA, Sheth SG, Chopra S. Liver biopsy. The New England journal of medicine. 2001;344(7):495-500. [PMID: 11172192 DOI: 10.1056/NEJM200102153440706]

3. ''Comparison between diffuse and focal lesions is needed''

The purpose of the study as stated in the manuscript was the Detection, Quantification and Grading of Hepatic Steatosis using MR Dixon technique. Hepatic steatosis, itself, being a diffuse parenchymal liver process.

The comparison between liver lesions (whether focal or diffuse) was not part of the scope of the study – this is a departure from the study aims, and merits a separate study on its own. We did not include patients with known focal liver lesions in our study.

The following appears in the manuscript text in the Materials and Methods section:

''all consecutive adult patients (≥ 18 years old) that underwent a per-protocol clinical MR examination for the combined assessment of liver fat, iron and fibrosis for suspected diffuse parenchymal liver disease and without known focal liver lesions, were entered into the study.''

4. *'''The parameters of instruments from different providers should be considered.''*

There are numerous different MR Dixon techniques worldwide, not just by varying manufacturers/providers but also by a host of different institutions, some of which have developed their own home-made Dixon based pulse sequences. It is not practical, nor is it part of the scope of the study to source the numerous and varied different parameters used in clinical practice – some of this data is either not available, very hard to get hold of, or forms part of trade secrets. In our study, we have focussed on the instruments that are available to us clinically at our home institution.

The following appear in the manuscript text in the Discussion section:

*''Finally, our study was performed exclusively on a single vendor 1.5-T MR platform using clinically available proprietary software. We did not evaluate the confounding effects of higher magnet strengths (e.g. 3-T), different manufacturer platforms, **software or protocols** as it was not within the scope of our study. An ex-vivo phantom study reported that multi-echo Dixon techniques were accurate and reliable across 1.5-T and 3-T, different clinical platfforms and multiple institutions (29).''*

Reference

29. Hu HH, Yokoo T, Bashir MR, et al. Linearity and Bias of Proton Density Fat Fraction as a Quantitative Imaging Biomarker: A Multicenter, Multiplatform, Multivendor Phantom Study. Radiology. 2021;298(3):640-651. [PMID: 33464181 DOI: 10.1148/radiol.2021202912]

THANK YOU !

By the Authors