Reviewer #1:
Scientific Quality: Grade B (Very good)
Language Quality: Grade A (Priority publishing)
Conclusion: Minor revision
Specific Comments to Authors: Collin and colleagues report an important study evaluating the capability of different ultrasonographic tools to detect and measure hepatic steatosis. The findings provide reference to clinical diagnoses that will enhance precision and efficiency. The overall study is sound. I only have minor comments for the authors to consider. 1. Figure 1. Panel A in straight line, while panels B-D in curve. Since the authors stated Spearman correlation was performed in Results section, I am curious what regression analysis was perform in Figure 1. Please indicate the statistics clearly (linear vs non-linear) and the correlation parameters in figures, e.g. R-square, actual P value rather than p <0.01. 2. Generally the order of citing figures should follow the order it appears in the manuscript. It's weird that the present Results describe Figure 1, then jump to Figure 4A, 4B, and then figure 2, figure 4D. 3. No citation of Figure 3 in the text. This should be at least mentioned in Results. 4. Table 1. Could the author show the clinical threshold of biological data in healthy population as a reference? 5. Considering the age difference (mean±SD, 56±14), could the author discuss the impact of age on the observed diagnostic capability? 6. Is there any correlation between the severity of steatosis identified from US and biological parameters?

Reviewer #2:
Scientific Quality: Grade C (Good)
Language Quality: Grade B (Minor language polishing)
Conclusion: Major revision
Specific Comments to Authors: The author evaluated several US tools to detect and measure hepatic steatosis and found that hepatorenal index had the best performance. MRI proton density fat fraction assessment diagnosis was used as the "gold standard". However, there are some concerns 1, The sample size of this study is quite small, while only 19% of included patients without steatosis. 2, Since the result of US is usually influenced by the examinators. The author said that AC, SSE and HRI were assessed by two different examinators in limitations. Please showed the kappa test to determine the concordance between two examinators. 3, Why one patient failed to receive Fibroscan/US? 4, Please describe the statistical method used in Figure 1. 5, Please show the data in SSSE, AC, cCAP between non-steatotic and steatotic patients.

Reviewer #3:
Scientific Quality: Grade C (Good)
Language Quality: Grade B (Minor language polishing)
Conclusion: Accept (General priority)
Specific Comments to Authors: Thank you for your submission. Your manuscript was an interesting read. The manuscript is well organized and follows a clear flow.

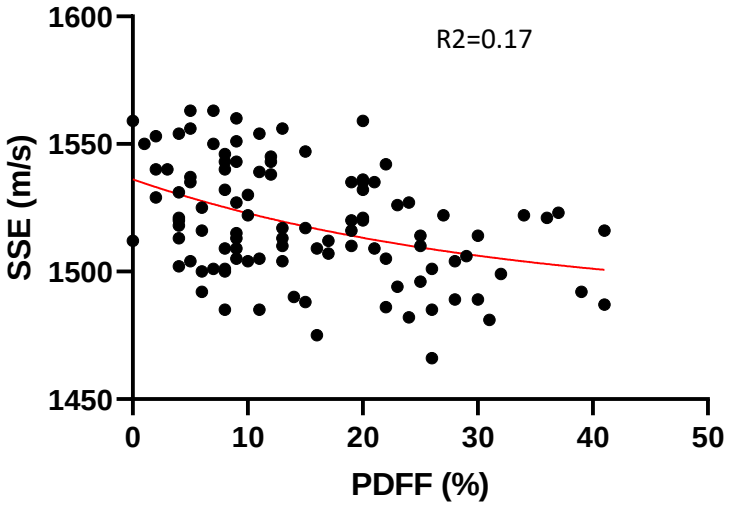
Reviewer #4:
Scientific Quality: Grade B (Very good)
Language Quality: Grade B (Minor language polishing)
Conclusion: Minor revision
Specific Comments to Authors: Overall, the study appears to be well-conducted and

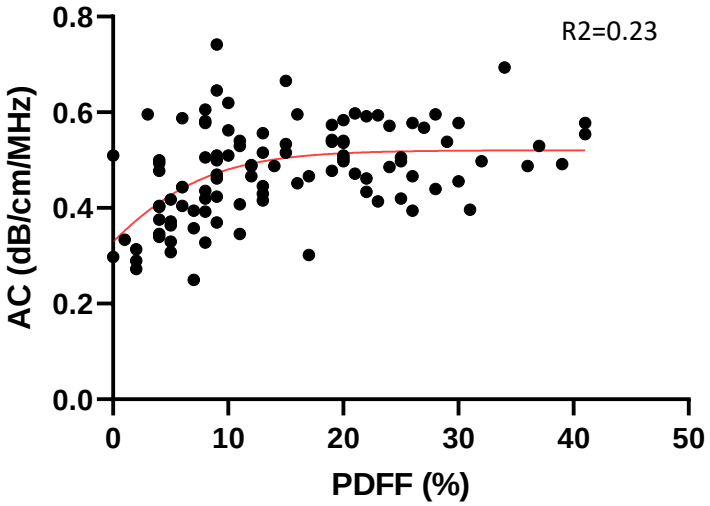
provides important insights into the diagnostic accuracy of various ultrasonographic tools for detecting hepatic steatosis in patients with NAFLD. The current study evaluated the diagnostic performance of various ultrasonographic tools for the detection of steatosis in an exclusively NAFLD patient population, using MRI-PDFF as the gold standard. The study found that standard ultrasound had poor sensitivity for mild steatosis and suffered from inter- and intra-observer variability. Hepato-renal index (HRI) was found to be the most reliable technique, followed by controlled attenuation parameter (cCAP), and acoustic radiation force impulse (ARFI) elastography-based liver fat quantification techniques, shear wave elastography (SSE), and ARFI elastography (AC). The study concluded that HRI had the best performance and was the simplest and most available method, while standard ultrasound should remain the first-line screening tool for steatosis. The study also noted that further validation of these results is needed in different populations and in a multicenter study. Recent studies suggest that VCTE may be the superior performing method available for assessing the degree of hepatic steatosis and fibrosis in the US population and in epidemiological studies (PMID: 36774231, PMID: 36460186), please discuss the advantages and disadvantages of other methods compared to VCTE.

Reviewer #5:
Scientific Quality: Grade C (Good)
Language Quality: Grade B (Minor language polishing)
Conclusion: Minor revision
Specific Comments to Authors: Authors have described different US and MRI based techniques to detect hepatic steatosis. I have few comments. 1. Authors have written it as a "Randomised clinical trial". They have used single cohort of patients prospectively and have evaluated different US and MRI based techniques on all of them. by definition "Randomised clinical trial" is "he process by which participants in clinical trials are assigned by chance to separate groups that are given different treatments or other interventions" Please explain as to why they think it to be RCT. 2. The comparison of time needed for acquisition of hepatic steatosis by a particular method, its cost should also be included. When we apply any method for a large population screening these parameters are relevant to be addressed.

We thank reviewer 1 for its relevant comments which can only improve the quality of our manuscript.

1) Indeed, we used non linear (logarithmic) regression for sound speed estimation, attenuation coefficient, hepatorenal index and controlled attenuation parameters as all these parameters reflects the attenuation of ultrasounds waves which are more and more dampened in liver parenchyma. This has been reported previously (Runge et al, Radiology 2018. Dioguardi et al. Ultraschall in Med 2018) and could explain why these tools failed to precisely quantifity liver steatosis. As requested we added R2 which allows to appreciate the goodness of fit and all non linear regression analysis were added in a supplementary figure 1. Linear correlation are presented for each tool in figure 2.





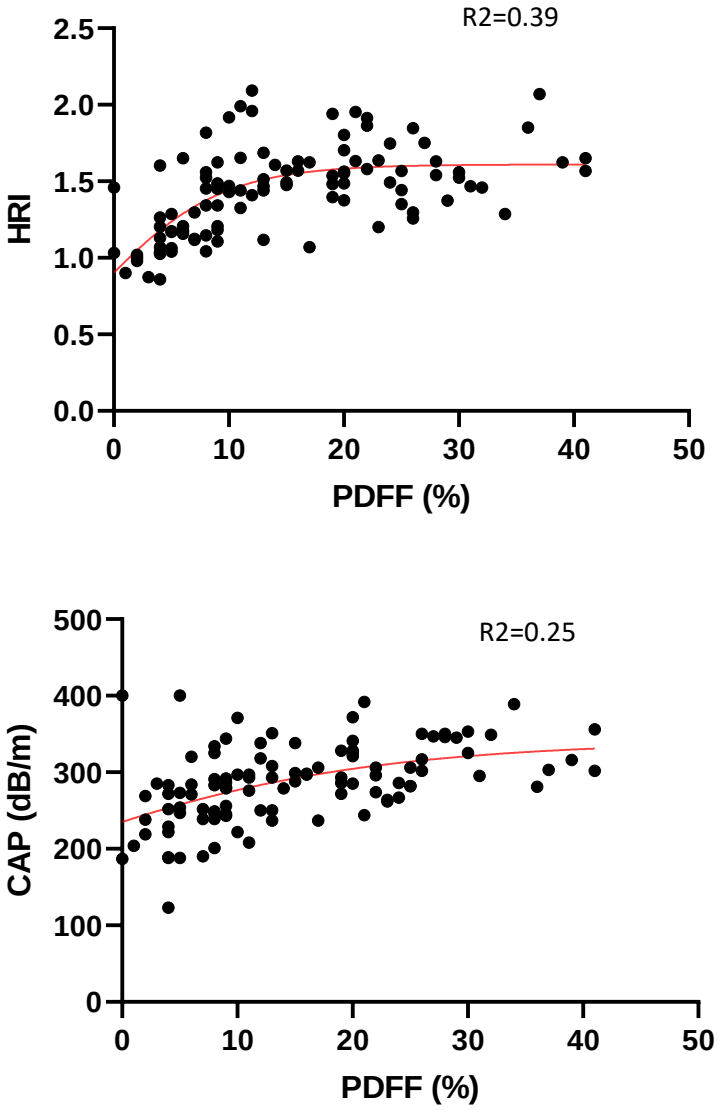


Figure 2 : Scatterplots showing non-linear (logarithmic) relationship between SSE (A), AC (B), HRI (C) and cCAP (D) with MRI PDFF using 6-echo gradient. PDFF = proton density fat fraction, cCAP = continuous controlled attenuation parameter, SSE = sound speed estimation, AC = attenuation coefficient, HRI = hepatorenal index

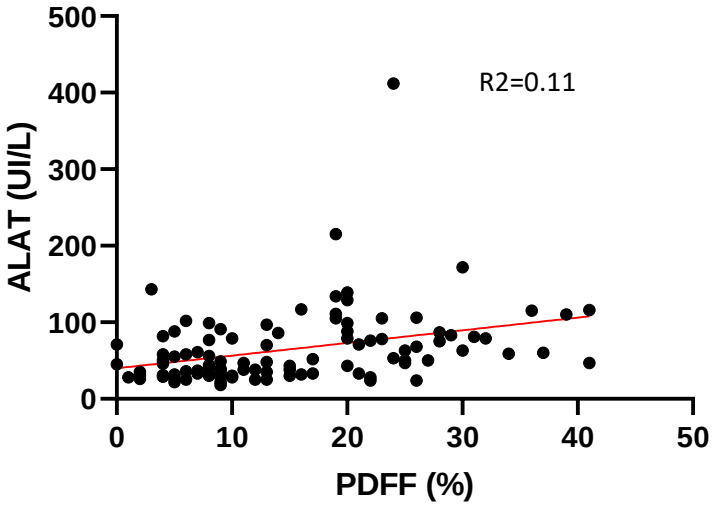
2) We corrected the order of appearance of each figure/table as requested. Figure 3 became figure 1 and was introduced in the patients and methods parts « HRI was calculated using the region of interest (ROI) measure tool, with average brightness ratio between two ROI at least 3 mm wide placed at the same depth in hepatic parenchyma and in renal cortex (figure 1) ».

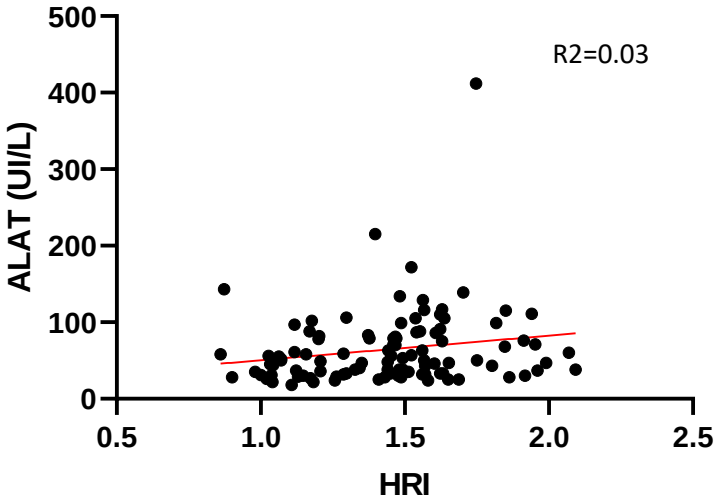
4) We added the main clinical and radiological parameters in healthy patients (table) and compared it to patients with steatosis.

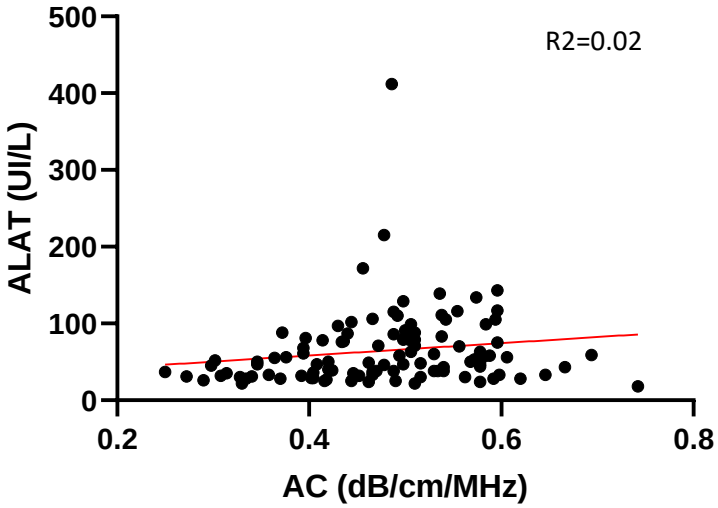
5) We did not find any correlation between age of patients and studied radiological parameters and therefore did not study the influence of age on the diagnosis

performances of these tools in multivariable analysis. Previous studies found a correlation between age and CAP but the increasing CAP values with age only reflected the increasing prevalence of obesity with age (Mjelle et al. Ultrasound Int Open 2021).

6) We only found a significant correlation between our gold standard PDFF and ALAT values as shown. Neither GGT nor ferritin correlated with PDFF. We also reported a significant correlation between AC, HRI and ALAT though the goodness of fit was rather low.






R2=0.02
ALAT (UI/L)
500
400
300
200
100
0
0.2
0.4
0.6
0.8
AC (dB/cm/MHz)

We thank reviewer 2 for its particular interest in our work and its relevant comments

1) We are aware of this important limitation of our study. All patients included were referred to our medical department for non-invasive fibrosis evaluation. More and more patients are adressed for suspected NAFLD. The unprecedented increase in the prevalence of metabolic diseases makes it difficult to include patients without steatosis. We are planning to expand our project to other hepatology departements in order to increase our cohort of patients with steatosis but mainly to obtain more healthy patients.

2) Unfortunately, the two examinators did not perform US liver examinations for all patients and we are unable to report an inter-observer concordance.

3) One patient failed to receive US scan/fibroscan due to very poor US signal. We consider the rate of failure very low compared to other studies. For instance, in the preliminary study evaluation SSE in patients, almost 25% had poor US signal which did not allow to calculate SSE.

4) Is the reviewer discussing figure 1 ? For figure 1, we used non linear (logarithmic) regression for sound speed estimation, attenuation coefficient, hepatorenal index and controlled attenuation parameters as all these parameters reflects the attenuation of ultrasounds waves which are more and more dampened in liver parenchyma.We consider that a linear correlation is not entirely appropriate for these US tools. For figure 5, a concordance table was designed and Fisher Test calculated. As requested we added the main clinical and radiological parameters in healthy patients (table) and compared it to patients with steatosis (table 1).

We thank reviewer 3 for its interest and very positive comments.

We thank reviewer 4 for its relevant comment and interest in our work.

We discussed the pros and cons of each tool including biological tools as suggested which will hopefully help clinicians and readers for the choice of the right tool to use.

We thank reviewer 5 for the interest he or she has taken in reading our work.

1) The "Randomised clinical trial" was an incorrect denomination and is the result of a mistake during the submission process.

2) Both cCAP and US measures of HRI requires a very short period of acquisition. HRI has the great advantage of being easily performed for someone already familiar with US scan evaluation. cCAP is measured during FibroScan® acquisition and does not raquire additional capabilities. SSE was the most difficult tool to tame, even after several weeks of training, as it requires a stable position, the strict absence of large hepatic vessels or artifacts in the image.