



# Reproducibility of hepatic MR elastography across field strengths, pulse sequences, scan intervals, and readers

Hye Jin Kim<sup>1</sup> · Bohyun Kim<sup>1</sup> · Hyun Jeong Yu<sup>1</sup> · Jimi Huh<sup>1</sup> · Jei Hee Lee<sup>1</sup> · Seung Soo Lee<sup>2</sup> · Kyung Won Kim<sup>2</sup> · Jai Keun Kim<sup>1</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

**Purpose** To evaluate the reproducibility of hepatic MRE under various combinations of settings of field strength, pulse sequence, scan interval, and reader in non-alcoholic fatty liver disease (NAFLD) patients.

**Methods** Adult NAFLD patients were prospectively enrolled for serial hepatic MRE with 1.5 T using 2D GRE sequence and 3.0 T using 2D SE-EPI sequence on the same day and after 2 weeks, resulting a total of four MRE examinations per patient. Three readers with various levels of background knowledge in MRE technique and liver anatomy measured liver stiffness after a training session. Linear regression, Bland–Altman analysis, within-subject coefficient of variation, and reproducibility coefficient (RDC) were used to determine reproducibility of hepatic MRE measurement.

**Results** Twenty patients completed the MRE sessions. Liver stiffness through MRE showed pooled RDC of 26% (upper 95% CI 30.6%) and corresponding limits of agreement (LOA) within 0.55 kPa across field strengths, MRE sequences, and 2-week interscan interval in three readers. Small mean biases and narrow LOA were observed among readers (0.05–0.19 kPa ± 0.53).

**Conclusion** The magnitude of change across combinations of scan parameters is within acceptable clinical range, rendering liver stiffness through MRE a reproducible quantitative imaging biomarker. A lower reproducibility was observed for measurements under different field strengths/MRE sequences at a longer (2 weeks) interscan interval. Operators should be trained to acquire region of interest consistently in repeat examinations.

**Keywords** Magnetic resonance elastography · Reproducibility of results · Non-alcoholic fatty liver disease

## Introduction

Magnetic resonance elastography (MRE) of the liver estimates the parenchymal stiffness that is positively correlated with the histopathologic stages of liver fibrosis induced by viral hepatitis and non-alcoholic fatty liver disease (NAFLD) [1–5]. Hepatic MRE uses phase contrast sequence and motion encoding gradient to capture the propagation of the shear wave generated by an external low frequency vibration [6–8]. Wave images of the hepatic shear wave in motion are

then processed to elastograms or stiffness maps to enable quantitative measurement of the hepatic parenchymal stiffness [8]. By reflecting liver fibrogenesis and ensuing collagen proliferation with crosslinking, MRE can yield high accuracy in predicting the severity of fibrosis noninvasively [9]. Hence, liver stiffness through MRE is a promising quantitative imaging biomarker (QIB) that could potentially replace biopsy [10, 11]. In recent clinical trials, liver stiffness obtained with MRE was shown to be a feasible QIB and adopted as one of the main outcomes in assessing the efficacy of drugs reducing hepatic steatosis in NAFLD [12–15].

Since scientific evidence increasingly indicates liver stiffness through MRE to serve as a QIB, studies to establish the variability of measurement, or precision, under a range of study conditions are needed [16]. Variability in the measurement of QIB is represented in terms of repeatability considered as the measurement variability under identical or nearly identical condition, and reproducibility, which refers to that under a variety of conditions [16, 17]. Reproducibility of

✉ Bohyun Kim  
kbh@ajou.ac.kr

<sup>1</sup> Department of Radiology, Ajou University School of Medicine, Ajou University Hospital, 164 World cup-ro, Yeongtong-gu, Suwon 16499, South Korea

<sup>2</sup> Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea

the QIB is more easily transferrable to actual clinical setting wherein the measurement value is acquired under inherently inconsistent conditions [17]. Specifically, in longitudinal study requiring comparisons of repeated measurements over time, a robust QIB should be constantly reproducible to ensure reliability of the study. In this setting, it is important to distinguish between change that represents a true response to intervention and variability in measurement, based on reproducibility of the QIB [17].

Prior studies have assessed the repeatability and reproducibility of hepatic MRE using the same or different parameters of field strength, MRE pulse sequence including gradient echo (GRE) and spin echo–echo planar (SE-EPI) sequences, imager manufacturer, scan interval, and driver [18–22]. These studies and a recent inclusive meta-analysis reported good repeatability/reproducibility of hepatic MRE and provided a pooled repeatability coefficient (RC) of 22% [23]; while variability of the measurement value under a single scan variant is discussed in depth, the effect of combining relevant factors might still needs to be validated. Our study aimed to demonstrate the reproducibility of liver stiffness measured with MRE under various combinations of settings of field strength, pulse sequence, scan interval, and readers expressed in reproducibility coefficient (RDC).

## Materials and methods

This prospective study was approved by our Institutional Review Board (IRB) and is compliant with Health Insurance Portability and Accountability Act (HIPAA). Written informed consent was obtained from all subjects.

### Study design

The current study population was included in the prior study which reported the repeatability and reproducibility of proton density fat fraction measurement [24]. On a sample size calculation, 19 subjects were required to detect a correlation of 0.6 or greater when statistical power was set at 0.8 under a significance level of 0.05. With a dropout rate of 20%, 24 subjects were planned to be recruited.

From August 2017 to December 2017, newly diagnosed NAFLD patients were prospectively enrolled from our outpatient clinic for liver disease. The diagnosis of NAFLD was based on radiologic evidence, clinical history, and biochemical profiles [25]: The inclusion criteria for the participants were (1)  $\geq 18$  years old, (2) hepatic parenchymal echogenicity brighter than renal cortex on screening ultrasound, (3) no significant amounts of alcohol consumption, (4) no coexisting chronic liver disease, and (5) no exposure to known hepatotoxins [25]. Subjects contraindicated to MRI were excluded from the study.

In all subjects, serial MRE was performed with 1.5 T (Signa HDxt; GE Healthcare, Milwaukee, WI, USA) and 3.0 T (Discovery 750 W; GE Healthcare) MR scanners on the same day less than 15 min apart, and after 2 weeks. MRE was performed at each field strength once per visit; thus, a total of four MRE examinations were performed per patient included in the study.

### MRE setup and protocol

MRE was obtained using a commercially available MRE technique (MR-Touch; GE Healthcare). Subjects were instructed to fast for at least 4 h before the examination, and positioned with a phase-array coil centered at the liver. A passive driver was securely placed against the right anterior chest wall over the liver to transmit the 60-Hz acoustic waves generated by an active driver located outside the MRI suite. Six experienced MRI technologists performed the MRE, and we did not specifically match the technologist with the patient. After finishing the 1st scan, the patient was transported to the other room for the 2nd scan.

Hepatic shear waves were imaged using motion-sensitized 2D GRE and 2D SE-EPI MRE pulse sequences with 1.5 T and 3.0 T MR systems, respectively [26, 27]. Four contiguous axial image slices through the mid liver were acquired during four cycles of end-expiratory breath-holds with 2D GRE sequence and two cycles of end-expiratory breath-holds with 2D SE-EPI sequence. Detailed parameters of MRE pulse sequences are shown in Table 1.

### Quantitative analysis of MRE

Wave images and elastograms (stiffness maps) were automatically generated for each image slice on the console, and

**Table 1** MR elastography pulse sequence parameters

Parameters	GE signa HDxt	GE discovery 750 W
Field strength	1.5 T	3.0 T
Pulse sequence	2D GRE	2D SE-EPI
TR	50 ms	1000 ms
TE	20 ms	50 ms
FOV	420×420 mm	420×420 mm
Matrix	256×64	64×64
Slice thickness	8 mm	8 mm
Number of slice	4	4
Slice spacing	2.5 mm	2.5 mm
Driver frequency	60 Hz	60 Hz
Acceleration factor	2	2
Acquisition time	56 s (14 s/slice)	28 s (7 s/slice)

GRE gradient recalled echo, SE-EPI spin echo–echo planar imaging, FOV field of view

these were submitted to a post processing software (READY View version 12.3; GE Healthcare). A 95% confidence map was superimposed on each elastogram to indicate the area of high-quality data to be included in the stiffness measurement [28].

Three readers with different levels of expertise in MRE technique and knowledge in liver anatomy independently measured the liver stiffness. Reader 1 and 2 (R1 and R2; K.B. and K.H.J.) were abdominal radiologists with more than 5 and 10 years of experience in abdominal MRI, respectively, and Reader 3 (R3; Y.H.J.) was an image analyst with 6 months of experience in processing body MR images. R1 had 3 years of experience in obtaining liver stiffness with MRE, and trained R2 and R3 with 20 clinical cases (4 image slices  $\times$  20 cases = 80 elastograms) before formal measurement was conducted. Regions of interest (ROIs) were manually drawn in the largest possible liver parenchyma bounded by confidence maps while avoiding the major vessels, biliary tree, and areas immediately under the liver capsule in the right liver [22, 29]. The readers were blinded to the clinical data of patients and measurement values by the other readers. Four separate review sessions were held for each reader with 1-week washout period during review of the MRE from the same patient. Liver stiffness in MRE was defined as the mean measurement value obtained from four image slices expressed in kilopascals (kPa) according to the method considered reproducible in prior studies [22, 30]. To assess intrareader agreement, repeat measurement values were obtained by R1 and R3 in the same manner. ROI sizes were recorded by R1.

### Statistical analysis

Liver stiffness values are presented as means and SD in kPa. A repeated measure analysis of variance (ANOVA) was used to assess differences in the value of liver stiffness among readers. A post-hoc Bonferroni test was used to identify significant difference in pairwise comparisons. Bonferroni-corrected *P* value of  $< .05$  was considered statistically significant.

Reproducibility of measurements across field strengths in the same day and at 2 weeks of interscan interval was assessed by within-subject coefficient of variation (wCV) and RDC [16, 17]. In cases of failed MRE acquisition, reproducibility analysis was conducted with the corresponding pairs excluded, and their values were only used for descriptive statistics. The bias of measurement was evaluated with Bland–Altman plots. The correlation of measurement was assessed by linear regression analysis. Interobserver agreements were assessed using two-way random-effects model intraclass correlation coefficients (ICCs). Intraobserver agreements for R1 and R3 were analyzed, and 95% confidence interval (CI) was calculated for each ICC. Paired

Student's *t* test was used to compare the size of ROIs between 1.5 and 3.0 T. *P* value of  $< .05$  was considered statistically significant.

Statistical analysis was performed using MedCalc version 19.0.3 (MedCalc, Ostend, Belgium) and Statistical Package for the Social Sciences (SPSS) version 25.0 (IBM, Armonk, NY, USA).

## Results

### Study population

Of 24 initially enrolled patients, four dropped out of the study: One patient was lost to follow-up and did not attend the 2nd MRE session, and the others withdrew their consents before the 1st MRE sessions. Finally, a total of 20 patients who completed the scheduled imaging sessions including 13 male individuals and seven female individuals with mean age  $\pm$  standard deviation (SD) of 43.2 years  $\pm$  15.4 (range 18–72) and mean body mass index  $\pm$  SD of 27.5 kg/m<sup>2</sup>  $\pm$  5.2 (range 19.2–38.2) were included.

### MRE failures

MRE was considered unsuccessful if the total number of the valid pixels bounded by the 95% confidence map was less than 500 in the four image slices [1]. MRE failure was present in four examinations (5%) of three patients. Of these, two failed MRE examinations in 3.0 T in two patients at the 1st visit in the early phase of the study was due to absence of waves on the wave images, which is probably associated with technical failure in the driver setup [1]. Subsequently, the function of active and passive drivers was double checked throughout the study. In one patient, there was failure of two consecutive MRE examinations in 1.5 T owing to lack of fit of the abdominal girth within the magnet bore.

### Liver stiffness values

The mean liver stiffness values by each reader in 1.5 T and 3.0 T at two visits are shown in Table 2. No significant intrareader difference in the measurement value at four different MRE examinations was obtained; whereas significant interreader difference in the measurement value was obtained between R3 and the other two readers, although the mean difference was small ( $< 0.3$  kPa).

### Reproducibility of the value of liver stiffness

The reproducibility of liver stiffness measurement at different field strengths/MRE sequences, and visits of each reader is shown in Table 3. At 2-week interscan interval, under

**Table 2** Mean liver stiffness measurement in MR elastography

	Reader 1	Reader 2	Reader 3	<i>P</i> value†
1.5 T on 1st visit	2.5 ± 0.5 (2.0–3.4)	2.5 ± 0.5 (2.0–3.7)	2.3 ± 0.3 (2.0–3.1)	< .01*
1.5 T on 2nd visit	2.5 ± 0.5 (1.8–3.5)	2.4 ± 0.5 (1.8–3.7)	2.3 ± 0.3 (1.8–3.0)	< .02**
3.0 T on 1st visit	2.6 ± 0.4 (2.0–3.5)	2.5 ± 0.4 (2.0–3.6)	2.4 ± 0.3 (2.1–3.4)	< .02*
3.0 T on 2nd visit	2.6 ± 0.4 (1.9–3.5)	2.5 ± 0.4 (1.7–3.4)	2.3 ± 0.3 (1.8–3.1)	< .02*
<i>P</i> value†	All > .05			

All values are in kPa. Numbers in parentheses are ranges

†Bonferroni-corrected *P* values

\*Significant difference between reader 3 and both reader 1 and 2

\*\*Significant difference between reader 3 and reader 1

**Table 3** Reproducibility of liver stiffness measurement in MR elastography

Scan conditions		Reader 1		Reader 2		Reader 3		Pooled	
Field strengths	Scan intervals	wCV (%)	RDC (%)	wCV (%)	RDC (%)	wCV (%)	RDC (%)	wCV (%)	RDC (%)
1.5 T	2-Week	6.8	18.8 (14.3, 27.4)	8.4	23.4 (17.7, 34.1)	6.5	18 (13.7, 26.3)	8.0	22.2 (19.5, 25.9)
3.0 T	2-Week	9.4	26 (19.6, 38.4)	7.3	20.1 (15.2, 29.7)	5.6	15.4 (11.6, 22.7)	7.8	21.6 (18.9, 25.9)
Crossover <sup>a</sup>	Same day	7.2–7.7	20–21 (15, 31.3)	5.4–7.2	15.4–19.8 (11.5, 28.9)	6–6.8	16.7–18.9 (12.7, 28.3)	7.1–8.0	19.8–22.1 (17.2, 25.7)
Crossover <sup>a</sup>	2-Week	7.2–9.8	20–27 (15.2, 40.8)	6.7–10.3	18.6–28.5 (14.1, 42.7)	4.9–8.3	13.5–23 (10.2, 34.5)	7.1–9.4	19.7–26.0 (17.3, 30.6)

All values are in %. Numbers in parentheses are 95% confidence interval for the %RDC  
wCV within-subject coefficient of variation, RDC reproducibility coefficient

<sup>a</sup>Across field strengths and MR elastography sequences

the same field strength, the range of wCV and RDC in each reader was 5.6–9.4% and 15.4–26%, respectively, and for pooled values of all readers, the RDC in both 1.5 T and 3.0 T was between 21.6 and 22%. Reproducibility tended to be lower for MRE performed at different field strengths/MRE sequences: largest values of wCV and RDC were obtained when stiffness was measured across the field strengths/MRE sequences over 2 weeks of interval (wCV range 4.9–10.3%; RDC range 13.5–28.5%). The range of pooled values of RDC was between 19.7 and 26%.

The mean biases and 95% limit of agreement (LOA) of the liver stiffness measurement are shown in Fig. 1. At the same field strength over 2-week interval, the mean bias ± LOA was 0.04 kPa ± 0.49 under 1.5 T and –0.03 kPa ± 0.53 under 3.0 T. Across field strengths/MRE sequences, the mean bias ± LOA was –0.03 kPa ± 0.46 at the same day, and –0.02 kPa ± 0.55 at 2 weeks of interscan interval.

### Linear regression of liver stiffness measurement

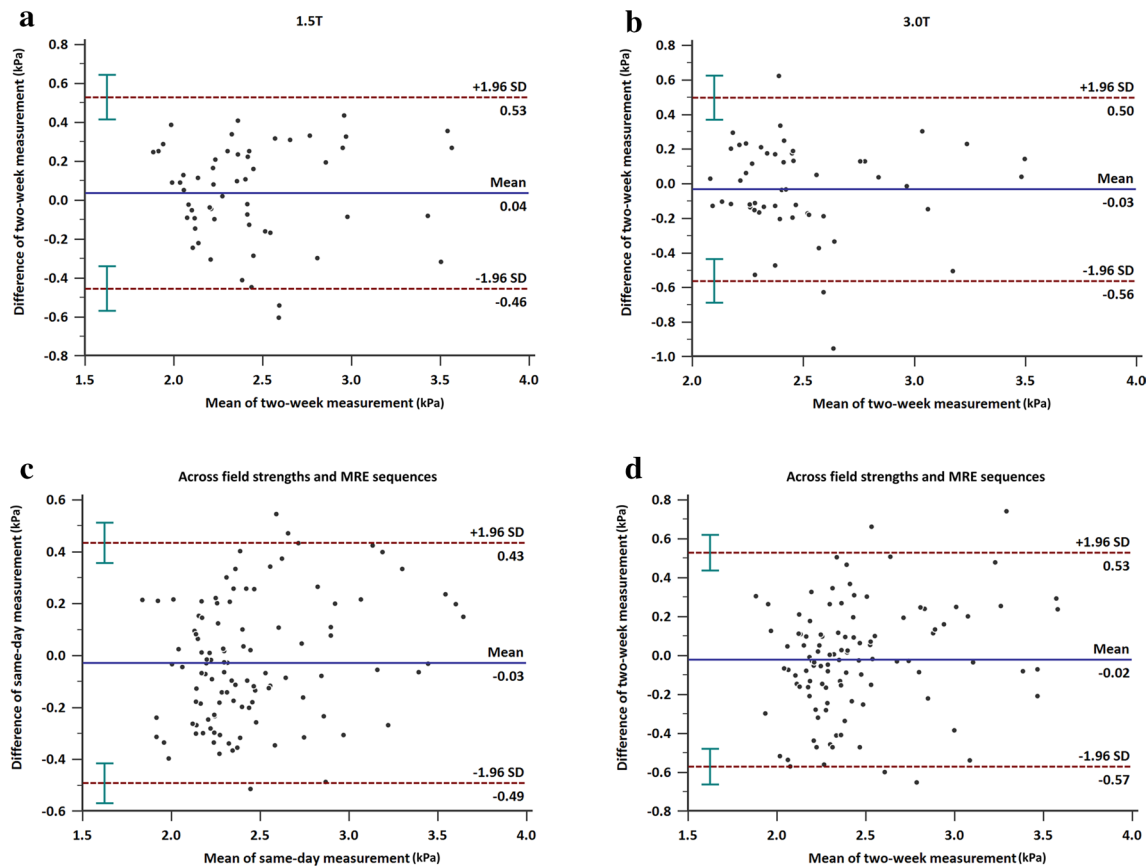
Linear correlations of measurement values under various scanning conditions and their correlation coefficients ( $r^2$ )

are shown in Fig. 2. For 114 measurements under 1.5 T (19 subjects × two scans × three readers) and 108 measurements under 3.0 T (18 subjects × two scans × three readers), moderate to strong linear correlation of the pooled values of liver stiffness at the same field strength over 2 weeks of interscan interval was observed ( $r^2=0.684$  for 1.5 T;  $r^2=0.515$  for 3.0 T). For 216 measurements [(17 subjects × two scans × three readers) + (19 subjects × two visits × three readers)], strong linear correlation of the measurement values across field strengths/MRE sequences on the same day ( $r^2=0.701$ ), and moderate linear correlation ( $r^2=0.586$ ) of those at 2-week interscan interval were observed.

### Intra- and interobserver agreements

The intrareader agreement for both R1 and R3 was excellent (ICC = 0.944; 95% CI 0.850–0.975 for R1; ICC = 0.817; 95% CI 0.677–0.900 for R3). Interreader agreement among all three readers was excellent (ICC = 0.958; 95% CI 0.939–0.972).

Mean biases of the measurement values by readers are shown in Fig. 3. The range of mean biases was 0.05–0.19 kPa, and LOAs were within 0.53 kPa.



**Fig. 1** Bland–Altman plots for the mean biases and limits of agreement (LOAs) of pooled liver stiffness measured under various scan conditions. **a** measurement at 2-week interscan interval with 1.5 T; **b** measurement at 2-week interscan interval with 3.0 T; **c** measurement

performed on the same day across field strengths/MRE sequences; **d** measurement at 2-week interscan interval across field strengths/MRE sequences. Mean biases range from  $-0.03$  to  $0.04$  kPa, and LOAs are within  $0.55$  kPa

## ROI size comparisons

The mean size of ROIs was significantly larger in 2D SE-EPI with 3.0 T MR system than in 2D GRE sequence with 1.5 T MR system ( $80.2 \text{ cm}^2 \pm 21.5$  vs.  $44.9 \text{ cm}^2 \pm 16.0$ ;  $P < .001$ ) (Fig. 4).

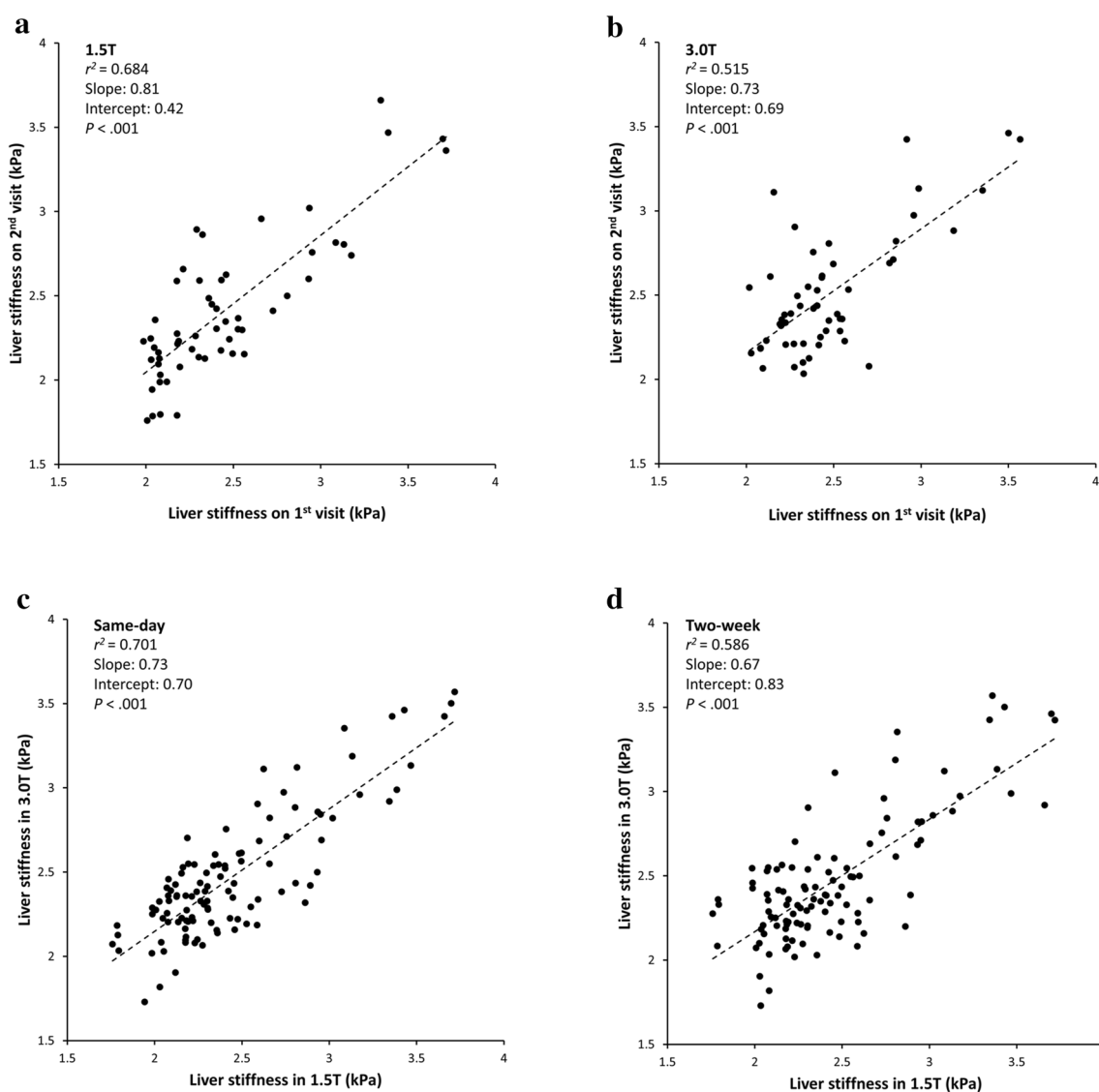
## Discussion

In our study, MRE-measured liver stiffness showed RDC of 13.5–28.5% in three readers under various combinations of scanning conditions, with a trend of lower reproducibility for measurement values acquired under different field strengths/MRE sequences at a longer (2 weeks) interscan interval but not in the specific reader. For measurements across field strengths/MRE sequences over 2-week of interval, the largest RDC in pooled measurement from all readers was 26% with upper 95% CI of 30.6%, which corresponded to a change of 0.55 kPa. Based on the liver stiffness of approximately 2.0–2.5 kPa reported in previous studies to

differentiate between advanced fibrosis (F3 and F4 stages) and no to mild to moderate fibrosis (F0, F1, and F2 stages), the reproducibility of the value of liver stiffness obtained in our study is acceptable, and MRE-measured liver stiffness could be used as a reproducible QIB [1, 11].

A recent meta-analysis reported that the summary RC of hepatic MRE was 22% (95% CI 16.1%, 28.2%), and the RC estimate tended to be higher in studies using untrained operators, 1.5 T rather than 3.0 T, and longer interscan periods [23]. In our study, the scan at longer interscan interval of 2 weeks resulted in a higher RDC than that acquired in the same day across field strengths/MRE sequences. Under the same field strength (1.5 T), RC ranged from 14 to 37% at 2- to 4-week of scan interval [19, 31, 32], which is in agreement with our result of similar magnitude of RDC at 2 weeks of interscan interval (18% to 23.4% in 1.5 T; 15.4% to 26% in 3.0 T). The impact of longer interscan period is also reflected in linear regression analysis across visits, where the slopes were 0.81 for 1.5 T and 0.73 for 3.0 T. Though these are close to 1, these discrepancies and higher RDC may be attributed to variables including and not limited to operators,





**Fig. 2** Linear correlations of pooled measurement values of liver stiffness. Moderate to strong linear correlations are observed among the measurement values from **a** 1.5 T at 2-week interscan inter-

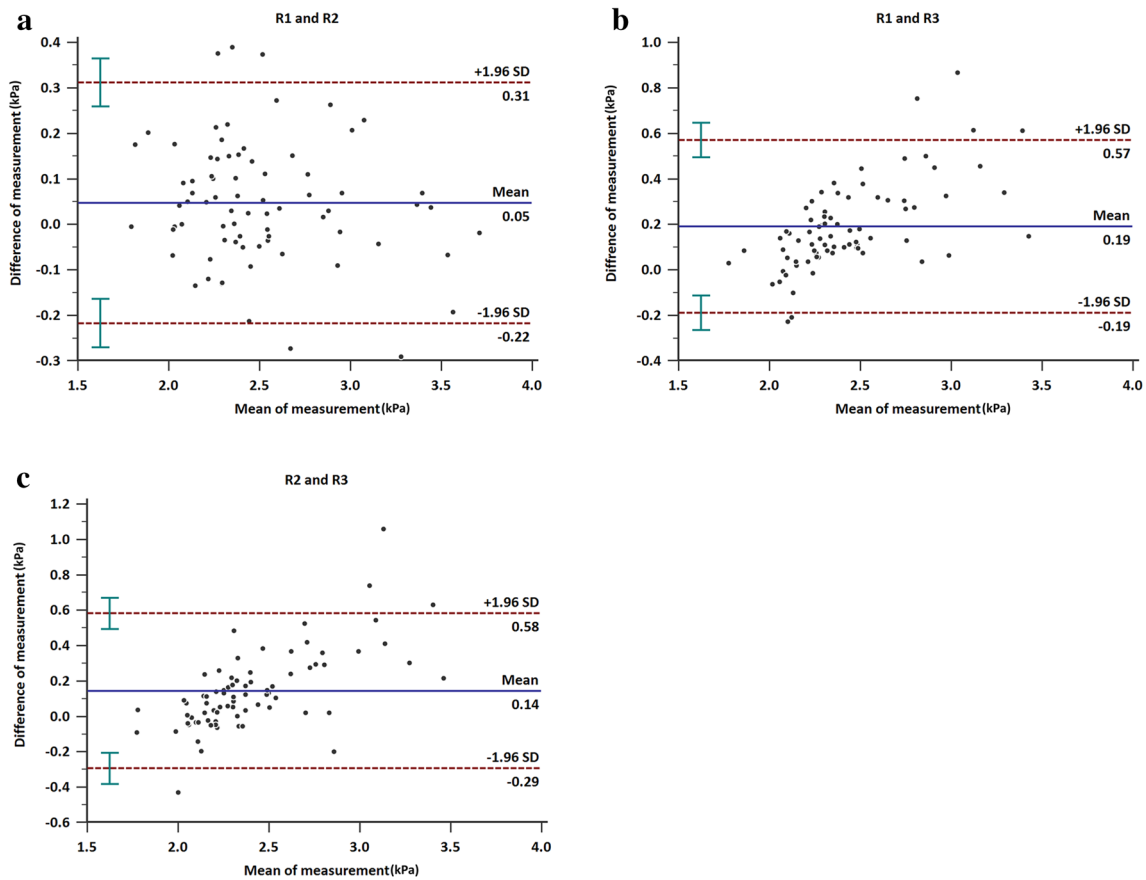
val, **b** 3.0 T at 2-week interscan interval, **c** across field strengths/MRE sequences in the same day, and **d** across field strengths/MRE sequences at 2-week interscan interval

technologists, reviewers, patient motion, and other patient characteristics [16]. Nevertheless, the percentage RDCs in pooled measurement are within 0.53 kPa of absolute value, which is considered a relatively narrow range of variability of the measurement.

Few studies have assessed the reproducibility of hepatic MRE across field strengths. Trout et al. conducted analysis of scans acquired on the same day and reported good to excellent agreement of the values (ICC of 0.62–0.84) obtained across field strengths using 2D GRE, 2D SE-EPI, and 3D SE-EPI MRE sequences [22]. Yasar et al. reported RDC of 25.4% to 29.7% for scans acquired on the same day with 1.5 T (Siemens) and 3.0 T (GE) platforms using 2D GRE which reflects the combined effect of field strength

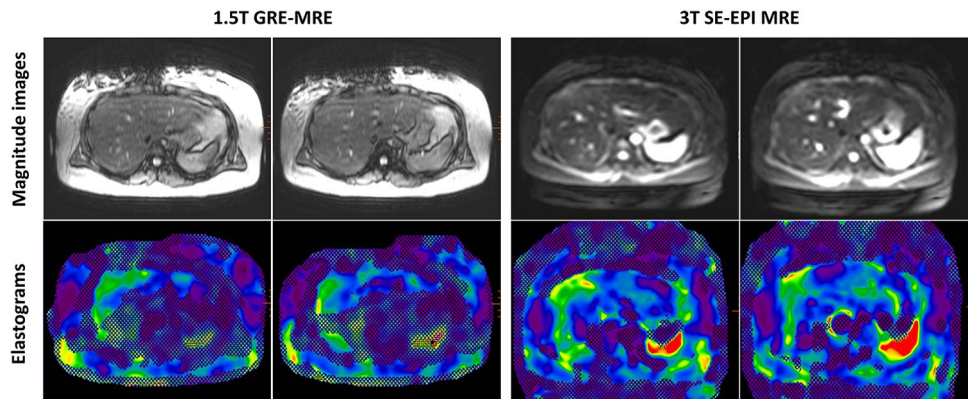
and vendor differences [33]. In our study, the RDC for scans acquired on the same day across field strengths/MRE sequences ranged from 15.4 to 21% in each reader, and the pooled value of RDC ranged from 19.8 to 22.1%. These values are similar to those obtained under the same field strength at 2 weeks of interscan interval, which suggests that switching of field strength and MRE sequence has as comparable impact to extending the scan interval to 2 weeks, on the reproducibility of hepatic MRE.

In our study, RDC under different field strengths tended to increase to 28.5% when the scan interval was extended to 2 weeks; although we used a distinct MRE sequence for each field strength, we obtained better reproducibility of results than those of prior studies, which may be explained



**Fig. 3** Bland–Altman plots for the mean biases and limits of agreement (LOAs) of liver stiffness between **a** Reader 1 and 2, **b** Reader 1 and 3, and **c** Reader 2 and 3. Mean biases range from 0.05 to 0.19 kPa, and LOAs are within 0.53 kPa

**Fig. 4** Representative 2D GRE and 2D SE-EPI MR elastography at 1.5 T and 3 T, respectively, in a 28-year-old female patient on her 1st visit. Mean liver stiffness was  $3.3 \text{ kPa} \pm 0.3$  at 1.5 T 2D GRE and  $3.0 \text{ kPa} \pm 0.2$  at 3 T 2D SE-EPI MR elastography. The mean size of ROI was  $53.7 \text{ cm}^2 \pm 8.7$  at 1.5 T 2D GRE and  $90.9 \text{ cm}^2 \pm 9.9$  at 3 T 2D SE-EPI MR elastography



by the fact that imaging was performed using scanners from the same manufacturer. Indeed, in the comparison of measurements from two imager manufacturers under fixed field strength and MRE sequence, statistically significant difference was noticed across manufactures at 3.0 T [22]. Thus, the largest pooled RDC of 26% with corresponding absolute difference within 0.55 kPa for measurements across field strengths/MRE sequences and 2-week interscan interval in

our study may have limited applicability to identical vendor platforms. Current reports have indicated cross-vendor validation of hepatic MRE for only few of the imager manufacturers and further verification is needed [22, 33, 34]. Our study provides the margin of liver stiffness through MRE that reflects true change across field strengths/MRE sequences over 2-week of interscan interval, and highlights that hepatic MRE is a robust QIB.

The newly introduced 2D SE-EPI MRE sequence has shorter scan time and higher signal-to-noise ratio than the 2D GRE MRE sequence [35, 36]. Studies have shown that 2D SE-EPI MRE is advantageous since it achieves a single breath-hold scan and high interobserver agreement while providing larger measurable ROI, lower failure rate, and comparable liver stiffness value as compared to 2D GRE [35–39]. In line with prior publications, the mean size of ROIs was nearly two-fold larger in 2D SE-EPI with 3.0 T than in 2D GRE sequence with 1.5 T units in our study. Currently, 2D GRE MRE sequence is most widely available, and more up-to-date MRE sequences including 2D SE-EPI MRE sequence are being developed; hence, reproducibility among sequences might become an important issue.

ROI drawn by a trained operator can substantially improve the reproducibility [22]. Therefore, before the start of our study, the readers underwent training for drawing a single free-hand ROI [29]. As a result, we obtained small variation in RDC with satisfactory interreader LOA ( $\pm 0.53$  kPa), and achieved excellent agreement among three independent readers (ICC = 0.958) with various levels of background knowledge in MRE technique and liver anatomy. Moreover, we achieved high intrareader agreements between the readers with the most and least experience in hepatic MRE measurement. However, the number of cases considered sufficient for training remains unclear, and evaluation of the actual impact of training is needed since we did not collect measurement values of readers before they underwent training. Operators should be trained to acquire ROIs consistently in repeat examinations.

Our study has several limitations. First, we were unable to assess the reproducibility across imager manufacturers due to limitation of available MR system in our institution. Second, although we prospectively enrolled patients with chronic liver disease with NAFLD, the cohort of our study did not include a wide spectrum of liver stiffness/fibrosis stage and the number of participants was small. Using threshold values by Hsu et al. assessing NAFLD, the majority of the patients belonged to F0 (< 2.61 kPa), and some were in F1 (< 2.97 kPa), F2 (< 3.62 kPa), and F3 (< 4.7 kPa) but none were in F4 [40]. Therefore, reproducibility in higher liver stiffness/fibrosis stage may still need further validation. Third, we were not able to assess the influence of MRE sequence on reproducibility separately from that of field strength, though the overall reproducibility was within an acceptable clinical range. One study performed direct assessment of the pairwise reproducibility of hepatic MRE under different pulse sequences [22]. Finally, we did not evaluate repeatability of measurement values acquired on the same day under the same field strength and therefore, there may be limitation in analyzing the influence of scanning interval on the reproducibility. Previous studies assessing hepatic MRE performed on the same day within the same

field strength have reported high repeatability of the results obtained [20, 22, 23].

In conclusion, liver stiffness through MRE showed pooled RDC of 26% (upper 95% CI 30.6%) and corresponding LOA within 0.55 kPa across field strengths, MRE sequences, and 2-week interscan interval for trained readers with various levels of experience in MRE. The measurement values were within acceptable clinical range of the magnitude of change, rendering liver stiffness through MRE a reproducible QIB.

**Acknowledgements** This work was supported by the new faculty research fund of Ajou University School of Medicine. The recipient of the fund is Bohyun Kim.

## References

1. Yin M, Glaser KJ, Talwalkar JA, et al. (2016) Hepatic MR Elastography: Clinical Performance in a Series of 1377 Consecutive Examinations. *Radiology* 278:114–124. <https://doi.org/10.1148/radiol.2015142141>.
2. Batheja M, Vargas H, Silva AM, et al. (2015) Magnetic resonance elastography (MRE) in assessing hepatic fibrosis: performance in a cohort of patients with histological data. *Abdom Imaging* 40:760–765. <https://doi.org/10.1007/s00261-014-0321-8>.
3. Venkatesh SK, Wang G, Lim SG, Wee A (2014) Magnetic resonance elastography for the detection and staging of liver fibrosis in chronic hepatitis B. *Eur Radiol* 24:70–78. <https://doi.org/10.1007/s00330-013-2978-8>.
4. Singh S, Venkatesh SK, Wang Z, et al. (2015) Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. *Clin Gastroenterol Hepatol* 13:440–451.e446. <https://doi.org/10.1016/j.cgh.2014.09.046>.
5. Singh S, Venkatesh SK, Loomba R, et al. (2016) Magnetic resonance elastography for staging liver fibrosis in non-alcoholic fatty liver disease: a diagnostic accuracy systematic review and individual participant data pooled analysis. *Eur Radiol* 26:1431–1440. <https://doi.org/10.1007/s00330-015-3949-z>.
6. Manduca A, Oliphant TE, Dresner MA, et al. (2001) Magnetic resonance elastography: non-invasive mapping of tissue elasticity. *Med Image Anal* 5:237–254.
7. Muthupillai R, Lomas DJ, Rossman PJ, et al. (1995) Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. *Science* 269:1854–1857.
8. Venkatesh SK, Ehman RL (2015) Magnetic resonance elastography of abdomen. *Abdom Imaging* 40:745–759. <https://doi.org/10.1007/s00261-014-0315-6>.
9. Reiter R, Freise C, Johrens K, et al. (2014) Wideband MRE and static mechanical indentation of human liver specimen: sensitivity of viscoelastic constants to the alteration of tissue structure in hepatic fibrosis. *J Biomech* 47:1665–1674. <https://doi.org/10.1016/j.jbiomech.2014.02.034>.
10. Dulai PS, Sirlin CB, Loomba R (2016) MRI and MRE for non-invasive quantitative assessment of hepatic steatosis and fibrosis in NAFLD and NASH: Clinical trials to clinical practice. *J Hepatol* 65:1006–1016. <https://doi.org/10.1016/j.jhep.2016.06.005>.
11. Morisaka H, Motosugi U, Ichikawa S, et al. (2018) Magnetic resonance elastography is as accurate as liver biopsy for liver fibrosis staging. *J Magn Reson Imaging* 47:1268–1275. <https://doi.org/10.1002/jmri.25868>.



12. Ajmera VH, Cachay E, Ramers C, et al. (2019) Novel MRI assessment of treatment response in HIV-associated NAFLD: a randomized trial of an SCD1 inhibitor (ARRIVE Trial). *Hepatology*. <https://doi.org/10.1002/hep.30674>.
13. Loomba R, Sirlin CB, Ang B, et al. (2015) Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). *Hepatology* 61:1239-1250. <https://doi.org/10.1002/hep.27647>.
14. Jayakumar S, Middleton MS, Lawitz EJ, et al. (2019) Longitudinal correlations between MRE, MRI-PDFF, and liver histology in patients with non-alcoholic steatohepatitis: Analysis of data from a phase II trial of selonsertib. *J Hepatol* 70:133-141. <https://doi.org/10.1016/j.jhep.2018.09.024>.
15. Cui J, Philo L, Nguyen P, et al. (2016) Sitagliptin vs. placebo for non-alcoholic fatty liver disease: A randomized controlled trial. *J Hepatol* 65:369-376. <https://doi.org/10.1016/j.jhep.2016.04.021>.
16. Sullivan DC, Obuchowski NA, Kessler LG, et al. (2015) Metrology Standards for Quantitative Imaging Biomarkers. *Radiology* 277:813-825. <https://doi.org/10.1148/radiol.2015142202>.
17. Raunig DL, McShane LM, Pennello G, et al. (2015) Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. *Stat Methods Med Res* 24:27-67. <https://doi.org/10.1177/0962280214537344>.
18. Wang K, Manning P, Szevenyi N, et al. (2017) Repeatability and reproducibility of 2D and 3D hepatic MR elastography with rigid and flexible drivers at end-expiration and end-inspiration in healthy volunteers. *Abdom Radiol (NY)* 42:2843-2854. <https://doi.org/10.1007/s00261-017-1206-4>.
19. Hines CD, Bley TA, Lindstrom MJ, Reeder SB (2010) Repeatability of magnetic resonance elastography for quantification of hepatic stiffness. *J Magn Reson Imaging* 31:725-731. <https://doi.org/10.1002/jmri.22066>.
20. Lee Y, Lee JM, Lee JE, et al. (2014) MR elastography for noninvasive assessment of hepatic fibrosis: reproducibility of the examination and reproducibility and repeatability of the liver stiffness value measurement. *J Magn Reson Imaging* 39:326-331. <https://doi.org/10.1002/jmri.24147>.
21. Shire NJ, Yin M, Chen J, et al. (2011) Test-retest repeatability of MR elastography for noninvasive liver fibrosis assessment in hepatitis C. *J Magn Reson Imaging* 34:947-955. <https://doi.org/10.1002/jmri.22716>.
22. Trout AT, Serai S, Mahley AD, et al. (2016) Liver Stiffness Measurements with MR Elastography: Agreement and Repeatability across Imaging Systems, Field Strengths, and Pulse Sequences. *Radiology* 281:793-804. <https://doi.org/10.1148/radiol.2016160209>.
23. Serai SD, Obuchowski NA, Venkatesh SK, et al. (2017) Repeatability of MR Elastography of Liver: A Meta-Analysis. *Radiology*:161398. <https://doi.org/10.1148/radiol.2017161398>.
24. Kim HJ, Cho HJ, Kim B, et al. (2019) Accuracy and precision of proton density fat fraction measurement across field strengths and scan intervals: A phantom and human study. *J Magn Reson Imaging* 50:305-314. <https://doi.org/10.1002/jmri.26575>.
25. Chalasani N, Younossi Z, Lavine JE, et al. (2018) The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 67:328-357. <https://doi.org/10.1002/hep.29367>.
26. Venkatesh SK, Yin M, Ehman RL (2013) Magnetic resonance elastography of liver: technique, analysis, and clinical applications. *J Magn Reson Imaging* 37:544-555. <https://doi.org/10.1002/jmri.23731>.
27. Yin M, Talwalkar JA, Glaser KJ, et al. (2007) Assessment of hepatic fibrosis with magnetic resonance elastography. *Clin Gastroenterol Hepatol* 5:1207-1213.e1202. <https://doi.org/10.1016/j.cgh.2007.06.012>.
28. Silva AM, Grimm RC, Glaser KJ, et al. (2015) Magnetic resonance elastography: evaluation of new inversion algorithm and quantitative analysis method. *Abdom Imaging* 40:810-817. <https://doi.org/10.1007/s00261-015-0372-5>.
29. MR Elastography Biomarker Committee. MR Elastography of the Liver, Quantitative Imaging Biomarker Alliance. Profile Stage: Consensus. QIBA, May 2, 2018. Available from: <http://qibawiki.rsna.org/index.php/Profiles>
30. Rusak G, Zawada E, Lemanowicz A, Serafin Z (2015) Whole-organ and segmental stiffness measured with liver magnetic resonance elastography in healthy adults: significance of the region of interest. *Abdom Imaging* 40:776-782. <https://doi.org/10.1007/s00261-014-0278-7>.
31. Wang Y, Ganger DR, Levitsky J, et al. (2011) Assessment of chronic hepatitis and fibrosis: comparison of MR elastography and diffusion-weighted imaging. *AJR Am J Roentgenol* 196:553-561. <https://doi.org/10.2214/ajr.10.4580>.
32. Shin SU, Lee JM, Yu MH, et al. (2014) Prediction of esophageal varices in patients with cirrhosis: usefulness of three-dimensional MR elastography with echo-planar imaging technique. *Radiology* 272:143-153. <https://doi.org/10.1148/radiol.14130916>.
33. Yasar TK, Wagner M, Bane O, et al. (2016) Interplatform reproducibility of liver and spleen stiffness measured with MR elastography. *J Magn Reson Imaging* 43:1064-1072. <https://doi.org/10.1002/jmri.25077>.
34. Serai SD, Yin M, Wang H, et al. (2015) Cross-vendor validation of liver magnetic resonance elastography. *Abdom Imaging* 40:789-794. <https://doi.org/10.1007/s00261-014-0282-y>.
35. Wagner M, Besa C, Bou Ayache J, et al. (2016) Magnetic Resonance Elastography of the Liver: Qualitative and Quantitative Comparison of Gradient Echo and Spin Echo Echo-planar Imaging Sequences. *Invest Radiol* 51:575-581. <https://doi.org/10.1097/rli.0000000000000269>.
36. Mariappan YK, Dzyubak B, Glaser KJ, et al. (2017) Application of Modified Spin-Echo-based Sequences for Hepatic MR Elastography: Evaluation, Comparison with the Conventional Gradient-Echo Sequence, and Preliminary Clinical Experience. *Radiology* 282:390-398. <https://doi.org/10.1148/radiol.2016160153>.
37. Serai SD, Dillman JR, Trout AT (2017) Spin-echo Echo-planar Imaging MR Elastography versus Gradient-echo MR Elastography for Assessment of Liver Stiffness in Children and Young Adults Suspected of Having Liver Disease. *Radiology* 282:761-770. <https://doi.org/10.1148/radiol.2016160589>.
38. Zhan C, Kannengiesser S, Chandarana H, et al. (2019) MR elastography of liver at 3 Tesla: comparison of gradient-recalled echo (GRE) and spin-echo (SE) echo-planar imaging (EPI) sequences and agreement across stiffness measurements. *Abdom Radiol (NY)*. <https://doi.org/10.1007/s00261-019-01932-5>.
39. Kim YS, Jang YN, Song JS (2018) Comparison of gradient-recalled echo and spin-echo echo-planar imaging MR elastography in staging liver fibrosis: a meta-analysis. *Eur Radiol* 28:1709-1718. <https://doi.org/10.1007/s00330-017-5149-5>.
40. Hsu C, Caussy C, Imajo K, et al. (2019) Magnetic Resonance vs Transient Elastography Analysis of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Pooled Analysis of Individual Participants. *Clin Gastroenterol Hepatol* 17:630-637.e638. <https://doi.org/10.1016/j.cgh.2018.05.059>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.