



Published in final edited form as:

*J Magn Reson Imaging*. 2010 March ; 31(3): 725–731. doi:10.1002/jmri.22066.

## Repeatability of Magnetic Resonance Elastography for Quantification of Hepatic Stiffness

Catherine D. G. Hines, MS<sup>1</sup>, Thorsten A. Bley, MD<sup>2,3</sup>, Mary J. Lindstrom, PhD<sup>4</sup>, and Scott B. Reeder, MD, PhD<sup>1,2,5,6</sup>

<sup>1</sup>Department of Biomedical Engineering, University of Wisconsin, Madison, WI

<sup>2</sup>Department of Radiology, University of Wisconsin, Madison, WI

<sup>3</sup>Department of Radiology, Freiburg University, Madison, WI

<sup>4</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI

<sup>5</sup>Department of Medical Physics, University of Wisconsin, Madison, WI

<sup>6</sup>Department of Medicine, University of Wisconsin, Madison, WI

### Abstract

**Purpose**—To determine the sources of variability of MRE hepatic stiffness measurements using healthy volunteers and patients and to calculate the minimum change required for statistical significance. Hepatic stiffness measured with magnetic resonance elastography (MRE) has demonstrated tremendous potential as a non-invasive surrogate of hepatic fibrosis, although the underlying repeatability of MRE for longitudinal tracking of liver disease has not been documented.

**Materials and Methods**—MRE stiffness measurements from twenty healthy volunteers and ten patients were obtained twice on the same day, and repeated 2–4 weeks later for volunteers in this IRB-approved study. A linear mixed effects model was used to estimate the component sources of variability in the data.

**Results**—The standard deviation of MRE measurements of the same individual on different days is 11.9% (percent of the measured stiffness) using the same reader and 12.0% using different readers. The standard deviation of the difference between two measurements (i.e. longitudinal change in an individual) is 17.4%; the corresponding 95% confidence interval for zero change is (-27.0%, 37.0%).

**Conclusion**—MRE is a repeatable method for quantifying liver stiffness. Using the described MRE technique, changes greater than 37.0% of the smaller measured stiffness value represent meaningful changes in longitudinal liver stiffness measurements.

### Keywords

Magnetic resonance elastography; repeatability; liver; chronic liver disease; fibrosis; cirrhosis

### Introduction

Cirrhosis, the end-stage of chronic liver injury, is widespread and increasingly prevalent, caused by alcohol abuse, viral hepatitis, non-alcoholic fatty liver disease (NAFLD) and

numerous other conditions (1,2). Attempts to intervene, halt, and even reverse liver damage requires accurate detection of fibrosis early in the disease progression, before end-stage fibrosis (cirrhosis) occurs. Detection with of early liver injury is particularly important in many conditions such as NAFLD, where patients may be asymptomatic until they present in the end stages (cryptogenic cirrhosis)(2). Recent research has also shown that some stages of fibrosis may be reversible, making early detection and quantitative assessment of fibrosis necessary for intervention and treatment (2-5).

Liver biopsy currently remains the gold standard for detection and assessment of fibrosis. Liver biopsy carries the potential risk of severe complications (death in 1:10,000) (6), is expensive, and is inherently limited due to sampling variability caused by the heterogeneous nature of features of chronic liver disease, including fibrosis (5,7,8). For these reasons, non-invasive methods that can detect and quantify fibrosis are of great interest in the management of chronic liver disease. In addition, the ability to perform repeated measurements is highly desirable for therapeutic monitoring. Other methods for the non-invasive measurement of hepatic fibrosis that are under development include ultrasound elastography techniques (9,10) and serology tests (7,10), although the efficacy of the latter remains to be determined (2,7).

MR elastography (MRE) is an established imaging method for the non-invasive assessment of the mechanical properties of soft tissue (5,11-15). The propagation of acoustic strain waves through the tissue is measured from motion encoded phase difference images using an inversion algorithm, and the elastic shear modulus of the tissue is subsequently calculated (16,17). It has recently been demonstrated that this technique is as an effective means of detecting and assessing liver fibrosis, by quantifying liver stiffness as a surrogate of fibrosis (5,11,18-20). Excellent correlation between liver stiffness and fibrosis staging has been shown by several groups, with sensitivity and specificity for the detection of any stage of hepatic fibrosis from normal, healthy livers to be 98% and 99%, respectively (19). Unlike ultrasound elastography methods, MRE allows for volumetric assessment over the entire liver. In addition, ultrasound elastography measurements may fail in up to 20% of patients, due to limited penetration in patients with a large body habitus or from the presence of ascites (7). Thus, MRE may provide safe, non-invasive assessment of hepatic fibrosis and cirrhosis suitable for repetitive evaluation, and is easily integrated into routine clinical abdominal MRI exams.

An important gauge of any quantitative biomarker is its repeatability. Although recent studies demonstrate the capability of MRE to quantify liver stiffness as a surrogate of liver fibrosis, the repeatability of this method from visit to visit is currently unknown. Other groups have performed brief repeatability studies (14), a thorough study quantifying the repeatability of MRE has yet to be performed. The purpose of this work is to determine the repeatability of MRE hepatic stiffness measurements using both patients and healthy volunteers.

## Materials and Methods

### Healthy Volunteer Study

The healthy volunteer population consisted of twenty volunteers (seven women, thirteen men) with no known liver disease. Mean age and BMI were 28.1 years (range, 22-40) and 22.9 (range, 17.8-30.3). All imaging was performed after obtaining IRB approval and informed consent. A passive pneumatic driver 19 cm in diameter was positioned on the rib cage and attached to an acoustic waveform generator. A 60 Hz waveform was applied to the driver. The MRE sequence was performed on a clinical 1.5T scanner (HDx TwinSpeed, GE Healthcare, Waukesha, WI) using an eight-channel phased array cardiac coil, which is the

most commonly used coil for liver imaging at our institution. A 2D gradient echo MRE sequence acquired anatomical magnitude and unwrapped phase difference wave images using the following image parameters: TE/TR = 24.2/100.0ms, flip = 30°, BW = ± 31.25 kHz, slice = 10 mm, 256 × 128 matrix, 4 slices, and an asymmetric 75% field of view (FOV) adjusted to fit each volunteer (range, 34-40 cm in the readout dimension).

The exam consisted of a localizer, followed by the MRE sequence. Four axial slices were acquired; slices were chosen from axial and coronal scout images such that one slice was placed through the caudate lobe (Couinaud segment 1), one above the caudate lobe, and two below, all equally spaced by 5cm. The slice through the caudate lobe was chosen such that the aorta and inferior vena cava were clearly visible with the caudate lobe at least partially between the two vascular structures. This prescription provided a reproducible approach for providing standardized image planes that included all segments of the liver in the axial plane. Four magnitude and unwrapped phase images were acquired per slice, each with a 90° phase offset obtained by shifting the motion-encoding gradients. Each slice required two 22-second breath-holds.

Following the first MRE scan, the coil was removed and volunteers were taken off the table. The entire exam was then repeated such that two exams were acquired sequentially on the same day (Exam 1 and Exam 2) with less than five minutes' delay. Two to four weeks later, the entire procedure was repeated for all volunteers (Exam 3 and Exam 4). The same operator performed all volunteer MRE exams.

Phase difference wave images were post-processed using an on-line reconstruction software package (15) that performs phase unwrapping of the generated wave images, magnitude image masking and mathematical inversion of the newly generated wave images (hereto referred as “wave images”) to generate a separate shear stiffness image (kPa) from which liver stiffness measurements could be made for each slice location. All images were transferred to an Advantage Workstation (AW 4.2, GE Healthcare, Waukesha, WI) for stiffness measurements.

### Patient Study

Ten patients (seven women, three men) undergoing a routine liver MRI exam were recruited for this IRB-approved study. Mean patient age and BMI were 50.0 years (range, 21- 68) and 26.6 kg (range, 22.3 - 35.3). MR elastography was performed immediately following the routine clinical exam. If contrast agents were used for the clinical exam, the elastography component was performed prior to contrast agent administration to avoid possible confounding effects from contrast, which might be different for the two MRE studies performed due to contrast wash-out. All imaging was performed after obtaining IRB approval and informed consent.

Pre-existing liver conditions in patients included the following: liver lesions (n = 1), primary sclerosing cholangitis (n = 2), alcoholic cirrhosis (n = 1), cryptogenic cirrhosis (n = 1), fatty liver disease (n = 1), hepatitis B (n = 1), hepatitis C (n = 2), and elevated liver enzymes, etiology unknown (n = 1).

The MRI system, sequence, and parameters were identical to the healthy volunteer study. However, unlike for the healthy volunteers, the exams were not repeated on a separate day because patients were rarely able to return for an additional exam. As with the healthy volunteer study, the same operator performed all patient MRE exams.

## Measurement of Liver Stiffness

Following imaging, two independent readers made measurements of the stiffness images from all exams of all patients and volunteers. A methodized region of interest (ROI) selection for both readers was established such that two readers could take ROI measurements in a reproducible, independent manner. This ROI selection method is described as follows. ROI were chosen from the wave images in a region of waves relatively free of reflections and interference patterns. The largest ROI that could be drawn was placed in such a region, was copied onto the exact location in the stiffness map, and a stiffness measurement was recorded. However, the wave inversion algorithm used is known to correctly handle reflections and interferences (15), and it is also assumed that the liver is a relatively homogeneous tissue with isotropic stiffness. ROI selection covered a range of liver segments, such that ROI were not exclusively recorded in what was perceived as the best part of the image.

One ROI per slice was measured. Average stiffness measurements and standard deviations were weighted by relative ROI area and calculated for all exams, patients, and volunteers for both readers. In order to assess intra-reader variability, both readers re-read Exams 1 and 2 of both patients and healthy volunteers. The time between re-reads was 3 days, in order to avoid memory bias.

## Statistical Analysis

A linear mixed effects (LME) model (21) was used to estimate the variability due to various sources in the data. A LME model is similar to a regression model except that instead of one error term for unexplained variability, a LME model includes an error terms for each identifiable sources of variability. These sources include: subject (separate terms for volunteer and patients), day (exams on different days), exam (replicate exams on the same day), reader (intra-reader), and reading (inter-reader). In addition, LME models include fixed effects (standard regression predictors). This model included fixed effects for the intercept and the type of subject (patient/healthy volunteer). The result is estimates for the fixed effects as in a standard regression model but also estimated standard deviations for each of the sources of variability (the random effect). These standard deviations can be combined to calculate the standard deviation of various quantities of clinical interest. The calculation of the total standard deviation from more than one source is done by squaring the individual standard deviations, adding the resulting variances and taking the square root of the result.

These data, like most biological data, have a constant percent variability rather than a constant absolute variability, i.e., larger numbers are more variable than smaller numbers and the distribution of the data is skewed instead of symmetric. Some accommodation must be made for this in the analysis of the data because, like most parametric statistical methods that assume normally distributed data, LME modeling requires constant variance across all levels of the response variable. In cases like these, the data are transformed to the log scale before analysis. This resulting log data values have constant variance since multiplicative changes become additive in the log scale. Since the log transformation is monotonic, no reordering of the data will occur.

The results of any analyses performed in the log scale must be transformed back to the original scale for interpretation. Means transform back to be equivalent to the geometric mean on the original scale, and standard deviations calculated in the log scale transform back to the original scale as percents. A percent standard deviation has the same interpretation as the coefficients of variation (standard deviation divided by the mean). However, the LME model allows us to compute percent standard deviations in complex

studies using all available data relevant to each standard deviation. The use of the log transformation adds an additional level of complication because all calculations (e.g., finding the standard deviation due to a combination of sources) must be done in the log scale and then transformed back.

This extra layer of complication also arises when computing the variability of the difference between two MRE measurements. The calculation must be done in the log scale and then transformed back to the original scale. The calculation in the log scale is the square root of 2 times the square of the log standard deviation of one MRE measurement. When transforming back to the original scale, a choice must be made as to whether the percent is of the smaller of the two values or the larger. We have chosen to use the smaller of the two values. If the larger is chosen the percent standard deviation will be smaller but when multiplied by the larger of the two values will result in the same change in the original scale.

The computation of 95% confidence requires multiplying the standard deviation in the log scale by  $\pm 1.96$  to construct the 95% confidence interval and then transforming back to the original scale. The value 1.96 is the 97.5 percentile of the normal distribution and is used to construct a 95% confidence interval. In practice the complexity added by the use of the log scale is minimized by keeping all calculations in the log scale until complete. Only then are the values of interest are transformed back to the original scale.

All statistical calculations were carried out in the R statistical programming language (22). The lme4 R library was used to fit the LME models (23).

## Results

Figure 1 displays representative corresponding axial magnitude image masks, stiffness maps, and wave images of Exams 1, 2, 3, and 4 from one healthy volunteer (male, age = 25, BMI = 22.3). This volunteer shows an example of normal stiffness measurements, as defined by Yin *et al* (19) to be below 2.93 kPa. The range of reported stiffness values for both readers across all exams is 1.85 to 2.15 kPa. This volunteer has no known liver disease.

Figure 2A displays representative corresponding axial magnitude image masks, stiffness maps, and wave images of Exams 1 and 2 from one of the patients, who had elevated stiffness values (19,24,25), although no biopsy correlation was performed. The range of reported stiffness values for both readers across all exams is 2.41 to 3.17 kPa. This patient had been previously diagnosed with fatty liver disease.

Lastly, Figure 3A displays representative axial magnitude image masks, stiffness maps, and wave images of Exam 1 and 2 from a second patient, also providing an example of elevated stiffness values (19,24,25). The reported range of stiffness values for both readers across all exams is 4.72 to 4.95 kPa. This patient was previously diagnosed with chronic hepatitis B, and biopsy confirmed Stage 2 fibrosis.

Figure 4 displays measured stiffness values from Exams 1-4 for all twenty volunteers and from Exams 1-2 for all ten patients, reported from the first reading by Reader 1 and Reader 2.

The variability of subjects differed substantially for patients and normal volunteers ( $p < 0.0001$ ) and differed to a much lesser extent by reader ( $p = 0.0004$ ). The standard deviation among all patients was 43.7% and 44.7% for Readers 1 and 2, respectively. This standard deviation is calculated between subjects in the same cohort (patients or volunteers), and since Reader 1 had different ROI than Reader 2, the standard deviations from patient-to-patient is slightly different. Similarly, the standard deviation among all healthy volunteers

was 10.4% and 7.5% for Readers 1 and 2, respectively. The percent standard deviation among patients is higher than that among volunteers because no specific disease was chosen for the patient population and a variety of disease conditions and disease severities was expected.

Table 1 summarizes the calculated component variabilities from the linear mixed effects model, expressed as percent of the mean measured stiffness. These component sources are: physiological changes in the subject from day to day (8.5%), replicate exams on the same subject on the same day (4.2%), inter-reader variability replicate readings by two different readers (1.9%), and intra-reader variability (1.4%). These last two sources were not significantly different from zero ( $p = 0.1479$  and  $0.6731$ , respectively), and contribute little to the total variability, but were included in the model for completeness. The residual standard deviation was 6.5%, and includes all sources of variability not explicitly accounted for in the LME model.

Table 2 summarizes the estimated standard deviations that would occur in specific clinical scenarios. These standard deviations are derived from the component variabilities in Table 1, and are expressed as percent of the measured stiffness. All total variabilities outlined in Table 2 include a contribution from the independent residual variability (6.5%) since other unexplained variabilities that affects all total variabilities, regardless of scenario. Inter- and intra-reader standard deviations are also given in Table 2, and are 6.9% and 6.7%, respectively. Because the additional variability due to reader is considered minor, all further results will include this variability unless stated otherwise. The standard deviation of stiffness measurements on the same subject on different days using the same machine, operator and reader was found to be 11.9%. If two different readers are used, then the standard deviation increases to 12.0%. Thus, the standard deviation for a single MRE measurement is 12.0%.

Of utmost clinical importance is the standard deviation of the difference of two MRE measurements. Using the variability for a single MRE measurement (12%), the standard deviation of the difference in two MRE measurements taken on different days on the same subject is calculated as 17.4%; this value is larger than the 12% standard deviation of one reading because it is the standard deviation of a function of two different MRE readings. A change in an individual's measurement is significant if it is larger than 1.96 times the log standard deviations of the difference which (once transformed back to the original scale is 37.0% of the smaller MRE value.

## Discussion

The purpose of performing this study was to understand the sources of variability in MRE measurements for the detection and monitoring of hepatic stiffness as a quantitative surrogate biomarker of hepatic fibrosis. The total standard deviation of one MRE measurement is 12.0% and the standard deviation of the difference in two measurements is 17.4%. These results quantify the variability in MRE measurements and allow us to determine which differences in measured MRE values are large compared to the underlying variability. For a typical clinical MRE exam changes greater than 37.0% represent meaningful changes in the liver stiffness over longitudinal exams with 95% confidence, related to natural progression of disease or intervention from treatment. Practical interpretation of these total standard deviations imply that if the absolute value of the difference between the values of two stiffness measurements made on two different days is not larger than 37.0% of the smaller of the two measurements, then no significant difference exists between the two stiffness measurements.

The observed biological variability due to day in healthy volunteers (variability due to day) of 8.5% may reflect normal diurnal variability of hepatic stiffness as well as the influence of recent meals. Increased splanchnic blood flow after a meal may affect hepatic stiffness, although this has not been documented. No meal restrictions were made for the volunteers, and although no food or liquid was ingested between scans, volunteers were scanned in a range of fasted and fed states. Per clinical routine, all patients fasted for 2-4 hours prior to scanning. Because this day to day component standard deviation of 8.5% was the largest contribution to the 12.0% within subject standard deviation, further work on the diurnal variation and the influence of fasting on hepatic stiffness is needed.

One limitation of this work is that the estimated variability from day to day is based on normal volunteer data only. All other individual standard deviations were calculated using both patient and volunteer data. Changes greater than 37.0% of the measured stiffness represent meaningful changes in liver stiffness for all individuals undergoing an MRE exam, with this being the best possible estimate for patients given the limitations of scanning patients on a second day. Repeat imaging at a later date (Day 2) for patients was impractical because many of our patients live outside our institution's metropolitan area from remote parts of the state. In addition, many patients were undergoing various treatments.

Unlike the healthy volunteers, it could not be guaranteed that the disease status of the patient livers, and hence the liver stiffness, would be constant over the 2-4 week time period of the study between Day 1 and Day 2. As such, component variability due to day for patients may include additional sources that would confound attempts to determine the true repeatability of MRE. On the same day, it was assumed that the stiffness was constant in patients and volunteers since the scanning session took approximately 10 minutes. Additionally, this work involved repeatability studies of elasticity only. While viscosity has been shown to change as a function of liver fibrosis (14,26), repeatability studies of viscosity will be addressed at a later date.

Finally, the healthy volunteers had no known liver diseases, documented only through oral questioning, without physical exam or serological testing of liver enzymes. Therefore, it was possible that some of the volunteers had cryptogenic disease. Previous work has established a threshold between early fibrosis and truly healthy livers to be 2.93 kPa (19). The average of all four exams for the healthy volunteers showed only one individual to surpass this threshold, although this occurred for only one reader. For this reason, we are confident that all healthy volunteers had normal, healthy liver stiffness values. As expected, the mean liver stiffness values for patients were higher than those for healthy volunteers, with patients having a mean response of  $4.00 \pm 0.51$  kPa, and volunteers having a mean response of  $2.44 \pm 0.06$  kPa.

In conclusion, we have demonstrated and quantified the variability of and sources of variability of MRE for longitudinal hepatic stiffness measurements using the described MRE technique; changes greater than 37.0% of the smaller measured stiffness value represent meaningful changes in longitudinal liver stiffness measurements. However, future developments, such as those in the MRE system or imaging acquisition techniques, may potentially alter the repeatability of this technique. The estimates of component standard deviation determined from this work also provide clues as to what approaches may reduce the total standard deviation in MRE measurements.

## Acknowledgments

We wish to acknowledge Richard Ehman, Robert Grimm, and Meng Yin from the Mayo Clinic, and David Stanley from GE Healthcare.

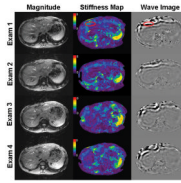
Grant support: 2007 NCRR/NIH Clinical and Translational Science Award

## References

1. Farrell GC, Larter C. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology*. 2006; 43(2, Suppl 1):S99–S112. [PubMed: 16447287]
2. Heidelbaugh JJ, Bruderly M. Cirrhosis and chronic liver failure: Part I. Diagnosis and evaluation. *American Family Physician*. 2006; 74(5):756–762. [PubMed: 16970019]
3. Arif A, Levine RA, Sanderson SO, et al. Regression of fibrosis in chronic hepatitis C after therapy with interferon and ribavirin. *Digestive Diseases and Sciences*. 2003; 48(7):1425–1430. [PubMed: 12870807]
4. Arthur MJ. Reversibility of liver fibrosis and cirrhosis following treatment for hepatitis C. *Gastroenterology*. 2002; 122(5):1525–1528. [PubMed: 11984538]
5. Huwart L, Peeters F, Sinkus R, et al. Liver fibrosis: non-invasive assessment with MR elastography. *NMR in Biomedicine*. 2006; 19(2):173–179. [PubMed: 16521091]
6. Bravo AA, Sheth S, Chopra S. Liver biopsy. *New England Journal of Medicine*. 2001; 344(7):495–500. [PubMed: 11172192]
7. Manning DS, Afdhal NH. Diagnosis and quantitation of fibrosis. *Gastroenterology*. 2008; 134(6): 1670–1681. [PubMed: 18471546]
8. Talkwalkar JA, Yin M, Fidler JL, Sanderson SO, Kamath PS, Ehman RL. Magnetic resonance imaging of hepatic fibrosis: emerging clinical applications. *Hepatology*. 2008; 47(1):332–342. [PubMed: 18161879]
9. Talwalkar JA, Kurtz DM, Schoenleber SJ, West CP, Montori VM. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systemic review and meta-analysis. *Clinical Gastroenterology and Hepatology*. 2007; 5(10):1214–1220. [PubMed: 17916549]
10. Nguyen-Khac E, Chatelain D, Tramier B, et al. Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: prospective comparison with 7 non-invasive laboratory tests. *Alimentary Pharmacology & Therapeutics*. 2008; 28(10):1188–1198. [PubMed: 18705692]
11. Rouviere O, Yin M, Dresner MA, et al. MR elastography of the liver: preliminary results. *Radiology*. 2006; 240(2):440–448. [PubMed: 16864671]
12. Klatt D, Asbach P, Rump J, et al. In vivo determination of hepatic stiffness using steady-state free precession magnetic resonance elastography. *Invest Radiol*. 2006; 41(12):841–848. [PubMed: 17099421]
13. Klatt D, Hamhaber U, Asbach P, Braun J, Sack I. Noninvasive assessment of the rheological behavior of human organs using multifrequency MR elastography: a study of brain and liver viscoelasticity. *Phys Med Biol*. 2007; 52(24):7281–7294. [PubMed: 18065839]
14. Asbach P, Klatt D, Hamhaber U, et al. Assessment of liver viscoelasticity using multifrequency MR elastography. *Magn Reson Med*. 2008; 60(2):373–379. [PubMed: 18666132]
15. Manduca A, Oliphant TE, Dresner MA, et al. Magnetic resonance elastography: non-invasive mapping of tissue elasticity. *Med Image Anal*. 2001; 5(4):237–254. [PubMed: 11731304]
16. Muthupillai R, Ehman RL. Magnetic resonance elastography. *Nature Medicine*. 1996; 2(5):601–603.
17. Muthupillai R, Lomas DJ, Rossman PJ, Greenleaf JF, Manduca A, Ehman RL. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. *Science*. 1995; 269(5232):1854–1857. [PubMed: 7569924]
18. Yin M, Woollard J, Wang X, et al. Quantitative assessment of hepatic fibrosis in an animal model with magnetic resonance elastography. *Magn Reson Med*. 2007; 58(2):346–353. [PubMed: 17654577]
19. Yin M, Talwalkar JA, Glaser KJ, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. *Clin Gastroenterol Hepatol*. 2007; 5(10):1207–1213. e1202. [PubMed: 17916548]
20. Kruse SA, Smith JA, Lawrence AJ, et al. Tissue characterization using magnetic resonance elastography: preliminary results. *Phys Med Biol*. 2000; 45(6):1579–1590. [PubMed: 10870712]
21. Pinheiro, JC.; Bates, DM. *Mixed-effects models in S and S-PLUS*. Springer; New York: 2000.

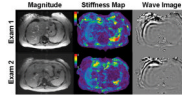


22. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna: 2008.
23. Bates, D.; Maechler, M.; Dai, B. R package version 0.999375-28. 2008. lme4: Linear mixed-effects models using S4 classes.
24. Huwart L, Sempoux C, Salameh N, et al. Liver fibrosis: noninvasive assessment with MR elastography versus aspartate aminotransferase-to-platelet ratio index. *Radiology*. 2007; 245(2): 458–466. [PubMed: 17940304]
25. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. *Gastroenterology*. 2008; 135(1):32–40. [PubMed: 18471441]
26. Sinkus R, Tanter M, Catheline S, Lorenzen J, Kuhl C, Sondermann E, Fink M. Imaging anisotropic and viscous properties of breast tissue by magnetic resonance-elastography. *Magn Reson Med*. 2005; 53(2):372–387. [PubMed: 15678538]



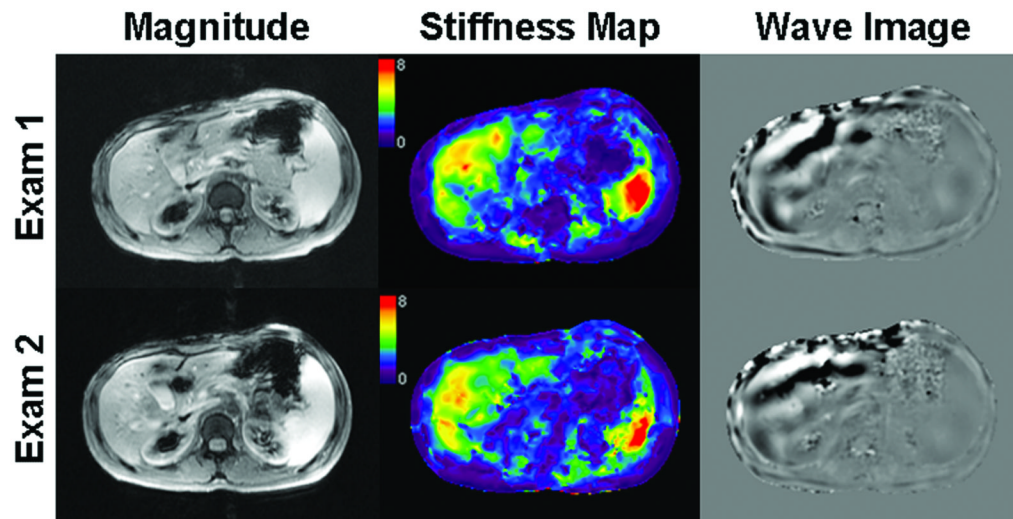
**Figure 1.**

Axial magnitude images, stiffness maps, and wave images of a healthy volunteer for Exams 1-4. The color bar on all stiffness images ranges from 0 to 8 kPa. The red circle indicates a region on the wave image that was chosen as an ROI due to minimal wave interference, and copied onto the same location of the stiffness map to record a stiffness measurement. Reader 1 reported stiffness values of  $1.84 \pm 0.29$ ,  $1.93 \pm 0.51$ ,  $2.01 \pm 0.32$ , and  $1.85 \pm 0.18$  kPa for Exams 1-4 (mean, 1.91 kPa), respectively. Reader 2 reported stiffness values of  $2.13 \pm 0.39$ ,  $1.94 \pm 0.40$ ,  $2.15 \pm 0.33$ , and  $2.09 \pm 0.44$  kPa for Exams 1-4 (mean, 2.08 kPa), respectively. All images were chosen from the same slice location.



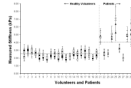
**Figure 2.**

Representative axial magnitude images, stiffness maps, and wave images of a patient with mild elevation of hepatic stiffness for Exams 1 and 2. Reader 1 reported stiffness values of  $2.78 \pm 0.40$  kPa and  $2.41 \pm 0.59$  kPa for Exam 1 and 2 (mean, 2.60 kPa), respectively, and Reader 2 reported stiffness values of  $2.64 \pm 0.34$  kPa and  $3.17 \pm 0.67$  kPa for Exam 1 and 2 (mean, 2.91 kPa), respectively. All images were chosen from the same slice location.



**Figure 3.**

Representative axial magnitude images, stiffness maps, and wave images of a patient with more severely increased liver stiffness for Exams 1 and 2. Reader 1 reported stiffness values of  $4.72 \pm 0.60$  kPa and  $4.87 \pm 0.92$  kPa for Exam 1 and 2 (mean, 4.79 kPa), respectively, and Reader 2 reported stiffness values of  $4.95 \pm 0.79$  kPa and  $4.79 \pm 0.81$  kPa for Exam 1 and 2 (mean, 4.87 kPa), respectively. All images were chosen from the same slice location.

**Figure 4.**

Results of measured stiffness values from the first reading of Reader 1 and 2 for Exams 1-4 of all twenty volunteers and Exams 1-2 of all ten patients. The vertical dashed line separates healthy volunteers (left, squares) from patients (right, triangles), black shapes represent Reader 1 values, and white shapes represent Reader 2 values. The dashed horizontal line through the volunteer data represents the mean stiffness of all twenty volunteers for Exams 1-4 ( $2.44 \pm 0.06$  kPa), and similarly for the patient data ( $4.00 \pm 0.51$  kPa) for both readings. Error bars represent measured standard deviation for each individual's exam (kPa).

**Table 1**

Component Sources of Variability, Expressed as Percent of Measured Stiffness

<b>Short Name</b>	<b>Source of Variability due to:</b>	<b>Standard deviation (%)</b>	<b>P-value *</b>
Days	Different exams on different days	8.5	p = 0.0001
Exams	Different exams on same day	4.2	p = 0.0025
Reader	Multiple readers (inter-reader)	1.9	p = 0.1479
Reading	Multiple readings (intra-reader)	1.4	p = 0.6731
Residual	Residual (unexplained variability)	6.5	N/A

\* P-value for tests of whether the standard deviation is significantly different than zero.

**Table 2**

Total Variability, Expressed as Percent of Measured Stiffness

Clinical Scenario	Total Variability (%)	Component Variabilities* included in total
Between exams on different days, same patient, different readers	12.0	Day, Exam, Reader, Reading, & Residual
Between exams on different days, same patient, same reader	11.9	Day, Exam, Reading, & Residual
Between exams on same day, different readers	8.2	Exam, Reader, Reading & Residual
Inter-observer	6.9	Reader, Reading, & Residual
Intra-observer	6.7	Reading & Residual

\*The component sources of variability are listed in Table 1