Published in final edited form as:

Hepatology. 2017 November ; 66(5): 1474–1485. doi:10.1002/hep.29241.

Magnetic Resonance Elastography Measured Shear Stiffness as a Biomarker of Fibrosis in Pediatric Nonalcoholic Fatty Liver Disease

Jeffrey B. Schwimmer, M.D.[1,2,11], **Cynthia Behling, M.D., Ph.D.**[1], **Jorge Eduardo Angeles, M.D.**[1], **Melissa Paiz, B.S.**[1], **Janis Durelle, B.S.**[1], **Jonathan Africa, M.D.**[1], **Kimberly P. Newton, M.D.**[1,2], **Elizabeth M. Brunt, M.D.**[3], **Joel E. Lavine, M.D.**[4], **Stephanie H. Abrams, M.D., M.S.**[5,6,7], **Prakash Masand, M.D.**[5], **Rajesh Krishnamurthy, M.D.**[5], **Kelvin Wong, Ph.D.**[8], **Richard L. Ehman, M.D.**[9], **Meng Yin, Ph.D.**[9], **Kevin J. Glaser, Ph.D.**[9], **Bogdan Dzyubak, B.S.**[9], **Tanya Wolfson, M.A.**[10], **Anthony C. Gamst, Ph.D.**[10], **Jonathan Hooker, B.S.**[11], **William Haufe, B.S.**[11], **Alexandra Schlein, B.S.**[11], **Gavin Hamilton, Ph.D.**[11], **Michael S. Middleton, M.D., Ph.D.**[11], and **Claude B. Sirlin, M.D.**[11]

[1]Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of California, San Diego School of Medicine, San Diego, California

[2]Department of Gastroenterology, Rady Children's Hospital San Diego, San Diego, California

[3]Department of Pathology and Immunology, Washington University, St. Louis, Missouri

[4]Texas Children's Hospital, Houston, Texas

[5]Columbia University, New York, NY

[6]Baylor College of Medicine, Houston, Texas

[7]Houston Methodist Hospital, Houston, Texas

[8]Miller Children's & Women's Hospital Long Beach, California

[9]Mayo Clinic, Rochester, Minnesota

[10]Computational and Applied Statistics Laboratory (CASL), San Diego Supercomputer Center, University of California, San Diego, California

[11]Liver Imaging Group, Department of Radiology, University of California, San Diego School of Medicine, San Diego, California

Abstract

Magnetic Resonance Elastography (MRE) is promising for non-invasive assessment of fibrosis, a major determinant of outcome in nonalcoholic fatty liver disease (NAFLD). However, data in children are limited. Study aims were to determine accuracy of MRE for detection of fibrosis and advanced fibrosis in children with NAFLD, and to assess agreement between manual and novel automated reading methods. We performed a prospective, multi-center study of 2D-MRE in

Correspondence: Jeffrey B. Schwimmer, M.D., Division of Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of California, San Diego, 3020 Children's Way, MC 5030 San Diego, CA 92123, jschwimmer@ucsd.edu, phone: 858-966-8907, fax: 858-560-6798.

children with NAFLD. MR-elastograms were analyzed manually at 2 reading centers and using a new automated technique. Analysis using each approach was done independently. Correlations were determined between MRE analysis methods and fibrosis stage. Thresholds for classifying the presence of fibrosis and of advanced fibrosis were computed and cross-validated. In 90 children with mean age of 13.1 ± 2.4 years, median hepatic stiffness was 2.35 kPa. Stiffness values derived by each reading center were strongly correlated with each other (r=0.83). All three analyses were significantly correlated with fibrosis stage (center 1, ρ=0.53; center 2, ρ=0.55; and automated analysis, ρ=0.52; p<0.001). Overall cross-validated accuracy for detecting any fibrosis was the same for all methods: 72.2% (61.8 – 81.1). Overall cross-validated accuracy for assessing advanced fibrosis varied by method: 88.9% (80.5% – 94.5%) for center 1, 90.0% (81.9% – 95.3%) for center 2, and 86.7% (77.9 – 92.9) for automated analysis.

Conclusions—2D-MRE can estimate hepatic stiffness in children with NAFLD. Further refinement and validation of automated analysis techniques will be an important step in standardizing MRE. How to best integrate MRE into clinical protocols for the assessment of NAFLD in children will need prospective evaluation.

Keywords

liver biopsy; obesity; pediatrics; fibrosis; NAFLD; NASH; MRI

Nonalcoholic fatty liver disease (NAFLD) is the most common cause of chronic liver disease in children[1]. The diagnosis of NAFLD requires that 5% or more hepatocytes have macrovesicular steatosis, and that other liver diseases and/or clinical conditions that may cause steatosis are excluded[2]. Approximately 25% of children with NAFLD have a progressive sub-phenotype known as nonalcoholic steatohepatitis (NASH)[3]. Some children with NAFLD will develop cirrhosis and end-stage liver disease[4–6]. NAFLD may also progress to hepatocellular carcinoma. In adults, the presence and severity of fibrosis are important predictors of the long-term risk for cirrhosis, hepatocellular carcinoma, and liver-related mortality [7, 8]. Natural history studies to determine significance of presence and severity of fibrosis on long-term outcomes in children are lacking[9]. Fibrosis has been reported in 40 to 90% [10–16], and advanced fibrosis has been reported in 6 to 34% of children with NAFLD[11, 14–17]. Identification and staging of liver fibrosis in children with NAFLD requires histologic evaluation of liver histology. The ability to determine the risk for progressive disease without an interventional procedure such as liver biopsy remains both a clinical and research challenge[18]. Hence, there is an urgent need for a reliable non-invasive biomarker of fibrosis in children with NAFLD.

Currently, the leading non-invasive imaging methods for assessing liver fibrosis are elastographic techniques that measure mechanical properties such as “stiffness” and stiffness-related parameters that are believed to be related to fibrosis. Existing techniques are either ultrasound-based (vibration-controlled transient elastography and point shear wave elastography) or magnetic resonance-based (magnetic resonance elastography (MRE)). Ultrasound-based techniques may be limited by anatomical factors such as obesity or narrow rib spacing. In adults with chronic liver disease, MRE has shown higher accuracy than ultrasound-based elastography techniques for predicting fibrosis, particularly in adults with

NAFLD [19–22]. However, data on MRE in children are extremely limited, and there have been no studies that focused on the performance of MRE in pediatric NAFLD.

MRI Assessment Guiding NAFLD Evaluation and Treatment (MAGNET) was a multi-center study developed as an ancillary study to the NIDDK NASH Clinical Research Network (CRN). The primary aim of MAGNET was to develop magnetic resonance (MR)-based biomarkers in children with NAFLD to provide high accuracy for non-invasive assessment of the histological lesions of NAFLD. The current study is the first report from MAGNET and had the following aims in children with known or suspected NAFLD:

1. To assess inter-reader agreement, and correlation between reading centers and methods for MRE-estimated hepatic stiffness.
2. To assess correlation between MRE-estimated hepatic stiffness and histologic fibrosis stage.
3. To determine accuracy of MRE-estimated hepatic stiffness for detection of any fibrosis (i.e., for differentiating stage ≥ 1 from stage 0) and of advanced fibrosis (i.e., for differentiating stage ≥ 3 from stage < 3 fibrosis).

EXPERIMENTAL PROCEDURES

Study Cohort

MAGNET was a multi-site, prospective, observational, cross-sectional ancillary study to the NASH CRN enrolling children with known or suspected NAFLD who had or were scheduled to have a liver biopsy. The determination to perform liver biopsy was a clinical decision and not part of this study. All children were age ≥ 8 and < 18 years and were enrolled in the NASH CRN at UC San Diego or Texas Children's Hospital. Exclusion criteria were extreme claustrophobia, pregnancy, weight exceeding scanner table limit, girth exceeding scanner bore diameter, or other contraindications to MR, e.g. metal in the eyes, implanted electronic devices, aneurysm clips, pacemaker, and cochlear implants. The parent(s) or legal guardian of all subjects provided written informed consent. Assent was obtained from all children. The protocol was approved by the institutional review boards of the University of California, San Diego, Rady Children's Hospital San Diego, Texas Children's Hospital, and Baylor College of Medicine. For this analysis, we included only those children enrolled in MAGNET who had MRE attempted within six months of liver biopsy.

Clinical Data Collection

Clinical data were obtained for each participant at a single fasting intake visit conducted at the Clinical and Translational Research Institute at the University of California, San Diego Medical Center or the Texas Children's Liver Center. Age and gender were recorded. Height was measured to the nearest tenth of a centimeter on a clinical stadiometer. Weight was measured on a clinical scale to the nearest tenth of a kilogram. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Phlebotomy was performed after a 12-hour overnight fast. Assays performed at local clinical laboratories

included platelet count (cells/mL), alanine aminotransferase ALT (U/L), aspartate aminotransferase (AST) (U/L), and gamma-glutamyltransferase (GGT) (U/L).

Liver Tissue Evaluation

Liver tissue slides were stained in a central laboratory with hematoxylin-eosin and Masson's trichrome stains. The slides were reviewed and scored by members of the NASH CRN Pathology Committee, a panel of 9 expert liver pathologists, who interpreted each feature of NAFLD in consensus. The Pathology Committee was blinded to any clinical or demographic information. A diagnosis was assigned as 1) not NAFLD, 2) NAFLD, but not NASH, 3) Borderline Zone 1 NASH, 4) Borderline Zone 3 NASH, or 5) Definite NASH based on the aggregate presence, zonal location and degree of the individual histologic features of fatty liver disease. Although no single histologic feature is considered diagnostic of NASH, a typical set of minimum criteria includes zone 3 macrovesicular steatosis (more than 5%), lobular inflammation and hepatocyte injury as manifest by ballooning in a pattern that is recognizable. Borderline cases demonstrated either a lesser degree of one or more findings, or an alteration of the pattern of NASH. A commonly found "pediatric" pattern of NAFLD is the borderline zone 1 (portal/periportal) pattern[5]. In addition to the diagnosis, an ordinal score was given for amounts of various histologic features, including, but not limited to, macrovesicular steatosis (0–3), lobular inflammation (0–2) and hepatocellular ballooning (0–2). The NAFLD Activity Score was calculated as the combination of steatosis, lobular inflammation, and hepatocellular ballooning on a scale of 0 to 8. Fibrosis was staged as follows: a) none, b) stage 1a – mild zone 3 perisinusoidal fibrosis requiring trichrome stain; c) stage 1b – moderate zone 3 perisinusoidal fibrosis not requiring trichrome stain; d) stage 1c – zone 1 portal/periportal fibrosis only; e) stage 2 –zone 3 perisinusoidal fibrosis and zone 1 periportal fibrosis; f) stage 3 – bridging fibrosis; and g) stage 4 – cirrhosis.

Magnetic Resonance Elastography

MRE Acquisition—MRE was performed on a 3T MR scanner (Signa HDxt 3.0T (UCSD) or Signa Excite 3.0T (Methodist Hospital, Houston TX for subjects at Texas Children's Hospital/Baylor College of Medicine) (GE Healthcare, Waukesha, WI, USA) using a previously described technique[23, 24] Children were instructed to fast for at least four hours before the MRE exam to reduce possible confounding physiological effects, and they were positioned supine and feet first with a torso phase-array surface coil centered at the level of the liver. A dielectric pad was placed anteriorly between the body wall and the surface coil. A passive acoustic driver was secured to the body wall anterior to the liver with an elastic band and continuous vibrations at 60 Hz were applied during data acquisition. The gradient-recalled-echo (GRE) MRE sequence had the following parameters: 50ms repetition time, 20.2-ms echo time, 30-degree flip angle, 256×64 matrix, 36×36 cm to 48×48 cm field of view adjusted case by case by the technologist to accommodate body habitus, one-signal average, ±31.25 kHz receiver bandwidth, Array Spatial Sensitivity Encoding Technique with parallel imaging acceleration factor 2, and motion sensitization along the z-direction (superior to inferior) and four phase offsets. MRE data was acquired at four separate axial slices (10 mm thick, 10 mm interslice gap) at the widest transverse part of the liver, each with a 16-second breath-hold, resulting in total scan time of roughly two minutes including recovery periods between breath-holds.

MRE Processing—Using a previously described two-dimensional (2D) direct inversion algorithm[23], the MR scanner computer automatically produced hepatic stiffness maps (called elastograms) from the acquired data; these maps display the spatial distribution of a hepatic stiffness parameter called the magnitude of the complex shear modulus. This processing also provided a "confidence map" for each elastogram with values ranging from 0 to 1, with 1 representing highly reliable hepatic stiffness estimates from the model. To obtain reliable hepatic stiffness values, a confidence mask was created by keeping only pixels with confidence values greater than 0.95. Wave images, elastograms, and confidence masks were generated automatically by the MR scanner computer; processing time was usually less than two minutes. Examinations with incomplete, unanalyzable, or unreliable data were recorded.

MRE Reading Centers

MRE exams were analyzed centrally in two reading centers, one located at the University of California, San Diego (Reading Center 1) and one at Mayo Clinic, Rochester, MN (Reading Center 2). As discussed below, Reading Center 1 performed only manual MRE analyses, while Reading Center 2 performed both manual and automated analyses. All analyses were done independently and blinded to all other data, including the results of the other analyses and liver histology.

Manual Analysis—MRE-generated elastograms, magnitude images, wave images, and confidence masks were transferred offline for manual analysis. Using a custom software package (MRE Quant, Mayo Clinic; Rochester, MN), image analysts from Reading Center 1 and 2 independently drew regions of interest (ROIs) on magnitude images, so as to include only liver parenchyma, avoiding edges of liver, large blood vessels and bile ducts. ROIs were then co-localized to corresponding wave images and further modified to include only areas having clearly visualized, parallel propagating waves. Hepatic stiffness measurements were calculated based on averaged values within the intersection of manually drawn ROIs and confidence masks.

Automated Analysis—MRE images were also analyzed using an automated liver elasticity estimation algorithm (ALEC) developed at the Mayo Clinic. The algorithm uses known relative positions and intensities of the liver, background, and other tissues to calculate a liver mask from MRE magnitude images. Areas with sharp anatomical features in magnitude images, phase images, or elastograms are then removed from the ROI to avoid incorporating blood vessels, or partial volumes of blood vessels into the hepatic stiffness calculation. Finally, a confidence mask, such as was used in manual analysis but with a less strict cutoff of 0.9, was applied to remove areas with noisy wave data, and mean hepatic stiffness was calculated from the resulting ROIs. Analyses were fully automatic; no manual input or screening was required.

Data Analysis

Reasons for ineligibility of the excluded subjects were summarized. Demographic, clinical and biopsy measures of the subjects included in the study were summarized descriptively. MR examinations with incomplete, unanalyzable, or unreliable MRE data were excluded.

Bland-Altman analyses were performed pairwise for all three MRE analyses: done at Reading Center 1, Reading Center 2, and automated. Bland-Altman bias (mean of the paired differences) and its significance, standard deviation (SD) of the differences, and the limits of agreement (LOA = bias+/− 1.96*SD) were computed for each pair of methods. Spearman's rank correlations were computed pairwise for the three MRE analyses. The choice of Spearman's correlation was guided by the outliers in the data, which would have strongly influenced the results of a parametric Pearson's correlation or linear regression analysis (whereas Spearman's correlation is insensitive to outliers). Pairwise correlations were compared using Williams' test for dependent correlations sharing one variable.[25]

Spearman's correlations were computed between fibrosis stages and each of the MRE methods. Receiver operating characteristic (ROC) curves for classifying no fibrosis (stage 0) vs. any fibrosis (stage ≥ 1), and no or mild fibrosis (stage ≤ 2) vs. advanced fibrosis (stage ≥ 3), were computed for each of the three MRE methods. Area under the ROC curve (AUROC) was computed for each MRE method and each classification. For each method and each classification, the classification threshold was chosen that provided highest sensitivity at a minimum of 90% specificity. The following diagnostic parameters then were calculated at that threshold: sensitivity, specificity, positive predictive value, negative predictive value, and total accuracy. Stratified 6-fold cross-validation was applied to obtain more accurate estimates of the performance of the three methods. Bootstrap-based tests were used to determine 95% confidence intervals around the AUROCs and all diagnostic parameters. AUROCs were compared pairwise using bootstrap-based tests.

RESULTS

Study Population

A flowchart is shown in Figure 1 detailing the recruitment, inclusion, and exclusion of children for this study. We screened 121 children and found seven to be ineligible for MAGNET. Reasons for ineligibility were: weight too high (n = 2), declined MRI (n = 2), metal in the eye (n = 1), and parent felt child not able to cooperate with MRI (n = 2). The remaining 114 children were enrolled in MAGNET. Of these, seven were scanned during a period in which the MRE software was temporarily not operational after a system upgrade; these children were excluded from the current analysis since MRE was not attempted. In the 107 children in whom MRE was attempted, MRE was incomplete in 5 (5%) due to girth exceeding scanner bore (n = 1), claustrophobia (n = 1), difficulty with breath holding (n = 1), or MRE passive driver reported uncomfortable (n = 1). In one child, data was not analyzable due to such low signal to noise ratio that no wave motion could be detected. Twelve additional children had unreliable MRE data due to severe breathing artifacts (n = 5), other severe imaging artifacts (n = 2), and poor signal to noise ratio causing the ROI area to be too small for reliable measurement (n = 5). Thus, the overall rate of failing to obtain reliable MRE data was 16% (17/107). There was no significant difference in age, sex, race, ethnicity, weight, or BMI z-socre for children with reliable MRE data vs those with none or unreliable MRE data (supplementary Table 1). Children with unreliable MRE data did have significantly more advanced fibrosis than children with reliable MRE data (supplementary

Table 1). For the 90 children with reliable MRE data, demographics and clinical features are shown in Table 1. Mean age was 13.1 ± 2.4 years, and 73% of participants were boys.

Liver biopsy results of the 90 children showed that three had steatosis < 5% and were determined to not have NAFLD. Of the 87 children with NAFLD, the overall diagnostic determinations were: NAFLD, but not NASH 44% (38/87), Borderline Zone 3 pattern 17% (15/87), Borderline Zone 1 pattern 20% (17/87), and Definite NASH 20% (17/87). The distribution of fibrosis severity was stage 0: 60% (54/90), stage 1: 27% (24/90), stage 2: 7% (6/90), stage 3: 6% (5/90), and stage 4: 1% (1/90).

Inter-reader Comparison of MRE Readings

The median 2D MRE hepatic stiffness value was 2.35 kPa (range 1.81 to 9.67) for Reading Center 1 manual analysis, 2.35 kPa (range 1.85 to 10.46) for Reading Center 2 manual analysis, and 2.31 kPa (range 1.74 to 9.65) for the automated analysis. Bland-Altman plots and scatterplots for each pair of methods are presented in Figure 2. The stiffness values obtained from the manual analysis at Reading Centers 1 and 2 were similar, on average (Figure 2c, bias=0.001, p=0.9606). However, the stiffness values obtained from the automated analysis were lower, on average, than the stiffness values obtained at Reading Centers 1 and 2(Figure 2a and 2e, bias: 0.074 and 0.073 kPa, p: 0.0006 and 0.0335, respectively). As shown in Figure 2d, the 2D MRE stiffness values obtained from the manual analysis at Reading Centers 1 and 2 were strongly correlated with each other (r=0.83), but not identical. Each manual analysis was also compared against the automated analysis. The correlation between the Reading Center 1 manual analysis and the automated analysis (Figure 2b) (ρ=0.90) was significantly higher than the correlation between the Reading Center 1 and Reading Center 2 manual analyses (ρ=0.83) or the correlation between the Reading Center 2 manual analysis and the automated analysis (Figure 2f) (ρ=0.79). (Williams' test p=0.0121 and p<0.001, respectively). There was no significant difference in correlations between Reading Center 1 and Reading Center 2, and between Reading Center 1 and Automated Reading (p=0.0116).

MRE vs Liver Histopathology

A representative example of a MRE-generated elastogram is shown for each fibrosis stage grade in Figure 3. All three MRE analyses were significantly correlated overall with stages of fibrosis (ρ=0.53 for Reading Center 1 manual analysis, ρ = 0.55 for Reading Center 2 manual analysis and ρ = 0.52 for automated analysis, all p<0.001). Figure 4 shows a dot plot of shear stiffness for each participant sorted by fibrosis stage and reading approach. Of note, there was a significant correlation between ALT values and MRE stiffness values: rho = 0.35 (p = 0.0008) for Reading Center 1, rho = 0.46 (p < 0.0001) for Reading Center 2, rho = 0.31 (p = 0.0033) for automated analysis. However, there was not a consistently significant correlation between NAS and MRE stiffness values: rho = 0.13 (p = 0.2337) for Reading Center 1, rho = 0.27 (p = 0.0101) for Reading Center 2, rho = 0.05 (p = 0.6375) for automated analysis.

Figure 5A shows ROC curves for distinguishing children without liver fibrosis (stage 0) from those with any liver fibrosis (stage ≥1). The AUROC for classifying presence of any

fibrosis was not significantly different between Reading Center 1 manual analysis and Reading Center 2 manual analysis (0.77 vs. 0.79; bootstrap p>0.05). The best classification threshold at ≥ 90% specificity was 2.77 kPa for Reading Center 1 manual analysis and 2.70 kPa for Reading Center 2 manual analysis. The cross-validated diagnostic performance is shown in Table 2. Sensitivities, negative predictive values, and overall accuracies tended to be low or modest at ≥90% specificity; since the thresholds were based on ≥90% specificity, positive predictive value was higher. The cross-validated overall accuracy was 72.2% (95% CI 61.8% – 81.1%) for both Reading Center 1 manual analysis and Reading Center 2 manual analysis.

Figure 5B shows ROC curves for the identification of advanced fibrosis (stage ≥3) in children with NAFLD. After the Bonferroni correction, the AUROC for classifying advanced fibrosis was borderline significantly higher for Reading Center 1 manual analysis than for Reading Center 2 manual analysis (0.93 vs. 0.88; family-wise bootstrap p=0.05). There were no other significant differences between AUROCs. The best classification threshold at ≥ 90% specificity was 3.05 kPa for Reading Center 1 manual analysis and 3.03 kPa for Reading Center 2 manual analysis. The cross-validated diagnostic performance is shown in Table 2. As for the detection of any fibrosis, sensitivities tended to be low at ≥90% specificity, and, given the low number of subjects with advanced fibrosis, the positive predictive values tended to be quite low. Since the thresholds were based on ≥90% specificity, specificity, overall accuracy and negative predictive value were high. The cross-validated overall accuracy was 88.9% (95% CI, 80.5% – 94.5%) for Reading Center 1 manual analysis and 90.0% (95% CI, 81.9% – 95.3%) for Reading Center 2 manual analysis.

DISCUSSION

In MAGNET, we performed a dual-center study of MRE in children with NAFLD and compared MRE-measured hepatic stiffness to liver histology fibrosis stage. Notwithstanding a relatively high technical failure rate, we demonstrated that MRE could be done in most children with NAFLD, although it was not universally tolerated as currently performed. There was a detrimental effect of inadequate breath holding even in those children who were able to complete their exams. For those with reliable MRE data, the inter-reader agreement for liver stiffness was strong. We also demonstrated the potential for automated analysis of MRE data. Overall, the correlation between MRE estimated hepatic shear stiffness and liver fibrosis stage was good. Notably, MRE did not perform as well in children with NAFLD compared to previously reported studies in adults[26, 27].

These data substantially advance the knowledge base regarding MRE in children. In 2012, Binkovitz et al reported the initial experience with pediatric MRE—a case series of seven children with varied liver disease to demonstrate the potential feasibility of MRE in children[28]. That same year in a company report for consumers, Siegel et al reported that in six children with cystic fibrosis and four healthy controls, a cutoff of > 3.38 kPa was both 100% sensitive and specific for the detection of cirrhosis[29]. In 2014, Xanthakos et al presented a case series of 35 children with eight different chronic liver diseases and noted that a cutoff of 2.71 kPa had a sensitivity of 88% and specificity of 85% in the separation of fibrosis stages 0–1 from 2–4[30]. Anecdotally, we have seen children referred to our Fatty

Liver Clinic who have already had liver MRE performed at outside imaging centers along with interpretations about whether or not they have fibrosis based upon cutoffs in clinical use for adults. The results of this study suggest that MRE cutoffs derived from adult studies may not be directly applicable to children with NAFLD. Therefore, clinicians should be careful in interpreting MRE results in children. Further studies are needed to develop and confirm standardized pediatric-based cut-offs that can transcend diseases and institutions.

Few studies have evaluated inter-reader variability of MRE-estimated hepatic stiffness. Those that have done so were only in adults and mostly in those with viral hepatitis rather than NAFLD. The intraclass correlation coefficient (ICC) ranged from 0.74 to 0.99[31–34]. In MAGNET we found a correlation of 0.83. It is the variability in the values between 2.0 and 3.0 kPa that is most critical to the accurate classification of fibrosis stage in children with NAFLD. As a biomarker for patients with NAFLD, MRE in children was similar to adults in the ability to detect any fibrosis but did not perform as well for detection of advanced fibrosis. In adults, large single-center studies have been done in patients with NAFLD. Kim et al reported that in 142 adults with NAFLD, a cut-off of 4.15 kPa had a sensitivity of 0.85 and specificity of 0.93 for identifying advanced fibrosis[27]. Loomba et al reported in 117 adults with NAFLD, that for identifying any fibrosis, a cutoff of 3.02 kPa had a sensitivity of 0.55 and specificity of 0.91, and that for identifying advanced fibrosis a cutoff of 3.64 had a sensitivity of 0.86 and specificity of 0.91[26]. Finally in a large meta-analysis of 697 adults with heterogeneous liver disease including NAFLD, Singh et al reported that for identifying any fibrosis, a cutoff of 3.45 kPa had a sensitivity of 0.73 and specificity of 0.79, and that for identifying advanced fibrosis a cutoff of 4.11 had a sensitivity of 0.85 and specificity of 0.85[35].

The histologic features of NAFLD in children can differ from those seen in adults with NAFLD. Children tend to have higher steatosis grades, which conceivably could affect hepatic stiffness values since fat tends to be softer than lean tissue. In children, fibrosis is typically portal-based, whereas in adults fibrosis classically begins in the centrilobular region. It is not yet known how the locations and patterns of fibrosis affect hepatic stiffness. It is believed that fibrosis produces a lattice-like framework of struts that impart parenchymal rigidity; it is plausible that the intraparenchymal arrangement of the framework, in addition to the composition of the framework could alter observed hepatic stiffness.

The current dual-center study was notable for the large sample size of well-characterized children with NAFLD including liver histology evaluated in a standardized fashion by expert liver pathologists blinded to the clinical and imaging details. In addition, the MRE images were analyzed at two experienced centers to enable evaluation of inter-reader agreement. The study also incorporated a novel fully automated analysis method that minimizes subjectivity and operator dependence from the analysis; with further validation, this new method may provide a mechanism for objective high-throughput centralized analysis clinical trials. The major limitation of this study was the large number of children who were unable to undergo MRE, or for whom MRE did not produce an interpretable elastogram. In addition, spectrum bias such as the small number of children with stage 3–4 fibrosis, may have introduced some error into the reported thresholds and diagnostic performance for the

detection of advanced fibrosis. This was further complicated by the technical failure rate in children with stage 3–4 fibrosis. There were not any obvious clinical differences such as age contributing to differences in technical failure. Nor are there any physics or engineering reasons to explain these differences. However, we cannot exclude that there are underlying differences in some children with greater fibrosis that made breath holding more difficult. Because the total sample size with advanced fibrosis was still relatively small, we caution against over-interpreting the significance of this occurrence. We do note that future studies of MRE in children should prospectively evaluate factors that are related to technical success or failure. Furthermore, the variable time interval between liver biopsy and MRE studies may have introduced an unknown degree of error. However, the time interval was relatively small for progression or regression of fibrosis.

In MAGNET, we demonstrated that MRE could be used to estimate hepatic stiffness in children with NAFLD. However, the results of this study did not show the degree of diagnostic performance reported in adult-based studies. This may be due, in part, to technical difficulties in children, which might be addressed with improved pediatric MRE sequences that do not require breath holding. Further technical development of MRE and other non-invasive quantitative imaging biomarkers of fibrosis in children with NAFLD remains a major need. The further refinement and validation of automated analysis techniques will be an important step in standardizing MRE. How to best integrate MRE into clinical and research protocols for the assessment of NAFLD in children will need prospective evaluation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This work was supported in part by DK088925, DK088925-02S1, DK090350, EB001981, and UL1RR031980 from the NCRR for the Clinical and Translational Research Institute at UCSD. The funders did not participate in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Abbreviations

MRE	magnetic resonance elastography
MAGNET	MRI Assessment Guiding NAFLD Evaluation and Treatment
NAFLD	nonalcoholic fatty liver disease
NASH	nonalcoholic steatohepatitis

References

1. Schwimmer JB, Deutsch R, Kahen T, Lavine JE, Stanley C, Behling C. Prevalence of fatty liver in children and adolescents. Pediatrics. 2006; 118:1388–1393. [PubMed: 17015527]

2. Lindback SM, Gabbert C, Johnson BL, Smorodinsky E, Sirlin CB, Garcia N, Pardee PE, et al. Pediatric nonalcoholic fatty liver disease: a comprehensive review. Adv Pediatr. 2010; 57:85–140. [PubMed: 21056736]

3. Pardee PE, Lavine JE, Schwimmer JB. Diagnosis and treatment of pediatric nonalcoholic steatohepatitis and the implications for bariatric surgery. Semin Pediatr Surg. 2009; 18:144–151. [PubMed: 19573756]

4. Rubinstein E, Lavine JE, Schwimmer JB. Hepatic, cardiovascular, and endocrine outcomes of the histological subphenotypes of nonalcoholic fatty liver disease. Semin Liver Dis. 2008; 28:380–385. [PubMed: 18956294]

5. Schwimmer JB, Behling C, Newbury R, Deutsch R, Nievergelt C, Schork NJ, Lavine JE. Histopathology of pediatric nonalcoholic fatty liver disease. Hepatology. 2005; 42:641–649. [PubMed: 16116629]

6. Molleston JP, White F, Teckman J, Fitzgerald JF. Obese children with steatohepatitis can develop cirrhosis in childhood. Am J Gastroenterol. 2002; 97:2460–2462. [PubMed: 12358273]

7. Soderberg C, Stal P, Askling J, Glaumann H, Lindberg G, Marmur J, Hultcrantz R. Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology. 2010; 51:595–602. [PubMed: 20014114]

8. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, Mills PR, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2015; 149:389–397.e310. [PubMed: 25935633]

9. Goyal NP, Schwimmer JB. The Progression and Natural History of Pediatric Nonalcoholic Fatty Liver Disease. Clin Liver Dis. 2016; 20:325–338. [PubMed: 27063272]

10. Nobili V, Alisi A, Vania A, Tiribelli C, Pietrobattista A, Bedogni G. The pediatric NAFLD fibrosis index: a predictor of liver fibrosis in children with non-alcoholic fatty liver disease. BMC Med. 2009; 7:21. [PubMed: 19409076]

11. Schwimmer JB, Zepeda A, Newton KP, Xanthakos SA, Behling C, Hallinan EK, Donithan M, et al. Longitudinal assessment of high blood pressure in children with nonalcoholic Fatty liver disease. PLoS One. 2014; 9:e112569. [PubMed: 25419656]

12. Patton HM, Lavine JE, Van Natta ML, Schwimmer JB, Kleiner D, Molleston J. Nonalcoholic Steatohepatitis Clinical Research, N. Clinical correlates of histopathology in pediatric nonalcoholic steatohepatitis. Gastroenterology. 2008; 135:1961–1971. e1962. [PubMed: 19013463]

13. Schwimmer JB, Middleton MS, Behling C, Newton KP, Awai HI, Paiz MN, Lam J, et al. Magnetic resonance imaging and liver histology as biomarkers of hepatic steatosis in children with nonalcoholic fatty liver disease. Hepatology. 2014

14. Ko JS, Yoon JM, Yang HR, Myung JK, Kim H, Kang GH, Cheon JE, et al. Clinical and histological features of nonalcoholic fatty liver disease in children. Dig Dis Sci. 2009; 54:2225–2230. [PubMed: 19697129]

15. Eng K, Lopez R, Liccardo D, Nobili V, Alkhouri N. A non-invasive prediction model for non-alcoholic steatohepatitis in paediatric patients with non-alcoholic fatty liver disease. Dig Liver Dis. 2014; 46:1008–1013. [PubMed: 25106814]

16. Carter-Kent C, Yerian LM, Brunt EM, Angulo P, Kohli R, Ling SC, Xanthakos SA, et al. Nonalcoholic steatohepatitis in children: a multicenter clinicopathological study. Hepatology. 2009; 50:1113–1120. [PubMed: 19637190]

17. Schwimmer JB, Middleton MS, Behling C, Newton KP, Awai HI, Paiz MN, Lam J, et al. Magnetic resonance imaging and liver histology as biomarkers of hepatic steatosis in children with nonalcoholic fatty liver disease. Hepatology. 2015; 61:1887–1895. [PubMed: 25529941]

18. Schwimmer JB. Clinical advances in pediatric nonalcoholic fatty liver disease. Hepatology. 2016; 63:1718–1725. [PubMed: 27100147]

19. Huwart L, Sempoux C, Vicaut E, Salameh N, Annet L, Danse E, Peeters F, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology. 2008; 135:32–40. [PubMed: 18471441]

20. Ichikawa S, Motosugi U, Morisaka H, Sano K, Ichikawa T, Tatsumi A, Enomoto N, et al. Comparison of the diagnostic accuracies of magnetic resonance elastography and transient elastography for hepatic fibrosis. Magn Reson Imaging. 2015; 33:26–30. [PubMed: 25308096]
21. Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, Fujita K, et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. Gastroenterology. 2016; 150:626–637.e627. [PubMed: 26677985]
22. Cui J, Heba E, Hernandez C, Haufe W, Hooker J, Andre MP, Valasek MA, et al. Magnetic resonance elastography is superior to acoustic radiation force impulse for the Diagnosis of fibrosis in patients with biopsy-proven nonalcoholic fatty liver disease: A prospective study. Hepatology. 2016; 63:453–461. [PubMed: 26560734]
23. Yin M, Talwalkar JA, Glaser KJ, Manduca A, Grimm RC, Rossman PJ, Fidler JL, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. Clin Gastroenterol Hepatol. 2007; 5:1207–1213. e1202. [PubMed: 17916548]
24. Venkatesh SK, Yin M, Ehman RL. Magnetic Resonance Elastography of Liver: Technique, Analysis, and clinical applications. Journal of Magnetic Resonance Imaning. 2013; 37:544–555.
25. Williams EJ. The comparison of regression variables. Journal of the Royal Statistical Society. 1959:396–399. Series B.
26. Loomba R, Wolfson T, Ang B, Hooker J, Behling C, Peterson M, Valasek M, et al. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: A prospective study. Hepatology. 2014; 60:1920–1928. [PubMed: 25103310]
27. Kim D, Kim WR, Talwalkar JA, Kim HJ, Ehman RL. Advanced Fibrosis in Nonalcoholic Fatty Liver Disease: Noninvasice Assessment with MR Elastography. Radiology. 2013; 268:411–419. [PubMed: 23564711]
28. Binkovitz LA, El-Youssef M, Glaser KJ, Yin M, Binkovitz AK, Ehman RL. Pediatric MR elastography of hepatic fibrosis: principles, technique and early clinical experience. Pediatr Radiol. 2012; 42:402–409. [PubMed: 22120578]
29. Siegel MJ, Priatna A, Bolster B, Kotyk JJ, Pediatric MR. Elastography of the Liver. Clinical Pediatric Imaging. 2012:108–111.
30. Xanthakos SA, Podberesky DJ, Serai SD, Miles L, King EC, Balistreri WF, Kohli R. Use of magnetic resonance elastography to assess hepatic fibrosis in children with chronic liver disease. J Pediatr. 2014; 164:186–188. [PubMed: 24064151]
31. Rustogi R, Horowitz J, Harmath C, Wang Y, Chalian H, Ganger DR, Chen ZE, et al. Accuracy of MR elastography and anatomic MR imaging features in the diagnosis of severe hepatic fibrosis and cirrhosis. J Magn Reson Imaging. 2012; 35:1356–1364. [PubMed: 22246952]
32. Lee Y, Lee JM, Lee JE, Lee KB, Lee ES, Yoon JH, Yu MH, et al. 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. 2014; 39:326–331. [PubMed: 23589232]
33. Runge JH, Bohte AE, Verheij J, Terpstra V, Nederveen AJ, van Nieuwkerk KM, de Knegt RJ, et al. Comparison of interobserver agreement of magnetic resonance elastography with histopathological staging of liver fibrosis. Abdom Imaging. 2014; 39:283–290. [PubMed: 24366108]
34. Lee DH, Lee JM, Han JK, Choi BI. MR elastography of healthy liver parenchyma: Normal value and reliability of the liver stiffness value measurement. J Magn Reson Imaging. 2013; 38:1215–1223. [PubMed: 23281116]
35. Singh S, Venkatesh SK, Wang Z, Miller FH, Motosugi U, Low RN, Hassanein T, et al. Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. Clin Gastroenterol Hepatol. 2015; 13:440–451.e446. [PubMed: 25305349]

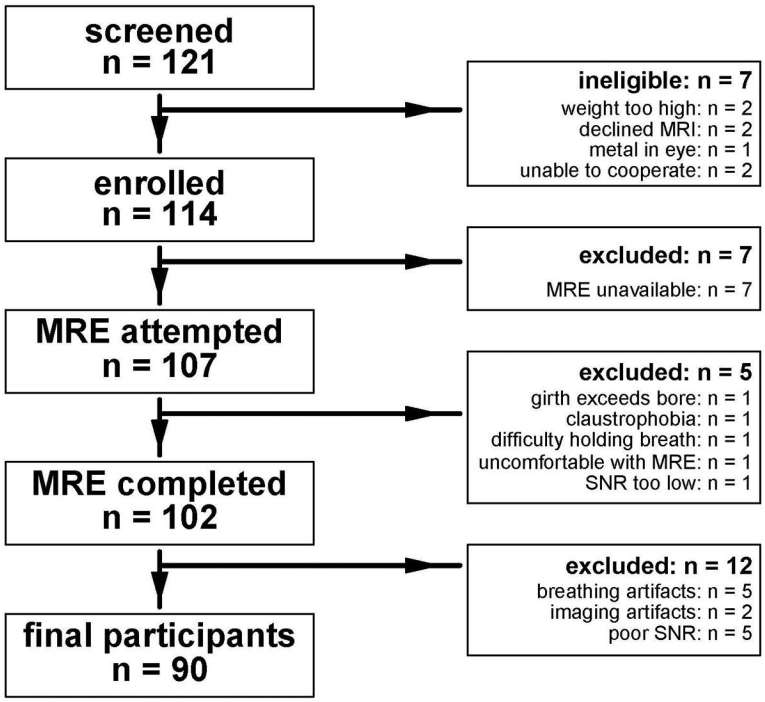

Figure 1.
Flowchart shows the application of study inclusion and exclusion criteria.

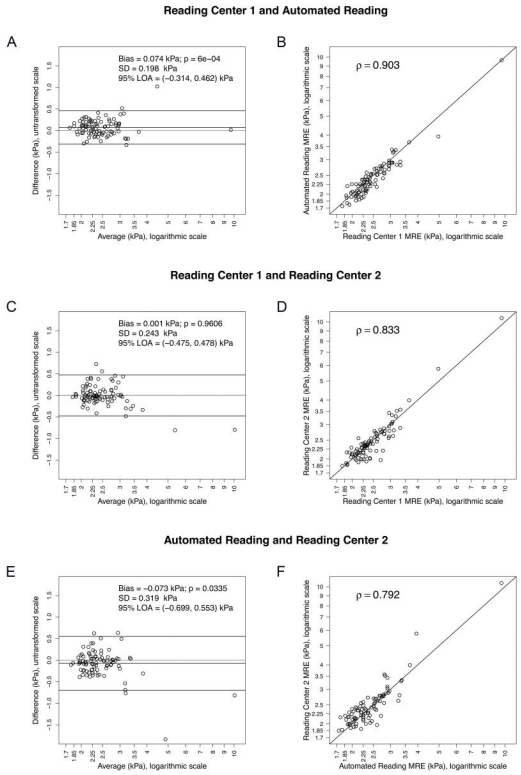

Figure 2.

Bland-Altman plots and scatterplots for each pair of elastorgram analysis methods. Logarithmic scale is used to plot raw and averaged MRE values for better visibility. Bland-Altman Bias (mean of the paired differences) and its p-value (p-value for a paired t-test), standard deviation (SD) of the differences, and the limits of agreement (LOA = Bias+/− 1.96*SD), are shown on the Bland-Altman plots A, C and E. Spearman's correlation coefficient is shown on the scatterplots B, D and F.

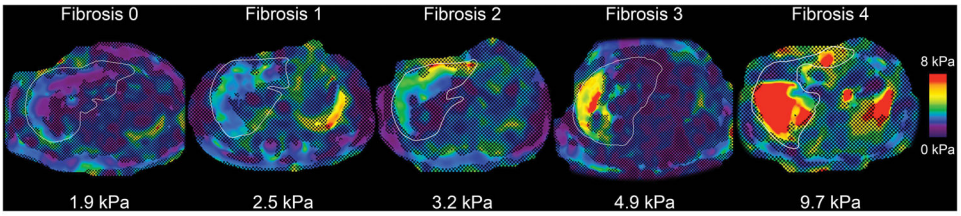

Figure 3.

Representative elastograms in children with stage 0, 1, 2, 3, and 4 fibrosis. Liver outlines were traced from anatomic images (not shown) and are overlain. Corresponding hepatic shear stiffness values in kPa are shown below each elastogram. Note greater stiffness visually and based on quantitative measurements in children with successively greater fibrosis stages. Scale bar 0–8 kPa.

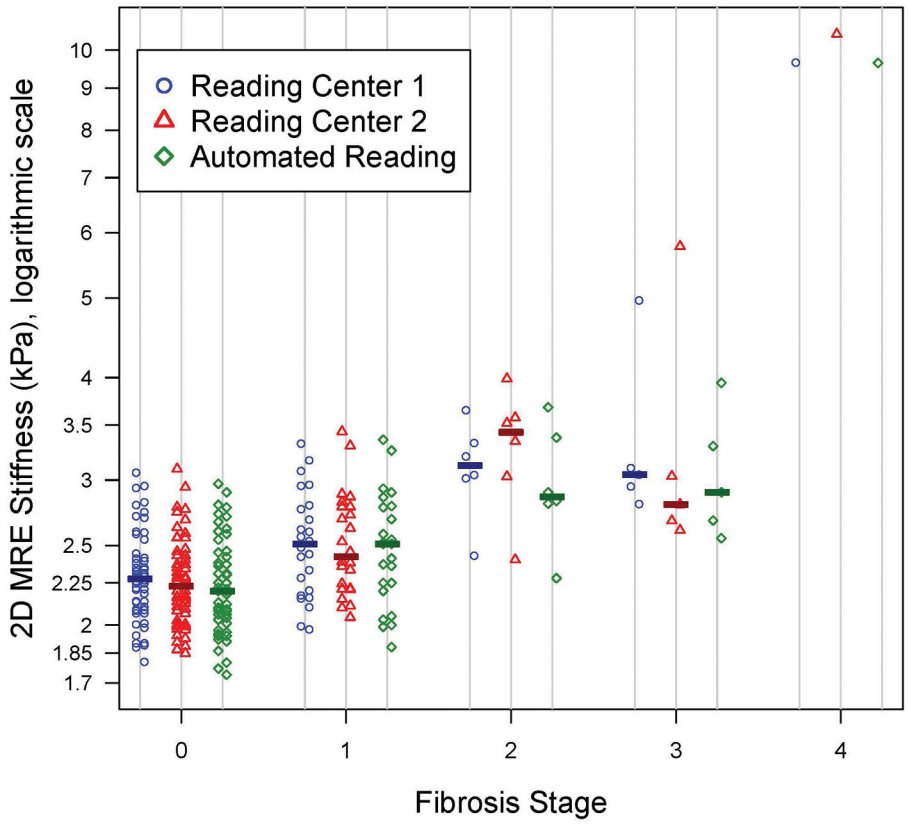

Figure 4.

Dot plot of the three elastogram analysis methods' stiffness values by fibrosis stage. The median stiffness value for each elastogram analysis method at each fibrosis stage is shown by a bar on the plot.

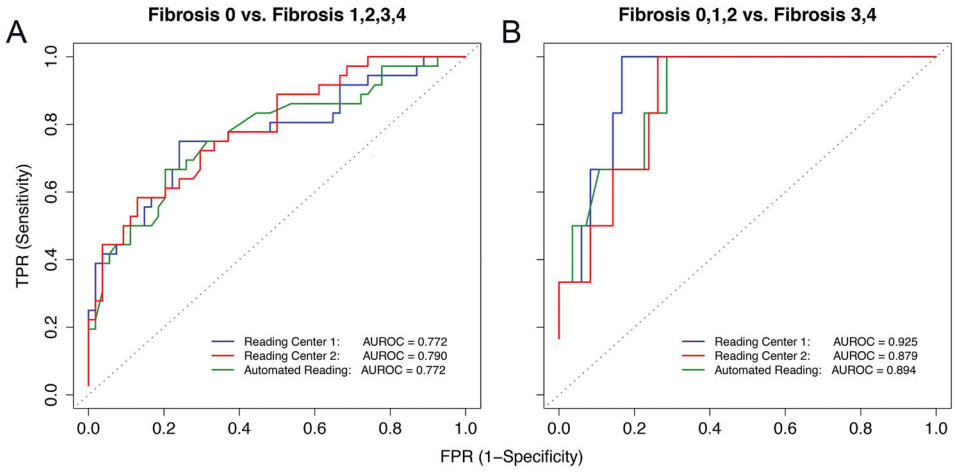

Figure 5.

ROC curves for classifying presence of any fibrosis stage > 0 (5A) and advanced fibrosis stage ≥ 3 (5B). AUROCs corresponding to the three elastogram analysis methods are shown on the plots.

Table 1

Characteristics of Study Population by Fibrosis Stage

Variables	Stage 0 N = 54	Stage 1 N = 24	Stage 2 N = 6	Stage 3 N = 5	Stage 4 N = 1
Age, mean (SD), y	13.3 (2.5)	12.7 (2.4)	14.8 (2.4)	12.2 (0.6)	12.5 (--)
Sex, N (%)					
Boys	37 (68.5)	18 (75)	6 (100)	5 (100)	0 (0)
Girls	17 (31.5)	6 (25)	0 (0)	0 (0)	1 (100)
Weight, mean (SD), Kg	79.2 (21.3)	80.3 (29.5)	98.8 (24.8)	80.5 (14.3)	47.2 (--)
Height, mean (SD), cm	161 (12)	159 (13)	168 (15)	160 (6)	138 (--)
BMI (Kg/m^2), mean (SD)	30.2 (5.6)	31.0 (7.6)	34.8 (5.6)	31.1 (3.5)	25.0 (--)
BMI Z score	2.09 (0.37)	2.15 (0.56)	2.37 (0.42)	2.26 (0.31)	1.55 (--)
ALT, median (range), U/L	50 (18–212)	80 (17–290)	131 (63–811)	206 (69–522)	302 (--)
AST, median (range), U/L	38 (16–150)	43 (16–155)	83 (38–364)	131 (51–330)	226 (--)
GGT, median (range), U/L	29 (11–177)	35 (14–109)	50 (28–232)	67 (22–144)	407 (--)
Platelet count, median (range), $\times 10^9/L$	278 (147–430)	288 (149–412)	294 (247–361)	344 (218–354)	69 (--)
Time between biopsy and MRI, mean (SD), days	76 (32)	84 (35)	62 (21)	90 (73)	115 (--)
2D MRE hepatic stiffness, median, kPa					
Reading Center 1	2.20	2.51	3.13	3.05	9.67
Reading Center 2	2.23	2.42	3.43	2.80	10.57
Automated Reading	2.28	2.51	2.87	2.90	9.66

Table 2

Cross-Validated Diagnostic Performance by Analysis Method and Fibrosis Outcomes

	Center 1 Manual Analysis	**Center 2 Manual Analysis**	**Automated Analysis**
	Any Fibrosis (Stage 0 vs Stages 1–4)		
Cut-off	2.77	2.69	2.78
Sensitivity	44.4% (27.9 – 61.9)	47.2% (30.4 – 64.5)	44.4% (27.9 – 61.9)
Specificity	90.7% (79.7 – 96.9)	88.9% (77.4 – 95.8)	90.7% (79.7 – 96.9)
PPV	76.2% (52.8 – 91.8)	73.9% (51.6 – 89.8)	76.2% (52.8 – 91.8)
NPV	71.0% (58.8 – 81.3)	71.6% (59.3 – 82.0)	71.0% (58.8 – 81.3)
Total Accuracy	72.2% (61.8 – 81.1)	72.2% (61.8 – 81.1)	72.2% (61.8 – 81.1)
	Advanced Fibrosis (Stages 0–2 vs Stages 3–4)		
Cut-off	3.05	3.03	3.33
Sensitivity	50.0% (11.8 – 88.2)	33.3% (4.3 – 77.7)	33.3% (4.3 – 77.7)
Specificity	91.7% (83.6 – 96.6)	94.0% (86.7 – 98.0)	90.5% (82.1 – 95.8)
PPV	30.0% (6.7 – 65.2)	28.6% (3.7 – 71.0)	20.0% (2.5 – 55.6)
NPV	96.2% (89.4 – 99.2)	95.2% (88.1 – 98.7)	95.0% (87.7 – 98.6)
Total Accuracy	88.9% (80.5 – 94.5)	90.0% (81.9 – 95.3)	86.7% (77.9 – 92.9)