Published in final edited form as:
Clin Gastroenterol Hepatol. 2015 March ; 13(3): 440–451.e6. doi:10.1016/j.cgh.2014.09.046.

Diagnostic Performance of Magnetic Resonance Elastography in Staging Liver Fibrosis: A Systematic Review and Meta-analysis of Individual Participant Data

Siddharth Singh, M.D.[1], **Sudhakar K. Venkatesh, M.D.**[2], **Zhen Wang, Ph.D.**[3], **Frank H. Miller, M.D.**[4], **Utaroh Motosugi, M.D.**[5], **Russell N. Low, M.D.**[6], **Tarek Hassanein, M.D.**[7], **Patrick Asbach, M.D.**[8], **Edmund M. Godfrey, B.M., B.Ch.**[9], **Meng Yin, Ph.D.**[2], **Jun Chen, Ph.D.**[2], **Andrew P. Keaveny, M.D.**[10], **Mellena Bridges, M.D.**[11], **Anneloes Bohte, M.D., Ph.D.**[12], **Mohammad Hassan Murad, M.D., M.S.**[3], **David J. Lomas, M.D.**[9], **Jayant A. Talwalkar, M.D., M.P.H.**[1], and **Richard L. Ehman, M.D.**[2]

[1]Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN [2]Department of Radiology, Mayo Clinic, Rochester, MN [3]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN [4]Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, IL [5]Department of Radiology University of Yamanashi, Chuo-shi, Yamanashi, Japan [6]Sharp and Children's MRI Center, San Diego, CA [7]Liver Centers of Southern California, Coronado, CA [8]Department of Radiology, Charité-Universitätsmedizin Berlin, Charité Campus Mitte, Berlin, Germany [9]Department of Radiology, Addenbrooke's Hospital, Cambridge, U.K. [10]Department of Transplant, Mayo Clinic, Jacksonville,

Corresponding author: Sudhakar K. Venkatesh, M.D., Associate Professor of Radiology, Department of Radiology, Mayo Clinic College of Medicine, Mayo Clinic, Rochester, MN, Venkatesh.Sudhakar@mayo.edu, Phone: 507-284-1728, Fax: 507-284-2405.

Author Contributions:

- Study concept and design: SS, SKV
- Acquisition of data: SS, SKV, FHM, UM, RNL, TH, PA, EMG, MY, JC, AK, JAT, RLE
- Statistical analysis: ZW, MHM, SS, SKV
- Interpretation of results: SS, SKV
- Drafting of the manuscript: SS, SKV
- Critical revision of the manuscript for important intellectual content: FHM, UM, RNL, TH, PA, EMG, MY, JC, AK, ZW, MHM, JAT, RLE
- Approval of the final manuscript: SS, FHM, UM, RNL, TH, PA, EMG, MY, JC, AK, ZW, MHM, JAT, RLE, SKV
- Study supervision: SKV

Disclosures: This work is supported in part by NIH grant EB001981 to MY, JC and RLE. MY, JC, RLE and the Mayo Clinic have intellectual property relating to the subject and may be eligible for royalties from licensing. RLE is CEO of Resoundant, Inc. This research has been reviewed by the Mayo Clinic Conflict of Interest Review Board and is being conducted in compliance with Mayo Clinic Conflict of Interest policies. None of the other authors have any disclosures.

FL [11]Department of Radiology, Department of Medicine, Mayo Clinic, Jacksonville, FL
[12]Department of Radiology, Academic Medical Center, Amsterdam, The Netherlands

Abstract

Background & Aims—Magnetic resonance elastography (MRE) is a non-invasive tool for staging liver fibrosis. We conducted a meta-analysis of individual participant data collected from published studies to assess the diagnostic accuracy of MRE and for staging liver fibrosis in patients with chronic liver diseases (CLD).

Methods—Through a systematic literature search of multiple databases (2003–2013), we identified studies on diagnostic performance of MRE for staging liver fibrosis in patients with CLD with native anatomy, using liver biopsy as the standard. We contacted study authors to collect data on each participant's age, sex, body mass index (BMI), liver stiffness (measured by MRE), fibrosis stage, staging system used, degree of inflammation, etiology of CLD, and interval between MRE and biopsy. Through pooled analysis, we calculated the cluster-adjusted area under receiver-operating curve (AUROC), sensitivity, and specificity of MRE for any fibrosis (≥stage 1), significant fibrosis (≥stage 2), advanced fibrosis (≥stage 3), and cirrhosis (stage 4)

Results—We analyzed data from 12 retrospective studies, comprising 697 patients (mean age, 55±13 years; 59.4% male; mean BMI, 26.9±6.7 kg/m^2; 92.1% with <1 year interval between MRE and biopsy; hepatitis C in 47.1%). Participants had fibrosis stages 0, 1, 2, 3, or 4 (19.5%, 19.4%, 15.5%, 15.9% and 29.7%, respectively). Mean AUROC values (and 95% confidence intervals) for diagnosis of any (≥stage 1), significant (≥stage 2), or advanced fibrosis (≥stage 3), and cirrhosis, were 0.84 (0.76–0.92), 0.88 (0.84–0.91), 0.93 (0.90–0.95), and 0.92 (0.90–0.94), respectively. Similar diagnostic performance was observed in stratified analysis based on sex, obesity, and etiology of CLD. The overall rate of failure of MRE was 4.3%.

Conclusion—Based on pooled analysis of data from individual participants, MRE has high accuracy for diagnosis of significant or advanced fibrosis and cirrhosis, independent of BMI and etiology of CLD. Prospective studies are warranted to better understand the diagnostic performance of MRE.

Keywords

IPD; non-invasive; elastography; diagnostic performance; pooled analysis

INTRODUCTION

Chronic liver diseases (CLD) are an important cause of morbidity and mortality in the United States – nearly 150,000 persons are diagnosed with CLD annually (of which 20% are diagnosed with cirrhosis), and 36,000 patients die of CLD, primarily attributable to complications of decompensated cirrhosis and/or hepatocellular cancer.[1, 2] Annually, these generate approximately 5.9 million CLD-related ambulatory care visits and 759,000 CLD-related hospitalizations, with healthcare costs exceeding $1.5 billion.[1]

Cirrhosis results from progressive hepatic fibrosis, a maladaptive response to chronic hepatocyte injury. The gold standard for the diagnosis and staging of fibrosis is liver biopsy. However, this procedure has several limitations – first, it is invasive and associated with an estimated morbidity and mortality rate of 3% and 0.01%, respectively; second, it is prone to sampling error resulting in misclassification of fibrosis stage in up to 25% of cases; and finally, there is considerable intra- and inter-observer variability in interpretation of histology, especially at lower stages of fibrosis.[3, 4] Because of these limitations, several non-invasive imaging-based methods of staging liver fibrosis have been developed.[5] These include ultrasound-based tests (transient elastography [TE], acoustic radiation force impulse imaging [ARFI]) or magnetic resonance-based magnetic resonance elastography (MRE). Ultrasound-based tests have the advantage of being low cost and easy to perform; however, these tests have several limitations – they evaluate a limited portion of the liver, have low applicability especially in obese patients, patients with narrow inter-rib space, presence of ascites, and findings are influenced by necroinflammatory activity, presence of hepatic congestion, cholestasis and fasting status.[5]

Using a modified phase-contrast imaging sequence to detect propagating shear waves within the liver, MRE provides a highly accurate, non-invasive measure of liver stiffness, evaluates a larger portion of the liver with the option of choosing the region of interest, can be performed in conjunction with a conventional MRI, has a high inter-observer correlation and overcomes limitations in interpretations due to obesity or ascites, making it highly applicable.[6] Several recent studies have characterized the diagnostic performance of MRE in various CLDs, including hepatitis C (HCV), hepatitis B (HBV) and non-alcoholic fatty liver disease (NAFLD).[7–9] The diagnostic performance of MRE has been the subject of two recent study-level conventional meta-analyses, and both have suggested a high diagnostic accuracy for differentiating different stages of fibrosis.[10, 11] However, study-level diagnostic accuracy meta-analysis of aggregate data have several limitations including (a) overestimation of diagnostic performance due to spectrum bias (inclusion of healthy controls), (b) selective reporting bias in individual studies (and inability to account for those at an aggregate level), (c) potential overlap of patients across studies which results in double-counting, (d) inability to identify an optimal diagnostic threshold, (e) high degree of heterogeneity (due to differences in patient characteristics, diagnostic thresholds in individual studies, etc.) and (f) limited subgroup analyses to examine stability of association and sources of heterogeneity.[12, 13] These limitations can be addressed through a pooled analysis of patient-level data from individual studies, and is ideal for meta-analyses of diagnostic accuracy.

Hence, in this systematic review, we sought to comprehensively evaluate the diagnostic performance of MRE for staging liver fibrosis in patients with CLD with native livers, through a collaborative individual participant data (IPD) meta-analysis.

METHODS

This collaborative IPD meta-analysis was conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and recommendations from Riley et al.[12, 14] The process followed an *a priori* established

protocol. This was exempt from ethical approval as the analysis involved only de-identified data, and all individual studies had received local ethics approval.

Search Strategy

First, we conducted a computer-aided systematic literature search of Medline, Embase, Web of Science and Scopus, from 2003 through September 22, 2013, with the help of an expert medical librarian, to identify all relevant articles on MRE in staging liver fibrosis. Details of the search strategy are available in the supplementary appendix. Briefly, a combination of keywords and medical subject heading (MeSH) terms were used including (mr OR "magnetic resonance") AND (elastography OR elasticity OR MRE) AND (liver OR hepatic OR fibrosis) AND (Sensitiv* OR value* OR performance OR accura* OR compar* OR predict*). Subsequently, two investigators (SS, SKV) independently reviewed the title and abstract of studies identified in the search to exclude studies that did not answer the research question of interest, based on pre-specified inclusion and exclusion criteria. The full text of the remaining articles was again independently reviewed, to determine whether it contained relevant information. Next, we manually searched the bibliographies of the selected articles, as well as review articles on the topic for additional articles. Third, we performed a manual search of conference proceedings from major gastroenterology and hepatology meetings (American Association for the Study of the Liver, European Association for the Study of the Liver, Digestive Disease Week, from 2010 to 2013) for additional abstracts on the topic. Finally, we consulted with experts in the field to identify additional published and unpublished primary studies.

Selection Criteria

We included all studies that met the following inclusion criteria: (a) evaluated the diagnostic performance of MRE as the index test, (b) using liver biopsy as the gold standard, (c) reporting fibrosis using a comparable liver biopsy staging system (METAVIR, Brunt, Ludwig, Knodell, Desmet and Scheuer), (d) in patients with intrinsic CLD with native livers, due to any etiology and stage of fibrosis. Inclusion was not otherwise restricted by study size, language or publication type. We excluded studies in which MRE was not the diagnostic test, patients with liver transplantation, liver biopsy was not the gold standard or sufficient IPD could not be obtained despite multiple attempts to contact study investigators.

Once relevant studies were identified, we contacted the corresponding author of eligible studies using electronic mail including a cover letter detailing the objectives of the collaborative meta-analysis, background information on IPD meta-analysis, and an Microsoft Excel document containing a data collection file for input of individual patient results for the project. In case of non-response, we sent another reminder email 2–4 weeks after the first; if there was no response to the 2nd email, then the study was excluded from our analysis. For investigators that responded, we obtained information on any potential overlap of patients in case of multiple related publications, and also sought unpublished data that may be eligible for inclusion in the collaborative meta-analysis if the inclusion criteria were met.

Data Abstraction

The following IPD from each study was requested and abstracted – age at time of index test, sex, body mass index (BMI), technique and reported liver stiffness on MRE, fibrosis stage on liver biopsy (and classification system used), degree of inflammation on liver biopsy (based on METAVIR activity grading – A0, no histologic necroinflammatory activity; A1, minimal activity, A2, moderate activity, A3, severe activity),[15] interval between MRE and liver biopsy, and etiology of underlying CLD (group into one of 7 categories – HCV, HBV, NAFLD, alcoholic liver disease, autoimmune hepatitis, cholestatic liver diseases including primary and secondary sclerosing cholangitis or primary biliary cirrhosis and miscellaneous/ others). To allow homogeneous comparison of liver fibrosis staging, we asked all groups to transform their reporting of fibrosis stage in accordance with a simplified 5-stage fibrosis scoring system, as reported in Supplementary Table 1.

Besides IPD, we also abstracted the following aggregate level data for each study: (a) study characteristics: primary author; time period of study/year of publication; country of study; (b) index test characteristics: failure rate and reason and (c) standard test characteristics: distribution of patients across fibrosis stages, average liver biopsy size and number of portal tracts per biopsy.

The quality assessment of included studies was performed by two investigators, independently (SS, SKV) using the quality assessment of diagnostic accuracy studies (QUADAS) questionnaire, which is designed to assess the internal and external validity of diagnostic accuracy studies included in systematic reviews.[16] This tool is a 14-item instrument that allows for the identification of important design elements in diagnostic accuracy studies such as patient spectrum, the presence or absence of observer blinding and verification bias, handling of indeterminate results, and reporting of patient loss to follow-up evaluation. Each item was scored ‘yes’ if reported (1 point) or ‘no’ if not reported or ‘unclear’ if there is no adequate information in the article to make an accurate judgment (0 point).

Outcomes Assessed

The primary outcome of interest was the diagnostic performance of MRE for the diagnosis of any (≥stage 1), significant (≥stage 2) and advanced fibrosis (≥stage 3) and cirrhosis (stage 4), compared with the reference standard of liver biopsy. Results were reported as sensitivity, specificity, area under receiver-operating curve (AUROC) with corresponding MRE stiffness cut-offs.

We performed several pre-planned subgroup and stratified analysis based on sex (males v. females), presence of obesity (BMI≥30kg/m^2 v. <30kg/m^2), degree of necroinflammatory activity (none-mild [A0–A1] v. moderate-severe [A2–A3]) and etiology of CLD (HBV, HCV, NAFLD and alcoholic liver disease). In addition, we performed a sensitivity analysis restricting only to studies in which the interval between MRE and liver biopsy was ≤1 year, to minimize risk of disease progression bias.

Statistical Analysis

We performed descriptive analyses, reporting mean (standard deviation) or median (interquartile range) for continuous variables. To investigate the association between age, sex, necroinflammatory activity, and MRE, we constructed simple linear regression models while clustering was used to account for difference between studies.

We then calculated the AUROC by pooling IPD across the included studies using the parametric two-stage model proposed by Alonzo and Pepe.[17] The correlation within each study was adjusted through clustering. We estimated the 95% confidence interval (95% CI) using bootstrapping with replacement in 10000 replications. Sensitivity and specificity of MRE and corresponding cut-offs were estimated using Youden index.[18] From pooled sensitivity and specificity, we estimated the positive and negative likelihood ratios (LR). Positive LR is the probability of a person who has the disease testing positive divided by the probability of a person who does not have the disease testing positive [i.e., positive LR = sensitivity/(1-specificity)]; negative LR is the probability of a person who has the disease testing negative divided by the probability of a person who does not have the disease testing negative [i.e., negative LR = (1-sensitivity)/specifity]. A positive LR higher than 5 and a negative LR less than 0.2 provides strong diagnostic evidence.[19] To compare the difference of AUROCs between subgroups, we used the interaction test proposed by Altman and Bland for comparisons with two estimates and one-way ANOVA for comparisons with more than two estimates.[20]

All statistical analyses were conducted using STATA version 12.1 (StataCorp LP, College Station, TX).

RESULTS

From 531 unique studies identified using our search strategy, 27 met our inclusion criteria. After contacting primary and/or corresponding authors of these studies, we were able to obtain IPD from 15 studies. Eleven studies were excluded due to non-response to electronic communication despite repeated attempts,[21–24] investigators declining to share data,[25, 26] or incomplete or missing data.[7, 27–30] Four studies were excluded from final pooled analysis despite availability of IPD – two studies due to the use of different technique of MRE (shear waves generated at 50Hz[31] or 90 Hz[32]) which result in different MRE stiffness values that could not be calibrated against the standard 60Hz MRE, and two studies performed exclusively in the patients after liver transplantation.[33, 34] Figure 1 shows the study identification and selection flowchart.

Characteristics and Quality of Included Studies

We analyzed IPD from 12 studies (9 published as full manuscripts, 3 as meeting abstracts) with 697 unique patients with CLD (Table 1);[8, 9, 35–44] all studies were retrospective in nature. Five studies were conducted in USA at 3 centers, five in Europe at 2 centers and two in Asia. All the studies used 1.5T MRI scanners, with shear waves generated at 60–62.5Hz, and had been published between 2006 and 2014. Overall, these studies were at low to moderate risk of bias – 10 of the included studies had a QUADAS score ≥10

(Supplementary Table 2).[8, 9, 35–39, 42–44] On assessment of individual QUADAS items, five studies were high risk of spectrum bias, especially since they reported on a control group of healthy patients and/or knew *a priori* the fibrosis stage of the patients;[35, 36, 39, 40, 44] however, in this IPD meta-analysis, we included only patients with CLD with individual patient liver stiffness values, minimizing the influence of spectrum bias on overall interpretation of diagnostic accuracy of MRE. Three studies provided insufficient information whether the results of MRE were interpreted while blinded to liver biopsy results, or vice versa, putting them at-risk for review bias;[8, 40, 41] two studies were performed in patients with established stage 2 or stage 3/4 fibrosis.[35, 39] The median interval between performance of MRE and liver biopsy was 20 days (IQR, 3–73 days); the interval was <1 year in 92.1% cases, and hence at low-risk of disease progression bias. Failure rate of MRE and/or liver biopsy was adequately reported in all studies. The overall failure rate of MRE was 4.3%, with the majority of failures due to iron overload. Indication for liver biopsy was not inconsistently reported.

The mean age of the pooled cohort was 55±13 years and 59.4% were males. Mean BMI was 26.9±6.7 kg/m^2 (n=410, 10 studies), with 24.1% classified as obese. Etiology of CLD in these studies included: HBV (11.6%), HCV (47.1%), NAFLD (16.5%), alcoholic liver disease (3.0%), autoimmune hepatitis (4.6%), cholestatic liver diseases (5.9%), and other miscellaneous causes (11.3%).

The distribution of fibrosis in the pooled cohort was: stage 0 19.5%, stage 1 19.4%, stage 2 15.5%, stage 3 15.9% and stage 4 29.7%; accordingly, 80.5% had any fibrosis (≥stage 1), 61.1% had significant fibrosis (≥stage 2), 45.6% had advanced fibrosis (≥stage 3) and 29.7% had cirrhosis. Distribution of histological necroinflammatory activity grade was available for 562 patients: 25.6% had no active inflammation, 41.8% had minimal inflammation, 26.0% had moderate inflammation and 6.6% had severe inflammation.

Diagnostic Accuracy of MRE

The mean liver stiffness across the entire cohort was 4.74±2.24kPa, ranging from 1.8–16.4 kPa. On cluster-adjusted pooled analysis, the AUROC of MRE for diagnosis of any (≥stage 1), significant (≥stage 2) or advanced fibrosis (≥stage 3) and cirrhosis was 0.84, 0.88, 0.93 and 0.92, respectively, suggesting excellent discriminative ability for detection of advanced fibrosis and cirrhosis, and good discriminative ability for detection of any and significant fibrosis (Table 2, Supplementary Figure 1). The corresponding MRE liver stiffness cut-offs were 3.45, 3.66, 4.11 and 4.71 kPa, respectively. Figure 2 shows the mean liver stiffness values corresponding to stage 0, stage 1, stage 2, stage 3 and stage 4 fibrosis. Based on these estimates of sensitivity and specificity, we estimated high positive and negative LR particularly for detection of advanced fibrosis and cirrhosis (Table 2). We were unable to estimate a positive and negative predictive value due to variability of prevalence depending on clinical situation in which MRE is used (primary care clinic vs. referral center hepatology practice); the prevalence estimates of fibrosis stages in this pooled cohort were skewed towards advanced fibrosis due to the retrospective nature of the studies and inherent diagnostic suspicion bias. Estimation of cluster-adjusted misclassification rate was also not

feasible in this pooled analysis due to inherent differences in the included studies, including variability in the etiology of CLD.[9]

There was no correlation between age (per unit age: regression coefficient=0.02, p=0.07) or sex (male vs. female: regression coefficient=0.21, p=0.40) or grade of inflammation [necroinflammatory activity: r=−0.52, p=0.27 (1 vs. 0); r=0.21, p=0.68 (2 vs. 0), r=0.78 p=0.16 (3 vs. 0)] and liver stiffness on MRE.

Subgroup and Sensitivity Analysis

On subgroup analysis, the diagnostic performance of MRE was comparable in males and females (Table 3). The presence or absence of obesity also did not influence the diagnostic accuracy for MRE at all stages of fibrosis; the AUROC for diagnosis of significant or advanced fibrosis in obese patients was 0.88 and 0.91, respectively. Likewise, the degree of necroinflammatory activity on liver biopsy did not significantly influence the diagnostic accuracy of MRE for detection of significant or advanced fibrosis. On stratified analysis by etiology of CLD, the diagnostic performance of MRE was comparable across patients with HCV, HBV, NAFLD and alcoholic liver disease (Table 4). The AUROC for diagnosis of any, significant and advanced fibrosis and cirrhosis in patients with NAFLD was 0.89, 0.90 and 0.94, respectively.

When we restricted analysis to patients in whom the interval between MRE and liver biopsy was <1 year (to minimize the risk of disease progression bias), MRE continued to have excellent discriminative ability for detection of advanced fibrosis and cirrhosis (Table 3).

DISCUSSION

In this systematic review and collaborative IPD meta-analysis of diagnostic performance of MRE in 12 studies with 697 patients with CLD, we made several key observations. First, the overall diagnostic accuracy of MRE for discriminating advanced fibrosis (≥stage 3) is excellent with an AUROC of 0.93; MRE's performance for diagnosis of significant (≥stage 2) and any fibrosis (≥stage 1) is also good (AUROC 0.84–0.88). The optimal cut-off of MRE for diagnosis of any, significant and advanced fibrosis and cirrhosis derived from this pooled analysis of patients with CLD is 3.45, 3.66, 4.11 and 4.71 kPa, respectively. Second, the diagnostic performance of MRE is robust and stable, independent of age, sex, obesity and degree of necroinflammatory activity. Third, MRE appears to have excellent diagnostic performance for chronic viral hepatitis.

Previous systematic reviews and study-level meta-analyses of diagnostic performance of MRE have suggested excellent diagnostic performance of MRE. In their conventional meta-analysis of aggregate data of 5 studies, Wang and colleagues demonstrated a high AUROC for differentiating stages of fibrosis using MRE (0.95–0.98).[11] Likewise, in another study-level meta-analysis of 11 studies, Guo and colleagues observed a similarly high AUROC for discriminating fibrosis stages (0.94–0.97).{Guo, 2014 #1002} However, as mentioned above, study-level meta-analysis of aggregate data have several inherent limitations decreasing confidence in estimates. Using participant level data, through collaboration with multiple research groups, we were able to overcome several of these limitations by (a) using

standardized statistical analysis across studies, (b) adjusting for baseline potential confounding factors (like age, sex, obesity, necroinflammatory activity etc.), (c) accounting for missing data and minimizing overlapping data in different studies, (d) decreasing selective reporting bias, (e) attempting to minimize spectrum bias by excluding data from healthy controls, (f) assessing robustness of association and sources of heterogeneity using subgroup and stratified analysis and (g) identifying differences in diagnostic performance based on underlying etiology of CLD. AUROC derived from this IPD meta-analysis represents a more reliable, accurate and real-world diagnostic performance of MRE for staging hepatic fibrosis.

Overall, we observed that the diagnostic performance of MRE was comparable, if not superior, to that of ultrasound-based methods of TE and ARFI. In their study-level meta-analysis of 50 studies on diagnostic accuracy of TE, Friedrich-Rust et al has estimated that the mean AUROC for diagnosis of significant and advanced fibrosis and cirrhosis is 0.84, 0.89 and 0.94, respectively; the diagnostic performance of TE for differentiating early stages of fibrosis (stage 0 vs. ≥stage 1) was not reported.[45] Similarly, a conventional meta-analysis of 15 studies on diagnostic accuracy of ARFI, Guo et al estimated that the mean AUROC for diagnosis of stage 1, stage 2, stage 3 and stage 4 fibrosis are 0.82, 0.85, 0.94, and 0.94, respectively; and the corresponding AUROCs for MRE in their analysis were 0.94, 0.97, 0.96, and 0.97.[10] As mentioned above, study-level meta-analyses of aggregate data tend to overestimate the diagnostic performance of an index test due to a variety of factors, in particular due to spectrum bias (inclusion of healthy controls) and selective reporting bias. Moreover, ultrasound-based techniques are limited by body habitus. In a single center prospective study of over 13000 TE exams, the rate of failed or unreliable TE measurements in obese patients was 16.9% and 35.4%.[46]; similarly, the rate of unreliable ARFI exams in obese patients was 17.6%.[47] Obesity, in particular high waist circumference, has also been associated with higher discordance with biopsy findings with both over- and underestimation of fibrosis stage.[48, 49] In contrast, we observed that the diagnostic performance of MRE was unaffected by obesity, with comparable AUROCs in obese and non-obese patients. This is particularly relevant given the epidemic of obesity and increasing prevalence of NAFLD. Besides body habitus, age and sex also affect the reliability of TE and ARFI.[46, 47] Neither of these factors influenced the performance of MRE in our pooled analysis. Failure rate of MRE in our pooled analysis was 4.3%, usually in patients with hemochromatosis with high hepatic iron content; newer improved sequences are available to perform MRE in patients with iron overload, and it is anticipated that the failure rate would decrease to <1%. Comparative studies of MRE and TE have suggested higher technical success rate as well as superior diagnostic accuracy of MRE.[30]

Acute viral or drug-induced hepatitis has been shown to modulate tissue elasticity and increase its stiffness, and hence use of elastography techniques in these situations may falsely classify patients as having advanced fibrosis or cirrhosis.[50] Recent studies have also suggested that chronic necroinflammatory activity due to viral hepatitis or NAFLD may also influence TE-measured liver stiffness in patients at all stages of fibrosis and should be regarded as a strong confounding variable.[51, 52] In a recent study, Ichikawa et al have also observed that hepatitis activity grade may also influence liver stiffness measured using MRE.[53] However, in our analysis, the diagnostic accuracy of MRE for detection of

significant or advanced fibrosis was not significantly influenced by presence of severe necroinflammation. Some statistically significant differences in performance of MRE were observed, for presence of any fibrosis and cirrhosis, in patients with severe inflammation; this may be related to multiple comparisons. Large, prospective studies are needed to study the influence of inflammation on MRE-measured liver stiffness.

Individual studies on diagnostic accuracy of MRE have often been criticized for combining patients with different etiology of CLD together (and hence increasing clinical heterogeneity). With the availability of IPD, we were able to estimate the diagnostic performance of MRE across various etiologies of CLD through subgroup analyses. We observed that MRE continues to have high diagnostic accuracy in patients with viral hepatitis, NAFLD or alcoholic liver diseases, the leading causes of CLD across the world. In contrast to ultrasound-based techniques, which have inferior performance characteristics in obese patients with NAFLD,[48, 49] the diagnostic accuracy of MRE continued to be superlative. In a recent prospective cross-sectional study of 117 patients with NAFLD, Loomba et al have confirmed high diagnostic accuracy of MRE in diagnosing advanced fibrosis in patients with NAFLD (AUROC for discriminating advanced fibrosis, 0.92). This attests to its high applicability for diagnosis and staging of fibrosis.

Besides being an IPD meta-analysis, our systematic review had several other strengths, including: (a) comprehensive and systematic literature search with well-defined inclusion criteria, carefully excluding redundant studies; (b) rigorous evaluation of study quality; (c) sub-group and sensitivity analyses to evaluate the stability of findings and identify potential factors responsible for inconsistencies and (d) being able to establish optimal diagnostic thresholds corresponding to the inflection point in the ROC.

There were several limitations in our study. First, though we tried to contact authors of all studies to participate in this collaborative meta-analysis, we were unable to obtain IPD from some investigators. In particular, our analysis was only able to evaluate the diagnostic performance of MRE performed at 60–62.5Hz, and not at 50Hz as is practiced in certain parts of Europe. Studies using MRE performed at 50Hz have suggested a similar high diagnostic accuracy for detection of significant and advanced fibrosis.[30, 31] Second, while IPD meta-analysis was able to alleviate several of the limitations of a conventional aggregate data meta-analysis, ours was still a retrospective analysis with several inherent variations due to lack of standardized performance of index test and lack of centralized reading of biopsies. All studies included in the meta-analysis were retrospective in nature, and liver histology was reviewed by multiple pathologists across 8 centers. We tried to minimize the impact of these by performing cluster-adjusted pooled analysis, subgroup analysis by center and sensitivity analyses to account for disease progression bias. Variable liver fibrosis staging systems were used in individual studies based on etiology of CLD. We tried to improve comparability by a priori requesting investigators to transform fibrosis stages into a simplified 5-stage fibrosis scoring system; however, such a transformation may result in misclassification. Third, there was incomplete capturing of some potential confounding factors in the included studies, such as BMI. However, as best as we could assess with available data, it is unlikely that these variables would significantly influence the diagnostic performance of MRE. Fourth, though we were able to identify optimal diagnostic

thresholds, these should be interpreted cautiously and require prospective validation in a well-defined population; these thresholds are likely to vary depending on practice where MRE is applied. Fifth, the gold standard in these included studies was liver biopsy. Liver biopsy itself is not a perfect gold standard, since it samples only 1/50,000 of total liver mass and significant discrepancy in fibrosis stage as high as 33% can be observed depending on site of liver biopsy.[3, 4] It is conceivable that the diagnostic accuracy of MRE may in fact be higher given its ability to globally evaluate the liver. Sixth, we restricted our analysis only to native livers and excluded patients with liver transplantation; altered anatomy as well as biological factors introduced due to transplantation (such as rejection) may alter the diagnostic performance of MRE in transplanted livers and merits further evaluation. Finally, though there is increasing interest in the prognostic utility of MRE and other elastographic techniques in patients with CLD,[54] our study was not designed to address this question.

In conclusion, through a systematic review and collaborative IPD meta-analysis, we observed that MRE is a highly accurate, non-invasive technique for diagnosis and staging of liver fibrosis, which is not significantly influenced by age, sex, obesity, degree of inflammation, and etiology of CLD. Prospective studies in patients who have a clinical indication for liver biopsy would further help us assess the diagnostic test performance of MRE in the real world setting. Additionally, longitudinal studies are needed to assess whether changes in MRE-derived liver stiffness as a result of treatment predicts improvement in long term clinical outcomes, and whether it can be used as endpoint in treatment trials. Future comparative accuracy, cost-effectiveness and patient preference analyses are needed to identify where MRE (as compared to ultrasound-based elastographic techniques) best fits in terms of diagnosis and monitoring of patients with CLD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We wish to thank Ms. Patricia Erwin, M.L.S., Senior Medical Librarian at the Mayo Clinic Library for helping in the literature search for this systematic review and meta-analysis.

References

1. Everhart JE, Ruhl CE. Burden of digestive diseases in the United States Part III: Liver, biliary tract, and pancreas. Gastroenterology. 2009; 136:1134–44. [PubMed: 19245868]
2. Asrani SK, Larson JJ, Yawn B, et al. Underestimation of liver-related mortality in the United States. Gastroenterology. 2013; 145:375–82. [PubMed: 23583430]
3. Bravo AA, Sheth SG, Chopra S. Liver biopsy. N Engl J Med. 2001; 344:495–500. [PubMed: 11172192]
4. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol. 2002; 97:2614–8. [PubMed: 12385448]
5. Castera L. Noninvasive methods to assess liver disease in patients with hepatitis B or C. Gastroenterology. 2012; 142:1293–1302. [PubMed: 22537436]
6. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. J Magn Reson Imaging. 2013; 37:544–55. [PubMed: 23423795]

7. Kim D, Kim WR, Talwalkar JA, et al. Advanced fibrosis in nonalcoholic fatty liver disease: noninvasive assessment with MR elastography. Radiology. 2013; 268:411–9. [PubMed: 23564711]
8. Ichikawa S, Motosugi U, Ichikawa T, et al. Magnetic resonance elastography for staging liver fibrosis in chronic hepatitis C. Magn Reson Med Sci. 2012; 11:291–7. [PubMed: 23269016]
9. Venkatesh SK, Wang G, Lim SG, et al. Magnetic resonance elastography for the detection and staging of liver fibrosis in chronic hepatitis B. Eur Radiol. 2013:1–9. [PubMed: 23184074]
10. Guo Y, Parthasarathy S, Goyal P, et al. Magnetic resonance elastography and acoustic radiation force impulse for staging hepatic fibrosis: a meta-analysis. Abdom Imaging. 201410.1007/s00261-014-0137-6
11. Wang QB, Zhu H, Liu HL, et al. Performance of magnetic resonance elastography and diffusion-weighted imaging for the staging of hepatic fibrosis: A meta-analysis. Hepatology. 2012; 56:239–47. [PubMed: 22278368]
12. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010; 340:c221. [PubMed: 20139215]
13. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet. 1993; 341:418–22. [PubMed: 8094183]
14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009; 151:264–9. W64. [PubMed: 19622511]
15. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology. 1996; 24:289–93. [PubMed: 8690394]
16. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003; 3:25. [PubMed: 14606960]
17. Alonzo TA, Pepe MS. Distribution-free ROC analysis using binary regression techniques. Biostatistics. 2002; 3:421–32. [PubMed: 12933607]
18. Youden WJ. Index for rating diagnostic tests. Cancer. 1950; 3:32–5. [PubMed: 15405679]
19. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA. 1994; 271:703–7. [PubMed: 8309035]
20. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. BMJ. 2003; 326:219. [PubMed: 12543843]
21. Batheja MJ, Silva AC, De Petris G, et al. Role of magnetic resonance elastography in assessing hepatic fibrosis. Gastroenterology. 2011; 140:S927.
22. Choi YR, Lee JM, Yoon JH, et al. Comparison of magnetic resonance elastography and gadoxetate disodium-enhanced magnetic resonance imaging for the evaluation of hepatic fibrosis. Invest Radiol. 2013; 48:607–13. [PubMed: 23538889]
23. Kim BH, Lee JM, Lee YJ, et al. MR elastography for noninvasive assessment of hepatic fibrosis: experience from a tertiary center in Asia. J Magn Reson Imaging. 2011; 34:1110–6. [PubMed: 21932355]
24. Yoon JH, Lee JM, Woo HS, et al. Staging of hepatic fibrosis: comparison of magnetic resonance elastography and shear wave elastography in the same individuals. Korean J Radiol. 2013; 14:202–12. [PubMed: 23483022]
25. Lee VS, Miller FH, Omary RA, et al. Magnetic resonance elastography and biomarkers to assess fibrosis from recurrent hepatitis C in liver transplant recipients. Transplantation. 2011; 92:581–6. [PubMed: 21822174]
26. Loomba R, Wolfson T, Ang B, et al. Magnetic resonance elastography predicts advanced fibrosis in patients with nonalcoholic fatty liver disease: A prospective study. Hepatology. 201410.1002/hep.27362
27. Huwart L, Peeters F, Sinkus R, et al. Liver fibrosis: non-invasive assessment with MR elastography. NMR Biomed. 2006; 19:173–9. [PubMed: 16521091]
28. Huwart L, Salameh N, ter Beek L, et al. MR elastography of liver fibrosis: preliminary results comparing spin-echo and echo-planar imaging. Eur Radiol. 2008; 18:2535–41. [PubMed: 18504591]

29. Huwart L, Sempoux C, Salameh N, et al. Liver fibrosis: noninvasive assessment with MR elastography versus aspartate aminotransferase-to-platelet ratio index. Radiology. 2007; 245:458–66. [PubMed: 17940304]
30. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology. 2008; 135:32–40. [PubMed: 18471441]
31. Bohte AE, de Niet A, Jansen L, et al. Non-invasive evaluation of liver fibrosis: a comparison of ultrasound-based transient elastography and MR elastography in patients with viral hepatitis B and C. Eur Radiol. 2014; 24:638–48. [PubMed: 24158528]
32. Rouviere O, Yin M, Dresner MA, et al. MR elastography of the liver: preliminary results. Radiology. 2006; 240:440–8. [PubMed: 16864671]
33. Crespo S, Bridges M, Nakhleh R, et al. Non-invasive assessment of liver fibrosis using magnetic resonance elastography in liver transplant recipients with hepatitis C. Clin Transplant. 2013; 27:652–8. [PubMed: 23837611]
34. Klatt DAP, Kamphues C, Hirsch S, Papazoglou S, Braun J, Sack I. MR elastography of liver transplant patients using parallel imaging techniques. Proc Intl Soc Mag Reson Med. 2011; 19:1485.
35. Asbach P, Klatt D, Hamhaber U, et al. Assessment of liver viscoelasticity using multifrequency MR elastography. Magn Reson Med. 2008; 60:373–9. [PubMed: 18666132]
36. Asbach P, Klatt D, Schlosser B, et al. Viscoelasticity-based staging of hepatic fibrosis with multifrequency MR elastography. Radiology. 2010; 257:80–6. [PubMed: 20679447]
37. Chen J, Talwalkar JA, Yin M, et al. Early detection of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease by using MR elastography. Radiology. 2011; 259:749–56. [PubMed: 21460032]
38. Godfrey EM, Patterson AJ, Priest AN, et al. A comparison of MR elastography and 31P MR spectroscopy with histological staging of liver fibrosis. Eur Radiol. 2012; 22:2790–7. [PubMed: 22752441]
39. 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; 41:841–8. [PubMed: 17099421]
40. Low RN, Hassanein T. MR elastography: Validation and reproducibility of measurements of mean liver stiffness and fibrosis. J Hepatol. 2012; 56:S415.
41. Nguyen D, Talwalkar JA, Yin M, et al. Assessment of hepatic fibrosis by magnetic resonance elastography in patients with sclerosing cholangitis. Gastroenterology. 2011; 140:S919.
42. Rustogi R, Horowitz J, Harmath C, 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; 35:1356–64. [PubMed: 22246952]
43. Wang Y, Ganger DR, Levitsky J, et al. Assessment of chronic hepatitis and fibrosis: comparison of MR elastography and diffusion-weighted imaging. AJR American Journal of Roentgenology. 2011; 196:553–61. [PubMed: 21343496]
44. Yin M, Talwalkar JA, Glaser KJ, et al. Assessment of Hepatic Fibrosis With Magnetic Resonance Elastography. Clin Gastroenterol Hepatol. 2007; 5:1207–1213. [PubMed: 17916548]
45. Friedrich-Rust M, Ong MF, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology. 2008; 134:960–74. [PubMed: 18395077]
46. Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. Hepatology. 2010; 51:828–35. [PubMed: 20063276]
47. Bota S, Sporea I, Sirli R, et al. Factors associated with the impossibility to obtain reliable liver stiffness measurements by means of Acoustic Radiation Force Impulse (ARFI) elastography--analysis of a cohort of 1,031 subjects. Eur J Radiol. 2014; 83:268–72. [PubMed: 24360231]
48. Myers RP, Pomier-Layrargues G, Kirsch R, et al. Discordance in fibrosis staging between liver biopsy and transient elastography using the FibroScan XL probe. J Hepatol. 2012; 56:564–70. [PubMed: 22027584]
49. Petta S, Di Marco V, Camma C, et al. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther. 2011; 33:1350–60. [PubMed: 21517924]

50. Sagir A, Erhardt A, Schmitt M, et al. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. Hepatology. 2008; 47:592–5. [PubMed: 18098325]

51. Kim SU, Kim JK, Park YN, et al. Discordance between liver biopsy and Fibroscan(R) in assessing liver fibrosis in chronic hepatitis b: risk factors and influence of necroinflammation. PLoS One. 2012; 7:e32233. [PubMed: 22384189]

52. Tapper EB, Cohen EB, Patel K, et al. Levels of alanine aminotransferase confound use of transient elastography to diagnose fibrosis in patients with chronic hepatitis C virus infection. Clin Gastroenterol Hepatol. 2012; 10:932–937. e1. [PubMed: 22289876]

53. Ichikawa S, Motosugi U, Nakazawa T, et al. Hepatitis activity should be considered a confounder of liver stiffness measured with MR elastography. J Magn Reson Imaging. 201410.1002/jmri.24666

54. Singh S, Fujii LL, Murad MH, et al. Liver stiffness is associated with risk of decompensation, liver cancer, and death in patients with chronic liver diseases: a systematic review and meta-analysis. Clin Gastroenterol Hepatol. 2013; 11:1573–84. [PubMed: 23954643]

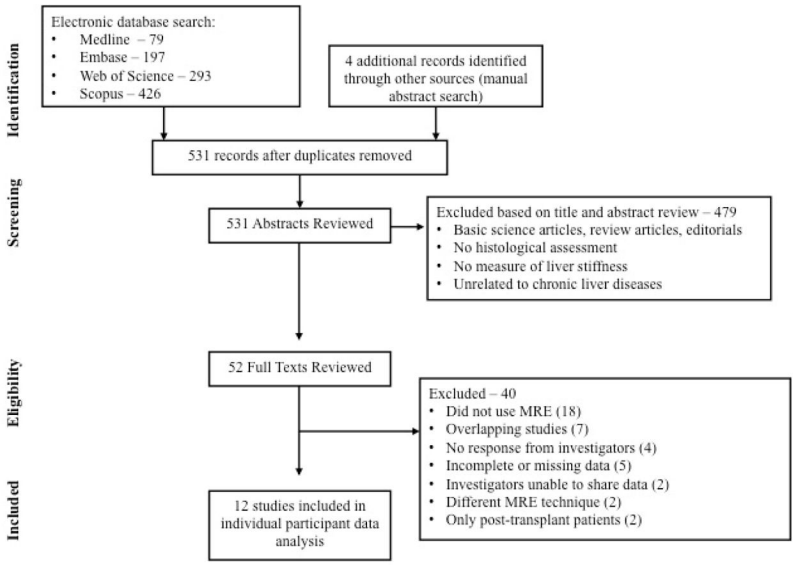

Figure 1.
Flow sheet summarizing study identification and selection.

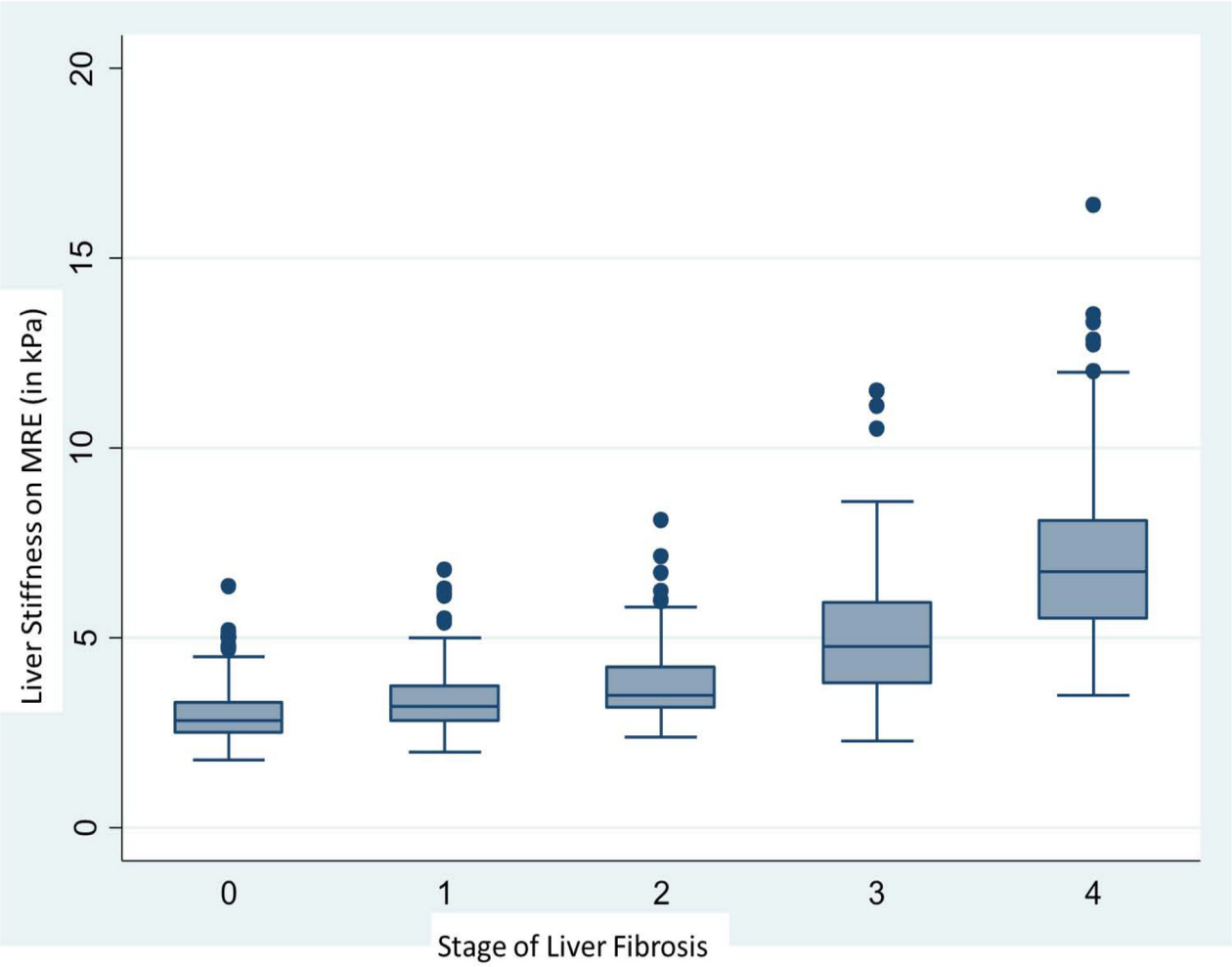

Figure 2.

Composite box-plot graph showing magnetic resonance elastography (MRE), stiffness values for various stages (METAVIR) of fibrosis. Horizontal line through each box represents a median value and each box top and bottom represent data from the 25th to 75th percentile (middle 50% of observations). Whiskers represent data from minimum to maximum excluding outliers which are represented as separate dots.

Table 1

Characteristics of 12 included studies. Please note some results were updated from the original published manuscript based on individual patient data; all groups were able to convert different fibrosis staging systems into the standard Metavir classification for reporting in the pooled analysis.

Study	Time Period	Total number of patients included in analysis	Failure rate (MRE/ liver biopsy)	Patient characteristics			Standard Test			Index Test	
				Age, mean; Sex (%males)	BMI, mean	Etiology of CLD	Fibrosis staging system	Biopsy quality – Length (cm), No. of portal tracts	Distribution of fibrosis stage (stage 0/1/2/3/4)	Interval between MRE and liver biopsy (days)	Stiffness, mean, by stage of fibrosis
Rochester, Minnesota, U.S.A. – MRE Scanner – GE Medical Systems, WI, USA											
Yin, 2007[44]	2005– 2007	48	0	56 (11); 46%	29.1 (5.4)	HCV – 15; NAFLD – 15; ALD – 1; AIH – 8; Cholestatic – 3; Others – 6	Metavir; Batts- Ludwig; Brunt	2.0 (0.9); 10 (4)	Stage 0/1/2/3/4 – 14/6/5/5/18	195 (117)	Stage 0 – 3.79 (1.03); Stage 1 – 4.01 (1.21); Stage 2 – 5.22 (1.51); Stage 3 – 6.01 (1.70); Stage 4 – 8.21 (2.26)
Chen, 2011[37]	2007– 2010	58	0	51 (13); 17%	37.0 (8.4)	NALFD - 58	Brunt	NR	Stage 0/1/2/3/4 – 29/15/3/5/6	30 (22)	Stage 0 – 2.71 (0.50); Stage 1 – 3.43 (0.77); Stage 2 – 4.58 (1.32); Stage 3 – 5.55 (2.01); Stage 4 – 5.82 (1.33)
Nguyen, 2011[41] [Abst]	NR	26	0	51 (18); 77%	27.4 (4.7)	Cholestatic – 26	Batts and Ludwig	NR	Stage 0/1/2/3/4 – 3/2/1/6/14	<365	Stage 0 – 3.53 (1.45); Stage 1 – 2.76 (0.93); Stage 2 – 4.00

Study	Time Period	Total number of patients included in analysis	Failure rate (MRE/ liver biopsy)	Patient characteristics			Standard Test			Index Test	
				Age, mean; Sex (%males)	BMI, mean	Etiology of CLD	Fibrosis staging system	Biopsy quality – Length (cm), No. of portal tracts	Distribution of fibrosis stage (stage 0/1/2/3/4)	Interval between MRE and liver biopsy (days)	Stiffness, mean, by stage of fibrosis
											(NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91 (NA); Stage 3 – 3.91
Chicago, Illinois, U.S.A. – MRE Scanner – Seimens Medical Solutions, Erlangen, Germany											
Wang, 2011[43]	2008– 2009	71	4/0	55 (10); 62%	28.0 (4.9)	HBV – 3; HCV – 40; NAFLD – 8; ALD – 1; AIH – 3; Cholestatic – 6; Others – 11	Metavir; Batts- Ludwig; Brunt	NR	Stage 0/1/2/3/4 – 28/12/6/6/20	60 (1–359)	Stage 0 – 3.33 (0.81); Stage 1 – 3.90 (0.91); Stage 2 – 5.43 (1.00); Stage 3 – 7.70 (2.26); Stage 4 – 8.87 (2.76)
Rustogi, 2012[42]	2008– 2011	76	18[a]/0	54 (12); 48%	28.0 (4.1)	HBV – 1; HCV – 30; NAFLD – 5; ALD – 2; AIH – 8; Cholestatic – 11; Others – 19	Metavir; Brunt; Ludwig	2.08 (0.2); NR	Stage 0/1/2/3/4 – 34/6/6/26	129 (108)	Stage 0 – 3.19 (0.89); Stage 1 – 4.52 (1.33); Stage 2 – 4.64 (1.61); Stage 3 – 7.29 (2.61); Stage 4 – 7.62 (2.05)
San Diego, California, U.S.A. – MRE Scanner – Not reported											
Low, 2012[40]	NR	101	NR	55 (15); 61%	NR	HBV – 4; HCV 0 54; NAFLD – 12; AIH – 2; Cholestatic – 2; Others – 27	Metavir	NR	Stage 0/1/2/3/4 – 14/19/19/18/ 31	120 (NR)	Stage 0 – 2.35 (0.35); Stage 4 – 3.92

Study	Time Period	Total number of patients included in analysis	Failure rate (MRE/ liver biopsy)	Patient characteristics			Standard Test			Index Test	
				Age, mean; Sex (%males)	BMI, mean	Etiology of CLD	Fibrosis staging system	Biopsy quality – Length (cm), No. of portal tracts	Distribution of fibrosis stage (stage 0/1/2/3/4)	Interval between MRE and liver biopsy (days)	Stiffness, mean, by stage of fibrosis
											(0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34 (0.52); Stage 2 – 4.34
Cambridge, United Kingdom – MRE Scanner – GE Medical Systems, WI, USA											
Godfrey, 2012[38]	2007– 2009	57 (of 71 – 14 patients excluded since they were post-liver transplant)	6/0	49 (12); 72%	NR	HBV – 3; HCV – 26; NAFLD – 8; ALD – 7; AIH – 8; Cholestatic – 1; Post- LT – 14; Others – 5	Ishak, NAS	NR	Stage 0/1/2/3/4 – 11/35/11/9/5	1 (0)	Stage 0 – 3.10 (0.32); Stage 1 – 3.46 (0.57); Stage 2 – 3.64 (0.68); Stage 3 – 5.08 (1.57); Stage 4 – 6.27 (1.38)
Berlin, Germany – MRE Scanner – Seimens Medical Solutions, Erlangen, Germany											
Klatt, 2006[39]	NR	2	0	52 (10); 50%	21.0 (2.8)	HCV - 2	Desmet	NR	Stage 3 – 2	20 (2)	Stage 3 – 4.39 (1.95)
Asbach, 2008[35]	NR	8	0	58 (9.5); 62%	23.5 (4.6)	HCV – 4; NAFLD – 1; ALD – 3	Desmet	NR	Stage 3/4 – 3/5	15 (8)	Stage 3/4 – 2.6 (0.94)
Asbach, 2010[36]	2006– 2009	72	1/1	53 (12); 63%	25.2 (4.4)	HCV – 42; HBV – 7; NAFLD – 8; ALD – 7; AIH – 3; Cholestatic – 3; Others – 2	Desmet	NR	Stage 0/1/2/3/4 – 0/20/17/16/19	21 (14)	Stage 1 – 2.61 (0.43); Stage 2 – 3.00 (0.63); Stage 3 – 3.86 (0.61); Stage 4 –

Study	Time Period	Total number of patients included in analysis	Failure rate (MRE/ liver biopsy)	Patient characteristics			Standard Test			Index Test	
				Age, mean; Sex (%males)	BMI, mean	Etiology of CLD	Fibrosis staging system	Biopsy quality – Length (cm), No. of portal tracts	Distribution of fibrosis stage (stage 0/1/2/3/4)	Interval between MRE and liver biopsy (days)	Stiffness, mean, by stage of fibrosis
											5.86 (1.22) 5.86 (1.22)
Singapore – MRE Scanner – GE Medical Systems, WI, USA											
Venkatesh, 2013[9]	2009– 2011	63	1/0	50 (12); 70%	24.8 (4.0)	HBV – 63	Metavir	2.1 (0.8); >10 portal tracts – 100%	Stage 0/1/2/3/4 – 12/12/10/8/21	63 (68)	Stage 0 – 2.52 (0.16); Stage 1 – 2.88 (0.15); Stage 2 – 3.49 (0.35); Stage 3 – 4.35 (0.73); Stage 4 – 6.54 (2.09)
Yamanashi, Japan – MRE Scanner – GE Medical Systems, WI, USA											
Ichikawa, 2012[8]	2010– 2012	114	5/0	66 (10); 75%	23.0 (3.2)	HCV – 114	Metavir	NR	Stage 0/1/2/3/4 – 3/15/28/25/43	23 (17)	Stage 0 – 2.10 (0.10); Stage 1 – 2.42 (0.29); Stage 2 – 3.16 (0.32); Stage 3 – 4.22 (0.78); Stage 4 – 6.21(1.09)

Note: Reason for failure of MRE was usually high hepatic iron content precluding valid estimation of stiffness using MRE;

[a] Rustogi et al excluded 18 patients from their original cohort of 100 patients with both MRE and liver biopsy due to failure of MRE (12 patients with hemochromatosis, 6 patients with poor MRE wave penetration)

Table 2

Pooled analysis of the diagnostic performance of magnetic resonance elastography for diagnosis and staging of liver fibrosis, based on 697 patients from 12 studies.

Fibrosis Stage	Optimal cut-off (kPa)	AUROC (95% CI)	Sensitivity	Specificity	Positive LR	Negative LR
Any Fibrosis (≥Stage 1)	3.45	0.84 (0.76–0.92)	0.73	0.79	3.48	0.34
Significant Fibrosis (≥Stage 2)	3.66	0.88 (0.84–0.91)	0.79	0.81	4.16	0.26
Advanced Fibrosis (≥Stage 3)	4.11	0.93 (0.90–0.95)	0.85	0.85	5.67	0.18
Cirrhosis (Stage 4)	4.71	0.92 (0.90–0.94)	0.91	0.81	4.79	0.11

Abbreviations: AUROC-Area under receiver-operating curve, CI-Confidence intervals, LR-likelihood ratio

Table 3

Subgroup and sensitivity analyses. Diagnostic performance of MRE, stratified by sex, presence or absence of obesity and degree of necroinflammatory activity, as well as on restricting to patients in whom the interval between MRE and liver biopsy was <1 year.

Categories	Subgroups¶	Fibrosis Stage	AUROC	Sensitivity	Specificity	$P_{interaction}$*
Sex	Males (n=413)	≥1	0.78 (0.67–0.87)	0.65	0.77	Males vs. Females: ≥F1: 0.07
		≥2	0.85 (0.80–0.89)	0.73	0.81	≥F2: 0.06
		≥3	0.92 (0.88–0.95)	0.82	0.87	≥F3: 0.76
		Stage 4	0.92 (0.78–0.94)	0.93	0.81	F4: 0.99
	Females (n=282)	≥1	0.89 (0.80–0.95)	0.79	0.83	
		≥2	0.91 (0.85–0.94)	0.83	0.82	
		≥3	0.93 (0.89–0.95)	0.90	0.82	
		Stage 4	0.92 (0.90–0.95)	0.89	0.81	
Obesity	BMI ≥30kg/m^2 (obese) (n=99)	≥1	0.88 (0.84–0.98)	0.79	0.81	Obese vs. non-obese ≥F1: 0.37
		≥2	0.92 (0.50–0.95)	0.97	0.83	≥F2: 0.87
		≥3	0.91 (0.74–0.94)	0.89	0.78	≥F3: 0.38
		Stage 4	0.92 (0.71–0.96)	0.87	0.82	F4: 0.83
	BMI <30kg/m^2 (n=311)	≥1	0.81 (0.61–0.92)	0.71	0.75	
		≥2	0.90 (0.83–0.97)	0.81	0.82	
		≥3	0.95 (0.91–0.97)	0.86	0.92	
		Stage 4	0.94 (0.89–0.97)	0.94	0.83	
Inflammation grade	Absent-mild (n=379)	≥1	0.76 (0.65–0.83)	0.64	0.75	Absent-mild inflammation vs. moderate-severe inflammation ≥F1: 0.03
		≥2	0.88 (0.82–0.92)	0.77	0.82	≥F2: 0.14
		≥3	0.93 (0.89–0.97)	0.85	0.87	≥F3: 0.23
		Stage 4	0.95 (0.90–0.97)	0.94	0.85	F4: 0.02
	Moderate-severe (n=183)	≥1	0.90 (0.79–0.97)	0.79	0.87	
		≥2	0.81 (0.76–0.90)	0.67	0.81	
		≥3	0.89 (0.85–0.95)	0.80	0.83	
		Stage 4	0.84 (0.75–0.91)	0.93	0.70	
Sensitivity Analysis						

Categories	Subgroups¶	Fibrosis Stage	AUROC	Sensitivity	Specificity	$P_{interaction}$ *
Interval between MRE and liver biopsy	<1 year (n=493)	≥1	0.80 (0.70–0.88)	0.70	0.76	N/A
		≥2	0.86 (0.81–0.91)	0.78	0.78	
		≥3	0.92 (0.87–0.96)	0.86	0.82	
		Stage 4	0.93 (0.90–0.95)	0.94	0.82	

[Abbreviations: AUROC-Area under receiver operating curve, BMI-Body mass index, MRE-Magnetic resonance elastography]

*Represents the comparison of diagnostic performance of MRE between subgroups (males vs. females, obese vs. non-obese, none-mild vs. moderate-severe inflammation) for each corresponding fibrosis stage (any, significant, advanced fibrosis and cirrhosis). To compare the difference of AUROCs between subgroups, we used the interaction test proposed by Altman and Bland for comparisons with two estimates.

¶Please note that numbers in subgroups may not add up to 697 due to missing data in individual studies

Table 4

Diagnostic performance of MRE for detection of fibrosis, stratified by etiology of chronic liver disease.

Subgroups	Fibrosis Stage	AUROC	Sensitivity	Specificity	$P_{interaction}$ *
Hepatitis C (n=328)	≥1	0.82 (0.75–0.92)	0.68	0.83	≥F1: 0.96
	≥2	0.88 (0.84–0.91)	0.77	0.83	≥F2: 0.90
	≥3	0.94 (0.89–0.96)	0.84	0.89	≥F3: 0.84
	Stage 4	0.92 (0.80–0.94)	0.94	0.81	F4: 0.82
Hepatitis B (n=81)	≥1	0.89 (0.60–0.99)	0.92	0.76	
	≥2	0.94 (0.71–0.99)	0.95	0.83	
	≥3	0.97 (0.88–0.99)	0.94	0.90	
	Stage 4	-	-	-	
NAFLD (n=115)	≥1	0.89 (0.81–0.97)	0.78	0.83	
	≥2	0.90 (0.79–0.93)	0.82	0.82	
	≥3	0.94 (0.91–0.98)	0.89	0.84	
	Stage 4	0.90 (0.64–0.94)	0.90	0.76	
Alcoholic liver disease (n=21)	≥1	0.76 (0.63–0.90)	0.57	0.89	
	≥2	0.81 (0.58–0.96)	0.70	0.77	
	≥3	0.93 (0.88–0.96)	0.92	0.81	
	Stage 4	-	-	-	

[Abbreviations: AUROC-Area under receiver operating curve, NAFLD-Nonalcoholic fatty liver disease]

* This represents the comparison of diagnostic performance of MRE between subgroups (HCV vs. HBV vs. NAFLD vs. alcoholic liver disease) for each corresponding fibrosis stage (any, significant, advanced fibrosis and cirrhosis). To compare the difference of AUROCs between subgroups, we used one-way ANOVA interaction test for comparisons with more than two estimates.