HHS Public Access

Author manuscript

Clin Gastroenterol Hepatol. Author manuscript; available in PMC 2020 March 01.

Published in final edited form as:

Clin Gastroenterol Hepatol. 2019 March ; 17(4): 630–637.e8. doi:10.1016/j.cgh.2018.05.059.

Magnetic Resonance vs Transient Elastography Analysis of Patients With Non-alcoholic Fatty Liver Disease: a Systematic Review and Pooled Analysis of Individual Participants

Cynthia Hsu[#1], **Cyrielle Caussy**[#1,2], **Kento Imajo**[3], **Jun Chen**[4], **Siddharth Singh**[1], **Kellee Kaulback**[5], **Minh-Da Le**[1], **Jonathan Hooker**[6], **Xin Tu**[7], **Ricki Bettencourt**[1], **Meng Yin**[4], **Claude B. Sirlin**[6], **Richard L. Ehman**[4], **Atsushi Nakajima**[3], and **Rohit Loomba**[1,8,9]

[1]Nonalcoholic Fatty Liver Disease Research Center, Department of Medicine, University of California at San Diego, La Jolla, California, USA

[2]Université Lyon 1, Hospices Civils de Lyon, Lyon, France

[3]Department of Gastroenterology, Yokohama City University Graduate School of Medicine, Yokohama, Japan

[4]Department of Radiology, Mayo Clinic, Rochester, Minnesota, USA

Please address correspondence to: Rohit Loomba, MD, MHSc, ACTRI Building, 1W202, 9452 Medical Center Drive, La Jolla, CA 92037, Ph: 858-246-2201, Fax: 858-246-2255, roloomba@ucsd.edu Web: http://fattyliver.ucsd.edu.

Author contributions:

Cynthia Hsu: analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission

Cyrielle Caussy: study design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission.

Kento Imajo: data collection, critical revision of the manuscript, approved final submission.

Jun Chen: data collection, critical revision of the manuscript, approved final submission.

Siddarth Singh: analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission

Kellee Kaulback: analysis of data, drafting of the manuscript, critical revision of the manuscript, approved final submission

Minh-Da Le: data collection and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission

Jonathan Hooker: data collection, imaging analysis, critical revision of the manuscript, approved final submission

Xin Tu: statistical analysis, critical revision of the manuscript, approved final submission

Ricki Bettencourt: statistical analysis, critical revision of the manuscript, approved final submission

Meng Yin: data collection, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, approved final submission.

Claude B. Sirlin: analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, obtained funding, study supervision, approved final submission

Richard L. Ehman: analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, obtained funding, approved final submission

Atsushi Nakajima: analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, obtained funding, approved final submission

Rohit Loomba: study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, obtained funding, study supervision, approved final submission

All authors approved the final version of this article.

Conflict of interests: Dr. Sirlin consults, advises, and is on the speakers' bureau for Bayer. He received grants from GE Healthcare. The Mayo Clinic and authors JC, MY and RLE have intellectual property rights and a financial interest through receipt of royalties and equity from licensing of MRE technology. Author RLE serves as uncompensated chief executive officer of Resoundant Inc. (Rochester, MN), majority-owned by Mayo Clinic. Mayo Clinic authors have control of the data collected at Mayo. The Mayo part of research has been conducted under the oversight of, and in compliance with, the Mayo Clinic Conflict of Interest Review Board. There are no conflicts of interest to declare for the other authors

[5]Research in Practice, Toronto, Ontario, Canada

[6]Liver Imaging Group, Department of Radiology, University of California at San Diego, La Jolla, California, USA

[7]Division of Biostatistics and Bioinformatics, Department of Family Medicine and Public Health, University of California at San Diego, La Jolla, California, USA

[8]Division of Gastroenterology, Department of Medicine, University of California at San Diego, La Jolla, California, USA

[9]Division of Epidemiology, Department of Family and Preventive Medicine, University of California at San Diego, La Jolla, California, USA

[#] These authors contributed equally to this work.

Abstract

Background & Aims: Magnetic resonance elastography (MRE) and transient elastography (TE) are noninvasive techniques for detection of liver fibrosis. Single-center studies have compared the diagnostic performance of MRE vs TE in patients with nonalcoholic fatty liver disease (NAFLD). We conducted a pooled analysis of individual participant data from published studies to compare the diagnostic performance of MRE vs TE for staging of liver fibrosis in patients with NAFLD, using liver biopsy as reference.

Methods: We performed a systematic search of publication databases, from 2005 through 2017. We identified 3 studies of adults with NAFLD who were assessed by MRE, TE, and liver biopsy. In a pooled analysis, we calculated the cluster-adjusted area under the curve (AUROC) of MRE and TE for the detection of each stage of fibrosis. AUROC comparisons between MRE and TE were performed using the Delong test.

Results: Our pooled analysis included 230 participants with biopsy-proven NAFLD with mean age of 52.2±13.9 years and a body mass index of 31.9±7.5 kg/m2. The proportions of patients with fibrosis of stages 0, 1, 2, 3, and 4 were: 31.7%, 27.8%, 15.7%, 13.9%, and 10.9%, respectively. The AUROC of TE vs MRE for detection of fibrosis stages ≥1 was 0.82 (95% CI, 0.76–0.88) vs 0.87 (95% CI, 0.82–0.91) (*P*=.04); for stage≥ 2 was 0.87 (95% CI, 0.82–0.91) vs 0.92 (95% CI, 0.88–0.96) (*P*=.03); for stage ≥3 was 0.84 (95% CI, 0.78–0.90) vs 0.93 (95% CI, 0.89–0.96) (*P*=.001); for stage ≥ 4 was 0.84 (95% CI, 0.73–0.94) vs 0.94 (95% CI, 0.89–0.99) (*P*=.005).

Conclusion: In a pooled analysis of data from individual participants with biopsy-proven NAFLD, we found MRE to have a statistically significantly higher diagnostic accuracy than TE in detection of each stage of fibrosis. MRE and TE each have roles in detection of fibrosis in patients with NAFLD, depending upon the level of accuracy desired.

Keywords

fibrosis measurement; comparison; diagnosis; BMI; Magnetic resonance elastography; Fibroscan; Transient Elastography; NAFLD

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is the most prevalent etiology of chronic liver disease worldwide affecting approximately 25% of the population worldwide[1,2]. Several studies have demonstrated that the presence of liver fibrosis is the most important predictor of mortality in NAFLD. Indeed, the risk of all-cause mortality is positively correlated with fibrosis stage and the risk of liver-related mortality increases exponentially with increase in fibrosis stage [3–5]. Therefore, the accurate detection and staging of liver fibrosis is important for the management of patients with NAFLD.

Currently, liver biopsy is considered the gold standard for the diagnosis and staging of liver fibrosis. Although liver biopsy is considered as a safe procedure, it may be associated with adverse effects including pain, infection, bleeding, and extremely rarely, even, death [6–8]. In addition, it is an expensive procedure requiring logistic coordination making its use unsuitable for the screening at the level of the population [9]. In addition, scoring and interpretation of liver biopsy are characterized by significant inter- and intra-observer variability [7,9–11].

Transient elastography (TE; FibroScan, Echosens, Paris, France) and magnetic resonance elastography (MRE) are two non-invasive methods used for assessing liver fibrosis in NAFLD patients. TE is an ultrasound-based imaging technique that assesses liver stiffness by measuring the speed of acoustic shear waves passing through the liver [12]. The liver stiffness measurement (LSM) using TE is correlated with stages of fibrosis especially in advanced fibrosis and cirrhosis [13,14]. Advantages of TE are a rapid, bed-side LSM and the development of an XL probe has significantly reduced the high failure rate observed in patients with BMI > 28 kg/m^2, which was observed with an M probe [15–17]. MRE is an MRI-based technique that images the propagation of acoustic shear waves in the liver and applies a mathematical algorithm to compute cross-sectional images displaying the magnitude of the complex shear modulus of liver tissue [18–20]. MRE has been shown to accurately diagnose fibrosis in NAFLD patients [21,22] and to be accurate and effective in patients with obesity [18–20]. However, an MRI-based technique like MRE is often considered as more expensive than an ultrasound-based technique such as TE.

Several studies comparing MRE and TE have found MRE to be more accurate than TE in identifying liver fibrosis[20,23,24]. However, these single center studies included participants with heterogeneous range of BMI, and different frequencies of use for M and XL probes for TE assessment. Finally, considerable differences in the identified thresholds for the dichotomized stage of fibrosis need to be addressed. Therefore, by performing a systematic review and pooled analysis of patient-level data from individual studies, we sought to address these discrepancies and compare the diagnostic performance of MRE and TE for detection of individual fibrosis stage(s) in patients with NAFLD, through a collaborative individual participant data (IPD) pooled analysis.

MATERIAL AND METHODS

This systematic review and collaborative IPD pooled analysis was conducted using an a priori established protocol and reported according to the Preferred Reporting Items for Systematic Reviews and MetaAnalyses guidelines [25] and recommendations from Riley et al. [26] Supplemental Table 1. This study was exempt from ethical approval as the analysis included only de-identified data, and all individual studies had received local ethics approval.

Search strategy

To identify all relevant articles comparing the diagnostic performance of MRE and TE for staging liver fibrosis in patients with NAFLD using the liver biopsy as a reference, the search strategy from a previously published technical review on the diagnostic performance of TE and MRE in diagnosing liver fibrosis by the American Gastroenterological Association was adopted [14]. An experienced medical librarian and study investigators subsequently updated the initial search including publication from January 01, 2005 to January 24, 2016, using a combination of controlled vocabulary supplemented with keywords, to include publications from January 01, 2016 to September 05, 2017. Studies were included if they met the following inclusion criteria: (1) liver stiffness assessment by both MRE and TE, (2) used liver biopsy as the gold standard, (3) reported fibrosis using a comparable liver biopsy staging system, and (4) included adult NAFLD patients (≥18 years). Inclusion was not otherwise restricted by study size or publication type. The flowchart summarizing the study identification and selection is provided in Supplemental Figure 1. Details of the search strategy and data abstraction are available in the Supplemental Data.

Statistical analysis

Patients' demographic data, laboratory, and imaging data were summarized with mean and standard deviation for continuous variables and with numbers and percentages for categorical variables. Mean and frequency were compared using an independent samples t-test, Wilcoxon Rank Sum Test or Chi-square test or Fisher's Exact Test, where appropriate.

In order to compare the diagnostic performance of MRE and TE, the participants with unreliable TE measurement defined in the pre-specified meta-analysis protocol as number of valid measurement <10 and/or IQR/median LSM >30% [27] were excluded from the analysis Supplemental Figure 2. Receiver operating characteristic (ROC) curve analyses were used to compare the performances of MRE versus TE for the diagnosis of dichotomized fibrosis stages using liver biopsy as a reference. The cluster-adjusted area under the ROC curves (AUROCs) were calculated using predicted probabilities from a random effects model. For each ROC analysis, the area under the ROC curve (AUROC), the optimal thresholds, and the following performance parameters were calculated: sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The optimal threshold and threshold for the prespecified sensitivity and specificity of 90% were determined by pooling the IPD across the included studies. The Delong test was used to compare the AUROCs of MRE versus TE for diagnosing fibrosis[28]. All statistical analyses were performed using SAS 9.4 (SAS Institute, Cary, NC), MedCalc 17.9 (Ostend, Belgium), or R 3.4.2 (The R

Foundation for Statistical Computing). A two-tailed P value ≤0.05 was considered statistically significant.

RESULTS

A total of 315 studies were identified using our primary search strategy and of these, three studies met our inclusion criteria [20,23,24]. After contacting the primary and/or corresponding authors of these studies, the IPD were obtained from all three studies. The study identification and selection process flowchart is shown in Supplemental Figure 1.

Characteristics and quality of included studies

All three studies included in this pooled analysis were prospective. Two studies were performed in the United States at different centers[20,24] while the third was performed in Japan[23]. These studies were at low risk of bias and had a QUADAS score ≥ 12 Supplemental Table 2. One study from Chen et al. included participant with several etiologies of chronic liver disease [20] and one study from Imajo et al. included NAFLD participants and non-NAFLD controls [23] while the study from Park et al. included only participants with NAFLD [24]. Only IPD from participants with NAFLD were collected for this pooled analysis.

The IPD from a total of 318 patients with biopsy-proven NAFLD were collected from the three studies. Based upon a pre-specified protocol, we excluded participant with unavailable MRE or TE measurement or with and unreliable LSM as defined by number of valid measurements <10 and/or IQR/median LSM >30%[27]. Out of the initial 318 IPD collected, a total of 230 participants were included in the pooled analysis, as presented in Supplemental Figure 2. The median time between biopsy and TE was 40.5 days and median time between biopsy and MRE was 34.5 days. Baseline characteristics, including demographics, clinical, biological, fibrosis stage, and imaging data, are presented in Table 1.

Comparison of MRE and TE for the diagnosis of liver fibrosis

The distribution of liver stiffness measurement by TE and MRE across the stage of fibrosis in the pooled cohort is illustrated in Figure 1. The direct comparison of the cluster-adjusted AUROC of TE and MRE demonstrated that MRE is more accurate than TE for diagnosing each dichotomized stage of fibrosis Figure 2. The detailed diagnostic performance of TE versus MRE for the detection of each dichotomized stage of fibrosis and the optimal thresholds are provided in **Table 2 and Figure 3** and the ROC curves provided in Supplemental Figure 3. MRE and TE images of 5 representative patients with stage of fibrosis 0, 1, 2, 3, and 4 based on the optimal threshold are shown in Figure 4.

Sensitivity analyses

In a multi-adjusted model including age, sex, BMI, type of probe and test to biopsy time, the diagnostic performance of VCTE and MRE remained similar and MRE had a significant higher diagnostic accuracy than TE for the detection of individual stages of fibrosis Supplemental Table 3.

Depending on the context of use, MRE and TE can be used either to rule in or rule out the stage of fibrosis. Therefore, sensitivity analyses were conducted to further assess the diagnostic accuracy and thresholds of MRE and TE when the sensitivity or the specificity was fixed at 90% **Figure 3.** The detailed thresholds and diagnostic performance at a high sensitivity fixed at 90%, to increase the probability to identify all individuals with a specific stage of fibrosis and at a high specificity fixed at 90%, to increase the probability to identify only individuals with a specific stage of fibrosis are detailed in Supplemental Table 4. Lower liver stiffness value of both MRE and TE for fixed sensitivity at 90% ruled out the presence of advanced fibrosis securely (negative predictive value of MRE: 95.8% and of TE: 94.3%) whereas higher liver stiffness values for fixed specificity of 90% has a lower performance to rule in the presence of advanced fibrosis (positive predictive value of MRE: 70.3% and TE: 65.4%) Supplemental Table 4.

DISCUSSION

Main findings

In this pooled analysis of data from individual participants with biopsy-proven NAFLD, MRE demonstrated a higher diagnostic accuracy than TE for the detection of individual stages of fibrosis. Additionally, this study establishes the optimal threshold of MRE and TE for the detection of any (≥stage 1), significant ≥ stage 2), advanced fibrosis (≥ stage 3) and cirrhosis (≥ stage 4). This study establishes the optimal MRE thresholds of 2.61, 2.97, 3.62, and 4.69 kPa, respectively, and the optimal TE thresholds of ***6.2,*** 7.6, 8.8, and 11.8 kPa, respectively. Finally, we provide thresholds of MRE and TE at a fixed sensitivity or specificity of 90% to rule in or rule out the disease stage, respectively. While low liver stiffness excludes the presence of advanced fibrosis securely, high liver stiffness values are less performant to diagnose the presence of advanced fibrosis as previously reported for high TE measurement in the study performed by Karlas et al. [29] Therefore, the presence of elevated liver stiffness would usually require further assessment to confirm the diagnosis of advanced fibrosis irrespective of the non-invasive method used. These results have important implications in developing an optimal approach for non-invasive assessment of the severity of NAFLD. Depending on the desirability of accuracy, the need to identify all the individuals or only individuals with a specific disease stage and availability of these two non-invasive modalities, this study provides much needed information that will help define the optimal strategies for the management of NAFLD patients. In situations where accurate staging of fibrosis is important, MRE may be preferable to TE due to the higher accuracy of MRE, whereas TE may be preferable in routine clinical care to rule out or rule in advanced fibrosis depending on the risk or the population screened. However, further cost-effectiveness studies are needed to draw more definite conclusions.

In context of published literature

The detection of liver fibrosis is a clinically important feature of NAFLD as liver fibrosis and especially the presence of advanced fibrosis (stages 3–4), have been shown to be associated with increased of liver-related morbidity and mortality in NAFLD patients [3–5].

This study confirms the accuracy of MRE for the diagnosis of liver fibrosis reported in previous metaanalyses. [30,31] Similarly, our study provides comparable diagnostic performance of TE for the diagnosis of fibrosis as provided by the meta-analysis performed by Friedrich-Rust et al. [32] This is the first pooled analysis to use individual participant-level data to compare the diagnostic performance of MRE and TE head-to-head in individuals with biopsy-proven NAFLD. This study confirms that MRE has a significantly higher accuracy than TE for the diagnosis of all individual stages of fibrosis.

Although, the previously published studies comparing the performance of MRE and TE have shown a higher diagnostic accuracy of MRE compared to TE for the diagnosis of fibrosis, discrepancies needed to be addressed. In the study by Park et al., MRE was significantly superior to TE for diagnosing any fibrosis only [24], in the study by Imajo et al., MRE was significantly superior to TE only for diagnosing significant fibrosis (stage ≥) and cirrhosis (stage ≥)[23]; whereas the study by Chen et al. showed MRE to be significantly superior to TE only for diagnosing significant (stage >2) and advanced fibrosis (stage ≥3) [20]. Additionally, the optimal thresholds for both MRE and TE varied across the three studies ranging from 3.35 to 6.7 kPa for MRE and from 6.9 to 14.6 kPa for TE. The different magnetic fields strength used for MRE measurement, (3T in Imajo et al. and by Park et al. study and 1.5T in Chen et al.), could not have impacted these results. Indeed, the liver stiffness measurement does not depend on magnetic field strength (3T or 1.5T) but depends on the frequency of the shear wave generating by the acoustic passive driver [33] which was standard at 60-Hz across all studies. However, the heterogeneity of the participants across the three studies, including different range of BMI and use of M and XL probes, may account for these discrepancies. We have recently demonstrated that the discordancy between MRE and TE for the assessment of fibrosis stage increases with increasing BMI [34]. In addition, Vuppalanchi et al. have shown that TE is highly reliable and reproducible, with low failure rates, when the appropriate probe is used [35]. Finally, the optimal thresholds for both MRE and TE varied across the three studies and more accurate thresholds needed to be determined. Therefore, by analyzing pooled data from individual participants with biopsy-proven NAFLD, this study allows us (1) to compare the diagnostic performances of MRE and TE; (2) to provide more accurate thresholds for the diagnosis of liver fibrosis, in a larger cohort that encompasses a wide range of BMI and TE measurements using M and XL probes when appropriate.

Strengths and limitations

The key novelty of this study is the utilization of participant level data to perform this meta-analysis, which overcomes limitations inherent in meta-analyses of aggregate data. Additionally, we performed a comprehensive and systematic literature search with well-defined inclusion criteria, carefully evaluated study quality and excluded redundant studies, applied standardized criteria across studies to define unreliable liver stiffness based upon a pre-specified meta-analysis protocol, and minimized spectrum bias by including only patients with biopsy-proven NAFLD and by excluding data from non-NAFLD controls or of etiology other than NAFLD. In addition, the thresholds of MRE and TE for the diagnosis of each stage of fibrosis derived from this IPD meta-analysis are more accurate and reflect the clinical context of use of MRE and TE.

However, the following limitations need to be acknowledged. The three studies included in this metaanalysis were conducted in advanced tertiary care centers and the majority of the patients were at high risk of advanced stage of NAFLD. Thus the generalizability of these findings in other clinical settings is unknown. The liver biopsy was the reference standard as it is considered as the gold standard for the diagnosis and staging of liver fibrosis. However, the scoring and interpretation of liver biopsy are characterized by significant inter- and intra-observer variability and it was not possible to obtain a centralized reading of the biopsy by a blinded pathologist [7,9–11]. Thus, potential bias regarding interpretation of the liver biopsy cannot be excluded. However, this bias would have impacted the diagnostic accuracy of both imaging modalities. Furthermore, the median time between biopsy and TE was 40.5 days and median time between biopsy and MRE was 34.5 days. A meta-analysis of paired liver biopsy studies has shown that the rate of fibrosis progression is slow, with an average progression of one stage to take 14.3 years in patients with NAFL and 7.1 years in patients with NASH [36]. Therefore, the test-to-biopsy time of the study is reasonable as the stage of fibrosis is unlikely to change within a year.

In addition, as the aim of this study was to compare the diagnostic performance of MRE versus TE for the staging of liver fibrosis in NAFLD, participants with unreliable TE measurement needed to be excluded especially to avoid potential bias due to inadequate type of probe use in obese individuals. Therefore, the analysis was not conducted in intention-to-diagnose. Finally, after multivariable adjustment including age, gender, BMI, type of probe and test-to-biopsy time the results remained consistent and statistically significant.

Implication for future research

By performing this IPD meta-analysis comparing MRE and TE head-to-head, we found that MRE is superior to TE for the diagnosis of fibrosis at every dichotomized stage in patients with biopsy-proven NAFLD. Additionally, we propose thresholds for MRE and TE for diagnosis of each level of fibrosis using data from heterogeneous populations, increasing the generalizability of the results. Further cost-effectiveness studies are needed to determine the optimal screening strategy for the diagnosis of liver fibrosis in NAFLD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: RL is supported in part by the American Gastroenterological Association (AGA) Foundation - Sucampo - ASP Designated Research Award in Geriatric Gastroenterology and by a T. Franklin Williams Scholarship Award; Funding provided by: Atlantic Philanthropies, Inc, the John A. Hartford Foundation, OM, the Association of Specialty Professors, and the American Gastroenterological Association and grant K23-DK090303. CS and RL serve as co-PIs on the grant R01-DK106419. CC is supported by grants from the Société Francophone du Diabète (SFD), the Philippe Foundation and Monahan Foundation under the Fulbright program. MY is supported by NIH grant EB017197. RLE is supported by NIH grant EB001981.

Appendix

Background

The non-invasive staging of liver fibrosis is important for the management of patients with nonalcoholic fatty liver disease. Individual patient data from three single center studies when pooled will provide more robust estimates of comparative efficacy of magnetic resonance elastography and transient elastography.

Findings

The systematic review and pooled analysis of individual participant data from 3 published studies demonstrates a higher diagnostic accuracy of magnetic resonance elastography compared to transient elastography for the detection of all individual stages of fibrosis.

Implication for patient care

This study will contribute to determine the optimal approach for the non-invasive detection of liver fibrosis. In addition, it provides thresholds of magnetic-resonance-elastography and transient elastography for the detection of dichotomized stage of fibrosis.

Systematic Review Protocol

Primary Outcome:

Comparison of the diagnostic performance of MRE and VCTE for the diagnosis of any fibrosis (stage≥1), significant fibrosis (≥ stage 2), advanced fibrosis (≥ stage 3) and cirrhosis (≥ stage 4) using the liver biopsy as the reference standard.

Search Strategy

An updated comprehensive search of several electronic databases, including Ovid MEDLINE and Ovid EMBASE, will be conducted. The search strategy will be designed and conducted by an experienced medical librarian with input from study investigators. Controlled vocabulary supplemented with keywords were used to search for cohort studies including NAFLD patient who underwent Magnetic Resonance Elastography (MRE), transient elastrography (Fibroscan) and a liver biopsy as reference standard.

Example search

#	Searches	Results
1	exp Non-alcoholic Fatty Liver Disease/ use ppez	7276
2	exp nonalcoholic fatty liver/ use emczd	26745
3	(fatty liver or nafld or steatosis).ti,ab.	79205
4	or/1–3	87331
5	exp Elasticity Imaging Techniques/	25075
6	elastography/ use emczd or exp elastography/ use emczd	18950

7	(elastogra* or elastometr* or sono?elastogra* or vibro?acoustogra* or fibroscan or fibro scan or (elasticity adj3 imag*) or (measure* adj2 (liver adj2 stiff*))).ti,ab.	18768
8	or/5–7	30905
9	4 and 8	2498
10	animals/ not (humans/ and animals/)	5803078
11	9 not 10	2489
12	Case Reports/ or Comment.pt. or Editorial.pt. or Letter.pt. or Congressess.pt.	4862875
13	Case Report/ or Comment/ or Editorial/ or Letter/ or conference abstract.pt.	9495723
14	11 not (12 or 13)	1473
15	limit 14 to (english language and yr="2016 -Current")	550
16	remove duplicates from 15	345

Inclusion criteria

1. Liver stiffness assessment by both MRE and TE
2. Use of liver biopsy as the gold standard for diagnosis of liver fibrosis
3. Reported fibrosis using a comparable liver biopsy staging system
4. Adult (≥18 years) patients with NAFLD

Exclusion criteria

1. Non-human studies
2. Studies not written in the English language
3. The diagnostic performance of MRE or TE was not assessed
4. Liver biopsy was not performed
5. Participants with etiologies other than NAFLD were included in the study and data from NAFLD participant specifically was not available
6. Sufficient IPD could not be obtained despite multiple attempts to contact study investigators

Study Identification, Screening, and Data

Two authors will independently review all titles and abstracts for relevance and inclusion. In case of disagreement between the two authors, then a consensus will be made by a third more senior author. The full text of identified articles will be reviewed by each of the two authors and reviewed by the third author for completeness and accuracy. Data

1. Study Characteristics: Primary author, publication year, time period of study, study location, total number of patients and number of patients with NAFLD.
2. Cohort Characteristics: age at time of index test, gender, body mass index (BMI), fibrosis stage on liver biopsy (and classification system used), biological data

(AST, ALT, alkaline phosphatase, albumin, INR, platelet count, GGT), and interval between TE, MRE and liver biopsy

3. TE assessment including liver stiffness measurement (LSM), interquartile range (IQR) of liver stiffness, IQR to LSM ratio, number of valid measurement, type of probes used (M or XL), unreliable measurement defined as number of valid measurement <10 and/or IQR/median LSM >30%[1].

4. Liver stiffness assessment using MRE and the MRE technique used

Quality Assessment

The quality of the study included will be assessed independently by 2 authors using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) questionnaire, developed to assess the internal and external validity of diagnostic accuracy studies included in systematic reviews [2]. Each item reported as “yes” will be scored 1 point, item reported as “no” or as “unclear” will be scored 0 point.

Search strategy

To identify all relevant articles comparing the diagnostic performance of MRE and TE for staging liver fibrosis in patients with NAFLD using the liver biopsy as a reference, the search strategy from a previously published technical review on the diagnostic performance of TE and MRE in diagnosing liver fibrosis by the American Gastroenterological Association was adopted [16]. An experienced medical librarian and study investigators subsequently updated the initial search including publication from January 01, 2005 to January 24, 2016, using a combination of controlled vocabulary supplemented with keywords, to include publications from January 01, 2016 to September 05, 2017. Any duplicates that were found, including duplicates to the initial search, were removed. Subsequently, two investigators (C.L.H. and M.L.) independently reviewed the titles and abstracts of the identified studies and excluded studies that did not address the research question of interest, based on pre-specified inclusion and exclusion criteria. The full text of the remaining articles was reviewed by the two investigators independently to determine whether they contained relevant information. Conflicts in study selection at this stage were resolved by consensus, referring back to the original article in consultation with a third reviewer (C.C.). This search was supplemented with a recursive search of the bibliographies of recently published systematic reviews on this topic to identify any additional studies. We also consulted with experts in the field to identify additional published and unpublished primary studies.

The corresponding author of eligible study was then contacted by e-mail, and provided with detailed objectives of the pooled analysis, background information on the IPD pooled analysis, and an Excel (Microsoft, Richmond, WA) document containing a data collection file for input of individual patient results for the project.

Data abstraction

The IPD from each study was abstracted using a standardized data abstraction form including: (a) study characteristics: primary author, location, time period of study, number of patients; (b) patient characteristics: age at time of index test, gender, body mass index (BMI), fibrosis stage on liver biopsy (and classification system used), biological data (AST, ALT, alkaline phosphatase, albumin, INR, platelet count, GGT); and interval between TE, MRE and liver biopsy; (c) TE assessment including liver stiffness measurement (LSM), interquartile range (IQR) of liver stiffness, IQR to LSM ratio, number of valid measurements, type of probes used (M or XL), unreliable measurement defined in the pre-specified meta-analysis protocol as number of valid measurement <10 and/or IQR/median LSM >30%[29]; and (d) liver stiffness assessment using MRE and the MRE technique used.

The quality of the included studies was assessed independently by 2 investigators (C.C, C.L.H) using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) questionnaire, which was developed to assess the internal and external validity of diagnostic accuracy studies included in systematic reviews [30]. This 14-item tool have been designed to identify important element in the design of diagnostic accuracy studies including blinding of the observer, patient spectrum, handling of uninterpretable results, verification bias, and explanation for withdrawal from the study. Each item was scored as “yes” when reported (1 point), “no” when not reported or as “unclear” if the information was not reported in the **study (0 point)** Supplemental Table 2.

Abbreviations

ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUROC	Area under the Receiver Operating Characteristic
BMI	body mass index
CI	confidence interval
GGT	Gamma-Glutamyl Transferase
INR	International Normalized Ratio
IPD	individual participant data
IQR	interquartile range
LSM	liver stiffness measurement
MRE	magnetic resonance elastography
NAFLD	nonalcoholic fatty liver disease
NASH	non-alcoholic steatohepatitis

NPV	negative predictive value
PPV	positive predictive value
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
TE	transient elastography
VCTE	vibration controlled transient elastography

REFERENCES

1. Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64(1): 73–84. [PubMed: 26707365]
2. Loomba R, Sanyal AJ. The global NAFLD epidemic. Nat Rev Gastroenterol Hepatol. 2013;10(11): 686–90. [PubMed: 24042449]
3. Angulo P, Kleiner DE, Dam-Larsen S, 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(2):389–97.e10 Epub 2015/05/04. [PubMed: 25935633]
4. Ekstedt M, Hagstrom H, Nasr P, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology. 2015;61(5):1547–54. Epub 2014/08/16. [PubMed: 25125077]
5. Dulai PS, Singh S, Patel J, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology. 2017;65(5):1557–65. [PubMed: 28130788]
6. Pasha T, Gabriel S, Therneau T, et al. Cost-effectiveness of ultrasound-guided liver biopsy. Hepatology. 1998;27(5):1220–6. [PubMed: 9581674]
7. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology. 2009;49(3):1017–44. [PubMed: 19243014]
8. Foster G, Goldin R, Main J, et al. Management of chronic hepatitis C: clinical audit of biopsy based management algorithm. Bmj. 1997;315(7106):453–8. [PubMed: 9284661]
9. Bravo AA, Sheth SG, Chopra S. Liver biopsy. N Engl J Med. 2001;344(7):495–500. [PubMed: 11172192]
10. 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(10):2614–8. [PubMed: 12385448]
11. Sumida Y, Nakajima A, Itoh Y. Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. World J Gastroenterol. 2014;20(2):475–85. [PubMed: 24574716]
12. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol. 2003;29(12):1705–13. [PubMed: 14698338]
13. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology. 2010;51(2):454–62. [PubMed: 20101745]
14. Singh S, Muir AJ, Dieterich DT, et al. American Gastroenterological Association Institute Technical Review on the Role of Elastography in Chronic Liver Diseases. Gastroenterology. 2017;152(6):1544–77. [PubMed: 28442120]
15. de Ledinghen V, Wong VW, Vergniol J, et al. Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: comparison between M and XL probe of FibroScan(R). J Hepatol. 2012;56(4):833–9. [PubMed: 22173167]

16. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol. 2012;107(12):1862–71. [PubMed: 23032979]
17. Myers RP, Pomier-Layrargues G, Kirsch R, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. Hepatology. 2012;55(1):199–208. [PubMed: 21898479]
18. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver: technique, analysis, and clinical applications. J Magn Reson Imaging. 2013;37(3):544–55. [PubMed: 23423795]
19. Yin M, Glaser KJ, Talwalkar JA, et al. Hepatic MR Elastography: Clinical Performance in a Series of 1377 Consecutive Examinations. Radiology. 2016;278(1):114–24. [PubMed: 26162026]
20. Chen J, Yin M, Talwalkar JA, et al. Diagnostic Performance of MR Elastography and Vibration controlled Transient Elastography in the Detection of Hepatic Fibrosis in Patients with Severe to Morbid Obesity. Radiology. 2017;283(2):418–28. [PubMed: 27861111]
21. Huwart L, Sempoux C, Vicaut E, et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology. 2008;135(1):32–40. [PubMed: 18471441]
22. 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. 2014;60(6):1920–8. [PubMed: 25103310]
23. Imajo K, Kessoku T, Honda Y, et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients with Nonalcoholic Fatty Liver Disease Than Transient Elastography. Gastroenterology. 2016;150(3):626–37e7. [PubMed: 26677985]
24. Park CC, Nguyen P, Hernandez C, et al. Magnetic Resonance Elastography vs Transient Elastography in Detection of Fibrosis and Noninvasive Measurement of Steatosis in Patients With Biopsy-Proven Nonalcoholic Fatty Liver Disease. Gastroenterology. 2017;152(3):598–607 e2. [PubMed: 27911262]
25. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and metaanalyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006–12. [PubMed: 19631508]
26. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010;340:c221. [PubMed: 20139215]
27. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol. 2008;48(5):835–47. [PubMed: 18334275]
28. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45. [PubMed: 3203132]
29. Karlas T, Petroff D, Sasso M, et al. Impact of controlled attenuation parameter on detecting fibrosis using liver stiffness measurement. Aliment Pharmacol Ther. 2018;47(7):989–1000. [PubMed: 29446106]
30. Singh S, Venkatesh SK, Wang Z, 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(3):440–51 e6. [PubMed: 25305349]
31. 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(1):239–47. [PubMed: 22278368]
32. 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(4):960–74. [PubMed: 18395077]
33. Venkatesh SK, Ehman RL. Magnetic resonance elastography of liver. Magn Reson Imaging Clin N Am. 2014;22(3):433–46. [PubMed: 25086938]
34. Caussy C, Chen J, Alquiraish MH, et al. Association Between Obesity and Discordance in Fibrosis Stage Determination by Magnetic Resonance vs Transient Elastography in Patients With Nonalcoholic Liver Disease. Clin Gastroenterol Hepatol. 2018.
35. Vuppalanchi R, Siddiqui MS, Van Natta ML, et al. Performance characteristics of vibration controlled transient elastography for evaluation of nonalcoholic fatty liver disease. Hepatology. 2018;67(1):134–44. [PubMed: 28859228]

36. Singh S, Allen AM, Wang Z, et al. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clin Gastroenterol Hepatol. 2015;13(4):643–54 e1–9; quiz e39–40. [PubMed: 24768810]

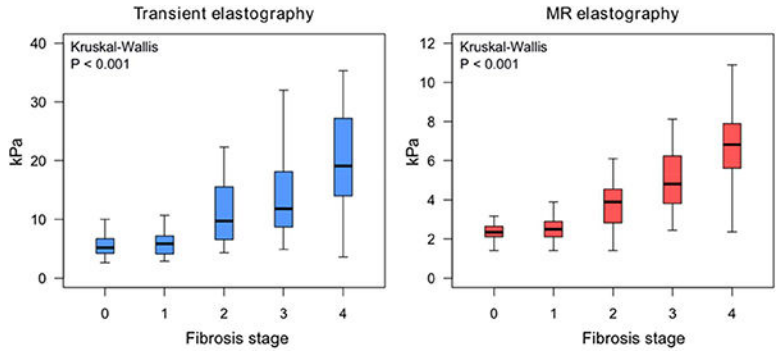

Figure 1.
Distribution of liver stiffness measurements by TE (A) and MRE (B).

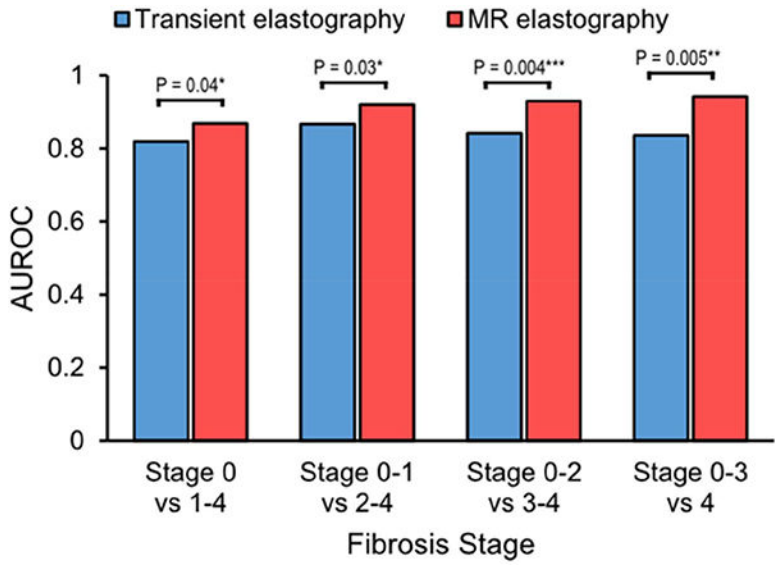

Figure 2.
Diagnostic accuracy of TE and MRE for diagnosing dichotomized stages of fibrosis.

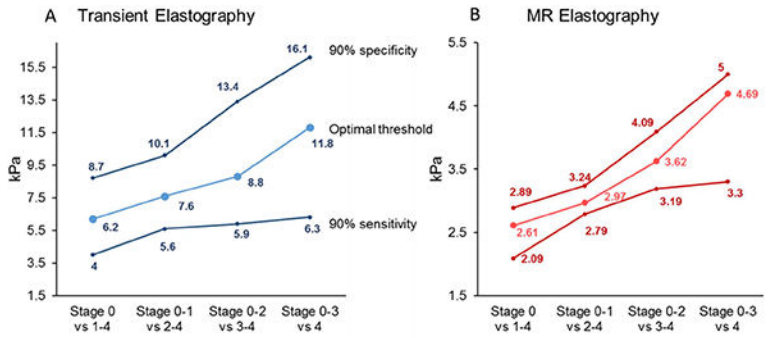

Figure 3.

Distribution of thresholds value for the diagnosis of dichotomized fibrosis stage of TE (A) or MRE (B). The lower line corresponds to the thresholds for a fixed sensitivity at 90%, line in the middle corresponds to the optimal threshold and the upper line corresponds to the thresholds for a fixed specificity at 90%.

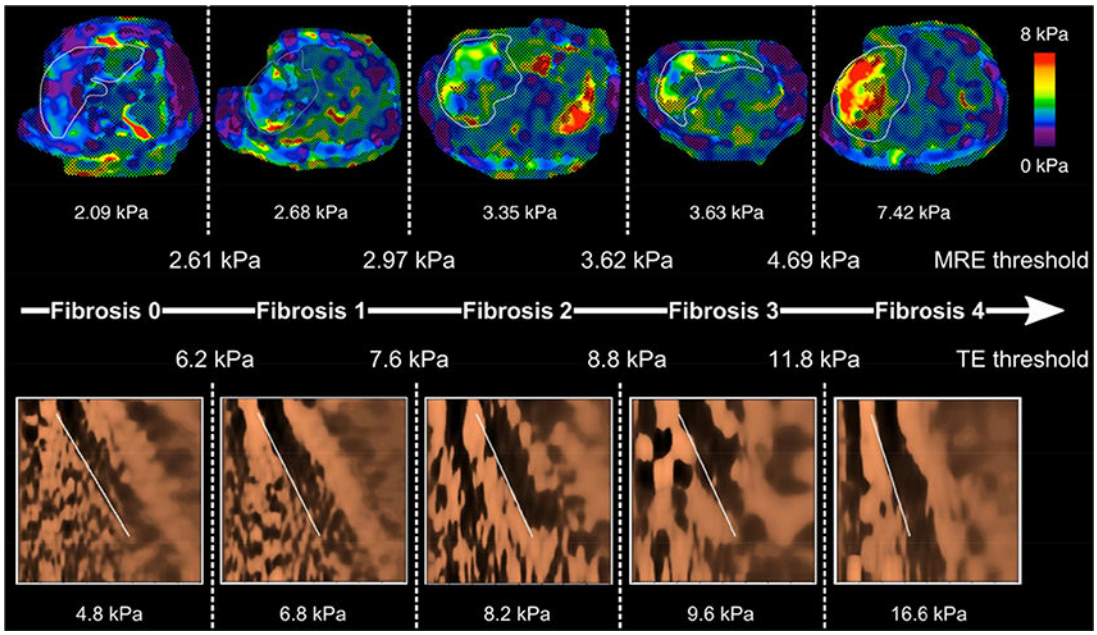

Figure 4.

MRE and TE images from representative patients with stage 0, 1, 2, 3, and 4 fibrosis, respectively.

Table 1

population characteristics

Characteristics	All	UCSD cohort	Japanese Cohort	Mayo Cohort	p-value*	Pairwise comparison
N	230	87	82	61		
Demographics						
Age (years)	52.2 ± 13.9	50.3 ± 14.5	56.6 ± 14.3	49.0 ± 11.1	**0.001**	α, Δ
Female, n (%)	124 (53.9%)	50 (57.5%)	33 (40.2%)	41 (67.2%)	**0.004**	α, Δ
Clinical						
Type 2 Diabetes, n (%)	67 (39.6%)	21 (24.1%)	46 (52.9%)	NA	**<0.001**	α
BMI (kg/m^2)	31.9 ± 7.5	30.3 ± 5.1	27.2 ± 4.2	40.5 ± 6.5	**<0.001**	α, β, Δ
Obesity (BMI⩾30), n (%)	122 (53.0%)	42 (48.3%)	21 (25.6%)	59 (96.7%)	**<0.001**	α, β, Δ
Biological data						
AST (U/L)	43.9 ± 46.8	35.2 ± 20.4	43.9 ± 24.0	62.6 ± 94.9	**0.026**	α, β
ALT (U/L)	57.3 ± 58.3	48.7 ± 42.7	57.6 ± 46.4	81.0 ± 104.5	0.197	
Alk P (U/L)	197.8 ± 158.0	76.9 ± 35.7	358.7 ± 111.8	91.2 ± 28.0	**<0.001**	α, β, Δ
Albumin (g/dL)	4.4 ± 0.4	4.5 ± 0.3	4.5 ± 0.5	4.2 ± 0.4	0.081	β, Δ
INR	1.1 ± 0.3	1.0 ± 0.1	1.0 ± 0.1	1.3 ± 0.9	0.156	β
Platelet count (10^9/L)	336.0 ± 167.1	237.9 ± 62.7	520.9 ± 118.5	215.9 ± 75.4	**<0.001**	α, Δ
GGT (U/L)	62.5 ± 63.9	50.8 ± 51.7	74.1 ± 72.5	NA	**<0.001**	α
Histology						
Fibrosis					**<0.001**	α, Δ
0	73 (31.7%)	40 (46.0%)	6 (7.3%)	27 (44.3%)		
1	64 (27.8%)	19 (21.8%)	29 (35.4%)	16 (26.2%)		
2	36 (15.7%)	11 (12.6%)	19 (23.2%)	6 (9.8%)		
3	32 (13.9%)	9 (10.3%)	20 (24.4%)	3 (4.9%)		
4	25 (10.9%)	8 (9.2%)	8 (9.8%)	9 (14.8%)		
Imaging data						
MRE (kPa)	3.5 ± 1.8	3.1 ± 1.3	4.2 ± 2.2	3.3 ± 1.7	**0.012**	α, Δ
VCTE (kPa)						
Median, median (IQR)	6.7 (7.3)	6.1 (4.8)	8.8 (10.7)	6.4 (4.8)	**0.005**	α
IQR, median (IQR)	1.0(1.5)	1.0(1.1)	1.2 (2.3)	1.0 (0.8)	**0.023**	α, Δ
IQR/M, median (IQR)	0.16 (0.09)	0.16 (0.09)	0.17 (0.10)	0.13 (0.10)	0.077	Δ
Probe size, n (%)					**<0.001**	α, β, Δ
Use of M probe	139 (60.4%)	44 (50.6%)	80 (97.6%)	15 (24.6%)		
Use of XL probe	91 (39.6%)	43 (49.4%)	2 (2.4%)	46 (75.4%)		

Continuous variables are expressed in mean with standard deviation in parentheses or median, unless otherwise noted as median with interquartile range (IQR) in parentheses or n (%).UCSD: University of California, San Diego; Japanese: Yokohama City University, NA: data not available.

*P-value determined by comparing characteristics across the 3 cohorts. Bold indicates significant P values <0.05. Pairwise comparison: α: UCSD vs Japanese cohort, β: UCSD vs Mayo cohort, Δ: Japanese vs Mayo cohort by Mann-Whitney-Wilcoxon test yielded significant p-value <0.05.

Table 2

Diagnostic performance of MRE and VCTE for the detection of liver fibrosis

Overall (n=230)	AUROC (95%CI)	Optimal Threshold (kPa)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	TE versus MRE P-value
Stage 1–4 (n=157) versus 0 (n=73)							
MRE	0.8685 (0.8230–0.9139)	2.61	71.3	72.6	84.9	54.1	**0.0394**
TE	0.8184 (0.7610–0.8759)	6.2	65.6	67.1	81.1	47.6	
Stage 2–4 (n=93) versus 0–1 (n=137)							
MRE	0.9193 (0.8826–0.9560)	2.97	84.9	85.4	79.8	89.3	**0.0265**
TE	0.8664 (0.8184–0.9144)	7.6	76.3	79.6	71.7	83.2	
Stage 3–4 (n=57) versus 0–2 (n=173)							
MRE	0.9295 (0.8948–0.9641)	3.62	82.5	83.2	61.8	93.5	**0.0004**
TE	0.8411 (0.7805–0.9016)	8.8	77.2	78.0	53.7	91.2	
Stage 4 (n=25) versus 0–3 (n=205)							
MRE	0.9423 (0.8920–0.9927)	4.7	80.0	85.9	40.8	97.2	**0.0046**
TE	0.8357 (0.7351–0.9364)	11.8	80.0	81.0	33.9	97.1	