European Radiology (2019) 29:1039–1047
https://doi.org/10.1007/s00330-018-5619-4

HEPATOBILIARY-PANCREAS

The development and validation of magnetic resonance elastography for fibrosis staging in primary sclerosing cholangitis

Kartik S. Jhaveri[1] · Hooman Hosseini-Nik[1] · Nima Sadoughi[1] · Harry Janssen[2] · Jordan J. Feld[2] · Sandra Fischer[3] · Ravi Menezes[1] · Angela C. Cheung[2]

Received: 24 April 2018 / Revised: 10 June 2018 / Accepted: 19 June 2018 / Published online: 26 July 2018

Abstract

Objectives To develop and internally validate MR elastography (MRE) quantified liver stiffness (LS) cut-off values for distinguishing early/moderate fibrosis from cirrhosis in primary sclerosing cholangitis (PSC) against non-invasive fibrosis test of vibration-controlled transient elastography (VCTE).

Methods Sixty-seven patients were enrolled prospectively at a tertiary care centre to undergo MRE and VCTE. MRE-quantified LS was calculated using three region-of-interest (ROI) methods: Trace, Average and Maximum. Each ROI method was compared with the reference standard of VCTE. Internal validation was performed with bootstrapping. Univariable and multivariable linear regression determined independent predictors for MRE-quantified LS and final Mayo Risk Score (MRS).

Results MRE-quantified LS by Trace ROI method had the highest sensitivity [87.5%; 95% confidence interval (CI), 66.0-96.8] and specificity (96.1%; 95%CI, 89.6-99.0) for distinguishing cirrhosis; and was the strongest predictor of final MRS (β, 0.44; 95% CI, 0.27-0.61). Alkaline phosphatase twice the normal upper limit (β, 1.55; 95% CI, 0.95-2.17), abnormal bilirubin (β, 1.27; 95% CI, 0.41-2.14) and thrombocytopaenia (β, 0.79; 95% CI, 0.12-1.46) were independent predictors of LS.

Conclusions MRE has a higher correlation with MRS than VCTE; and though MRE is possibly influenced by severe cholestasis and portal hypertension, MRE-quantified LS is an independent predictor of worse MRS.

Key Points

- *MRE is valid and reliable in assessing cirrhosis in PSC, and MRE-quantified Liver stiffness (LS) score was the strongest predictor of final Mayo Risk Score (MRS).*
- *Trace ROI performs best for distinguishing moderate fibrosis from cirrhosis and has the highest correlation with Mayo Risk Score (MRS).*
- *Cholestasis, hyperbilirubinaemia and portal hypertension may influence MRE LS score.*

Keywords Magnetic resonance imaging · Primary sclerosing cholangitis · Elasticity imaging techniques · Liver cirrhosis
Bile ducts

Abbreviations

AIH	Autoimmune hepatitis
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
APRI	Aspartate aminotransferase-to-platelet ratio
AST	Aspartate aminotransferase
AUROC	Area under the receiver operating characteristic
BMI	Body mass index
CI	Confidence interval
GGT	Gamma-glutamyltransferase
IBD	Inflammatory bowel disease
GRE	Gradient recalled echo
INR	International normalised ratio
IQR	Interquartile range
LS	Liver stiffness
MRCP	Magnetic resonance cholangiopancreaticogram
MRE	Magnetic resonance elastography

✉ Kartik S. Jhaveri
kartik.jhaveri@uhn.ca

[1] Joint Department of Medical Imaging, University of Toronto, University Health Network and Mount Sinai Hospital, 610 University Ave, 3-957, Toronto, Ontario M5G 2M9, Canada

[2] Toronto Centre for Liver Disease, University Health Network, University of Toronto, Toronto, Ontario, Canada

[3] Department of Laboratory Medicine and Pathology, University Health Network, University of Toronto, Toronto, Ontario, Canada

MRS	Mayo Risk score
Na-MELD	Sodium-model for end-stage liver disease
PBC	Primary biliary cirrhosis
PSC	Primary sclerosing cholangitis
PT	Prothrombin time
ROI	Region-of-interest
ULN	Upper limit of normal
VCTE	Vibration-controlled transient elastography

Introduction

Primary sclerosing cholangitis (PSC) is a chronic, progressive, cholestatic liver disease with bile duct stricturing [1]. Indications for liver transplantation include decompensated cirrhosis, recurrent cholangitis, refractory pruritus and early-stage malignancy not amenable to resection [2, 3]. Although up to 4% of PSC patients present with portal hypertension, 20% have cirrhosis at the time of diagnosis [4] and the majority develop decompensated cirrhosis within 10–15 years [5, 6]. Therefore, timely diagnosis of cirrhosis in PSC has implications for liver transplantation.

There is no accurate method to stratify PSC patients across the spectrum of disease severity. The Mayo Risk Score (MRS) is a prognostic score calculated based on age, variceal bleeding, serum albumin, aspartate aminotransferase (AST) and total bilirubin that can estimate transplant-free survival at 4 years and, thus, is reliable mostly in patients with advanced disease [1]. Liver biopsy has been the "gold standard" to stage liver fibrosis in chronic liver diseases; however, it is invasive and carries risks of bleeding and infection and also is limited by sampling issues due to disease heterogeneity in PSC [7]. Thus, it is no longer performed routinely for PSC at many centres and cannot be employed as a standard of reference for PSC studies. Non-invasive tools, such as vibration-controlled transient elastography (VCTE) and magnetic resonance elastography (MRE), are being increasingly utilised for quantification of hepatic fibrosis to replace liver biopsy, especially in patients with chronic HCV and HBV infections [8–14]. In contrast to VCTE, which only captures approximately 1% of hepatic volume [15], MRE provides liver stiffness (LS) maps that can encompass a major portion of the liver. This is especially important in PSC which unlike chronic viral hepatitis, primary biliary cirrhosis (PBC) and autoimmune hepatitis (AIH), causes heterogeneous injury and fibrosis [7, 16, 17]. Therefore, MRE may provide a more accurate measure of fibrosis in PSC. Our aim was to evaluate MRE as a non-invasive tool for staging fibrosis and distinguishing cirrhosis in PSC. We did not include other autoimmune liver diseases (i.e. PBC and AIH), as they are arguably more similar to viral, alcoholic and non-alcoholic fatty liver than PSC. Furthermore, from a clinical standpoint, neither PBC nor AIH would benefit as much as PSC from MRE, as PBC has a number of validated prognostic scores, and AIH patients require protocol biopsies to assess disease activity. The specific objectives of our prospective cohort study were to: (1) evaluate the performance characteristics of three region-of-interest (ROI) methods (Trace, Average, Maximum) for LS measurement methods in PSC with MRE; (2) develop and internally validate the optimal MRE-quantified LS cut-point between early/moderate fibrosis and cirrhosis; (3) determine the prognostic value of MRE as defined by the MRS.

Materials and methods

Population

This study was approved by the Institutional Research Ethics Board. Written consent was obtained. Consecutive patients were prospectively enrolled from a single tertiary liver centre from January 2014 to July 2016. Included patients were ≥18 years old, with a diagnosis of PSC, based on internationally accepted criteria with biliary stricturing on magnetic resonance cholangiopancreatogram (MRCP) [18]. Exclusion criteria were as follows: (1) concomitant diagnosis of other chronic liver diseases, including IgG4-related disease; (2) relative contraindications to VCTE (i.e. pregnancy, ascites, etc.); (3) contraindications for magnetic resonance imaging (MRI); (4) acute ascending cholangitis; (5) cholangiocarcinoma or hepatocellular carcinoma; (6) previous liver transplant; (7) failure to complete the MRE or VCTE.

A total of 71 patients were enrolled in the study. Of these, three patients were excluded following a normal follow-up MRCP (one with complete normalisation of liver enzymes and normal repeat MRI and two with biopsy-proven small-duct PSC) and one patient due to claustrophobia.

Study design

The patients underwent MRE along with routine MRI/MRCP. VCTE was performed on the same day or within 3 months of the MRE, using standard procedures (intercostal space near right liver lobe with the patient lying supine, taking the median of ten successful measurements) [19]. Study baselines comprised of laboratory investigations collected within 12 months of the MRE, history of inflammatory bowel disease (IBD), body mass index (BMI, kg/m^2) and outcomes such as first decompensating event (ascites, variceal bleeding, hepatic encephalopathy), liver-related death and transplant or wait-listing for transplant [both requiring sodium-model for end-stage liver disease (Na-MELD) score ≥14]. Clinically significant portal hypertension was defined as platelets $<150 \times 10^9$/L with esophageal or gastric varices or splenomegaly [20]. Since biopsy is not standard of care for PSC, paired biopsies were not protocolised for this study. The definition of cirrhosis was based on a combination of clinical findings (history of decompensation), the presence of portal hypertension or

ultrasonographic evidence of cirrhosis (nodularity with lobular redistribution and caudate lobe enlargement, with or without evidence of portal hypertension. MRSs were calculated at the end of follow-up as surrogate measures of outcome.

VCTE measurement

The majority of VCTE measurements were performed by a dedicated nurse operator (blinded to the clinical history), and the remainder by the clinician as part of clinical care. A total of ten measurements were taken and the interquartile range (IQR) of measurements was recorded. Patients were categorised into five groups of F0 (no fibrosis), F1 (mild fibrosis), F2 (moderate fibrosis), F3 (severe fibrosis) and F4 (cirrhosis) based on LS value cut-offs of 7.4 kPa, 8.6 kPa, 9.6 kPa and 14.4 kPa [21].

Radiological assessment

Magnetic resonance imaging (MRI) and MRE

All patients underwent MRI on a 1.5-T MRI unit (Avanto; Siemens Healthcare, Erlangen, Germany) with a combination of multichannel phase array coil and body coil. The upper abdomen was scanned with a comprehensive pancreaticobiliary protocol. A summary of the routine pulse sequences is depicted in Table 1.

MRE was performed in the supine position. Shear waves, generated by the acoustic active driver at a single fixed frequency of 60Hz, were transmitted to a passive pneumatic driver (paddle) connected to the active driver by a long plastic tubing. The passive driver was positioned placed on the patients' right upper abdomen and held in place by straps. The shear wave propagation through the liver was imaged with a multislice breath-hold 2D GRE MRE sequence (TR/TE = 50/23ms; flip angle, 25°; bandwidth, 260 Hz/Px; matrix, 128×128; acceleration factor of 2 and an asymmetric 75% field of view adjusted to fit each patient; field of view = 40 cm with acquisition time of 22s per slice). Up to five 10-mm slices were obtained through the mid two-thirds of the liver with an approximate breath-hold of 16-18 s per slice. Quantitative stiffness maps were generated automatically by software included in the MRE package on the MRI scanner.

Liver stiffness measurement with MRE

One reader (17 years of experience in MRI), blinded to both the clinical scenario and the VCTE results, measured the MRE LS values by each of the following three ways (Fig. 1) utilising post-processing software (OsiriX) on a standalone Macintosh Computer (Apple):

Trace ROI method

Manual contouring on all the MRE slices avoiding the liver edges, artefacts, focal lesions, large blood vessels and adhering to the areas shown on the stiffness maps as having high 95% confidence intervals (95% CI). This provided traced stiffness values for each slice which were summated to obtain a mean trace stiffness value for each patient; this method evaluated maximum possible liver portion with a mean liver area of 339 cm^2 (range, 142-614 cm^2).

Average ROI method

Two ROIs (1-5 cm^2) randomly placed on each MRE slice, taking precautions as in the Trace ROI method. Up to ten ROIs per MRE examination were recorded, to obtain an average LS value. Mean liver area of approximately 68 cm^2 (range, 7.8-155 cm^2) was included.

Table 1 Routine pancreaticobiliary MRI pulse sequences and MRE performed at 1.5 T

	TR (ms)	TE (ms)	FA	AF	VS (mm)	FOV (mm)	ST (mm)	SN	AT (s)
Axial 3D T1 VIBE Dixon	4	1.31, 2.4 9		2	1.1 × 1.1 × 3.0	360	3	72[a]	18
Coronal T2 HASTE	1,600	181	160	2	1.1 × 1.1 × 5.0	320	5	36	58
Axial T2 HASTE	1,600	80	160	2	1.2 × 1.3 × 5.0	400	5	35	56
DW MRI (b =100, 600)	6,100	68	-	2	1.1 × 1.1 × 5.0	400	5	35	3:58
MRCP: T2 HASTE Radial Slab	8,000	833	180	2	0.7 × 0.7 × 40.0	280	40	16[b]	2:08
MRCP: 3D T2 SPACE	2,500	600	140	3	1.1 × 1.1 × 1.5	380	1.5	88[a]	4:49
2D GRE MRE	50	23	25	2	1.6 × 1.6 × 5.0	400	5	5	22
Gadolinium contrast-enhanced 3D T1 VIBE (3 phases)	4.02	1.32	9	4	1.1 × 1.1 × 3.0	360	3	72[a]	41

AF acceleration factor, *AT* acquisition time, *FA* flip angle, *FOV* field of view, *SN* slice numbers, *ST* slice thickness, *VS* voxel size

[a] Per slab

[b] Slice groups

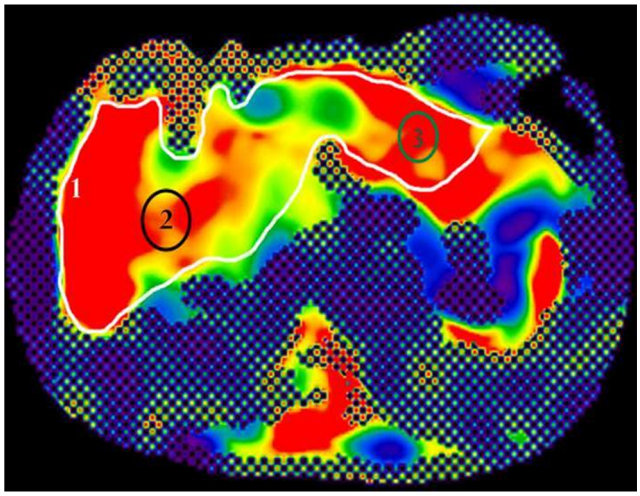

Fig. 1 Representative single axial MRE stiffness map image displaying methods of ROI analysis: (1) Trace ROI method (freehand drawn *white line* contouring liver margins), (2) Average ROI method (*black oval* around the randomly selected area) and (3) Maximum ROI method (*green oval*). MRE stiffness values (kPa) were 7.5, 6.8 and 8.7 respectively by the three ROI methods

Maximum ROI method

The ROI with the highest LS value recorded during the previous step was selected as the maximum ROI stiffness for each patient. A mean liver area of 2.4 cm^2 (range, 1.2–8.2 cm^2) was included. This method was done to emulate the histopathological scoring for fibrosis, wherein when areas of heterogeneous fibrosis are visible, the area score with the most severe fibrosis is utilised for the reported final histological fibrosis score. Figure 2 displays MRE LS maps with corresponding T1-weighted and T2-weighted MR images in three patients with varying degrees of LS.

Statistical analysis

Continuous parameters are reported as median and IQR or mean ± standard deviation (SD); proportions are reported as total numbers and proportions. Pearson's correlation (r) was also used to assess the association between the MRS at the end of follow-up, MRE-quantified LS and VCTE. The area under the receiver operating characteristic (AUROC) curve was used to determine the cut-points with the highest combined sensitivity and specificity, positive predictive values (PPV), negative predictive values (NPV) and accuracy using VCTE as a reference standard. The DeLong test was used to assess the differences between paired AUROC curves. Cut-points were internally validated using bootstrapping with 1,000 random replications of the original sample. Univariable and multivariable linear regression was used to test the associations between demographic (sex, age, BMI), clinical (any decompensation during follow-up or use of ursodeoxycholic acid at the time of MRE, presence of IBD) and laboratory parameters [alkaline phosphatase (ALP), AST, bilirubin, albumin, platelets, gamma-glutamyltransferase (GGT)] with MRE stiffness values or MRS at the end of follow-up. Where VCTE values had IQR >30% of the median measurement, post hoc sensitivity analysis was performed by excluding these subjects to determine the effect on the operating characteristic of MRE [19]. A two-sided p value <0.05 was deemed statistically significant. Statistics were performed using Stata software version 12.0 (StataCorp, College Station, TX, USA) and SPSS Statistical software version 20.0 (IBM, Armonk, New York).

Results

Sixty-seven patients with PSC completed the study. The patient demographics and characteristics are outlined in Table 2. Based on VCTE stiffness values, 36 patients (54%) had no fibrosis (F0), 4, 3 and 8 patients (6%, 4% and 12%) showed mild, moderate and severe fibrosis (F1, F2 and F3), and 16 patients (24%) were cirrhotic (F4). In total, 15 patients (22%) had IQR values >30% of the median VCTE stiffness value. Median MELD score was 7 (IGR, 6-8). Five patients were missing serum sodium values, the median Na-MELD score in the remaining 62 patients was 8.6 (IQR, 6.3-9.5). MRS was low in 38 patients (58%), intermediate in 27 patients (41%) and high in one patient. The median MRS was -0.10 (IQR, -0.82 to 0.49), where score less than or equal to 0 are in the low-risk group and those greater than 0 but less than 2 are in the intermediate risk group [1]. Correlation was highest between the MRS and the Trace ROI method ($r = 0.68$, $p < 0.0001$), followed by MRS and VCTE ($r = 0.62$, $p < 0.0001$), MRS and Average ROI method ($r = 0.66$, $p < 0.0001$) and MRS and Max ROI method ($r = 0.46$, $p < 0.0001$).

The Trace ROI method, using VCTE as the reference standard, had the highest c-statistic, sensitivity, specificity and accuracy compared with the Average and Maximum ROI methods (Table 3). Comparison of paired AUROC curves showed no significant difference in c-statistics and operating characteristics of Trace ROI method with the exclusion of the 15 patients with IQR values >30% on VCTE (p value, 0.50, 0.42 and 0.45 for ≥F1, ≥F2 and ≥F4). Thus, for all subsequent analysis, MRE stiffness using the Trace ROI method with the entire sample ($n = 67$) was used as the main diagnostic test. Using bootstrapping, cut-points for stages of fibrosis were internally validated (Table 4); these values overlapped with all the cut-points derived from the original AUROC. On univariable analysis, variables associated with advanced disease and cholestasis, including abnormal bilirubin, ALP, AST, ALT ($p < 0.001$), low albumin ($p = 0.002$) and thrombocytopaenia ($p = 0.02$), were predictive of higher MRE stiffness as measured by the Trace ROI method. On

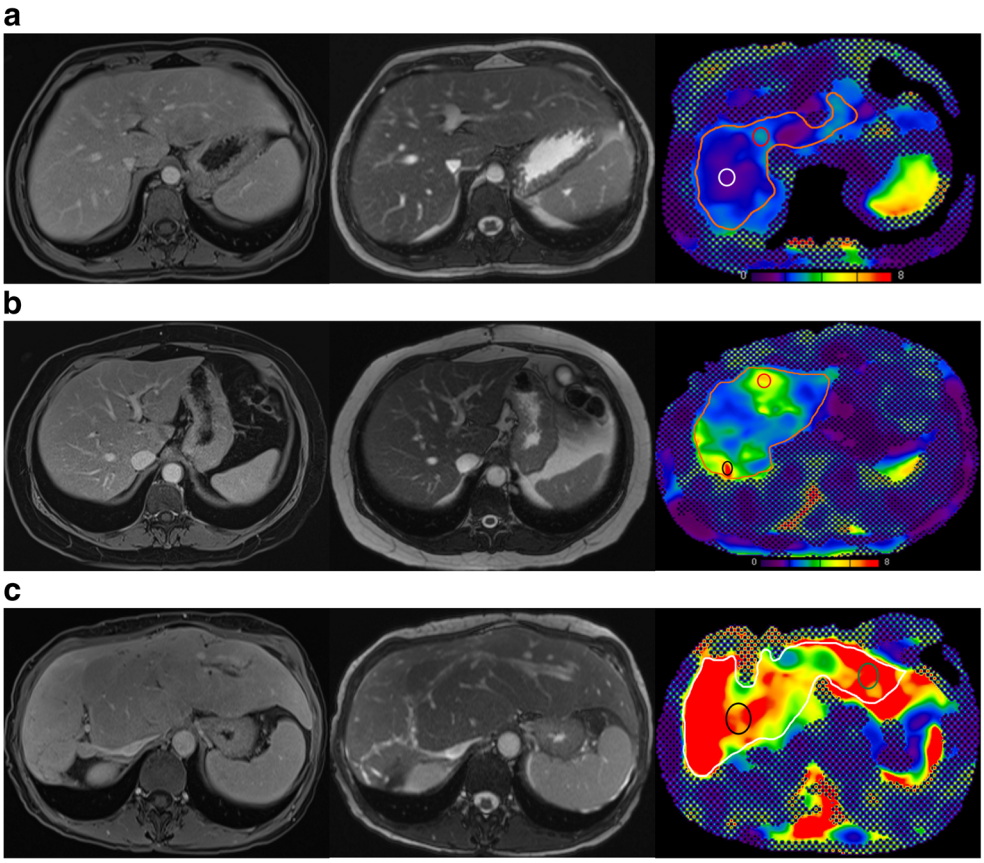

Fig. 2 Axial T1-weighted (*left*), T2-weighted (*middle*) and the corresponding MRE stiffness map images (*right*) in patients with mild (**a**), moderate (**b**) and severe (**c**) liver fibrosis

multivariable analysis, ALP was two times the upper limit of normal (2ULN) [β, 1.55 (95% CI, 0.95– 2.17) $p < 0.001$], abnormal bilirubin [β, 1.27 (95% CI, 0.41–2.14), $p = 0.02$] and thrombocytopaenia [β, 0.79 (95% CI, 0.12–1.46), $p = 0.02$] were independently associated with higher MRE-quantified LS. When stratified by features suggestive of mild and severe cholestasis (i.e. abnormal ALP and/or bilirubin), MRE-quantified LS scores were significantly different only in F3-F4 categories between patients with ALP lower or higher than 2UNL (4.18 ± 1.22 vs 5.69 ± 1.65; *p* value, 0.02).

With respect to predictors of final MRS, both MRE-quantified LS using Trace ROI [β, 0.44 (95% CI, 0.27-0.61), $p < 0.001$] and BMI at enrolment [β 0.09 (95% CI, 0.03-0.16), $p < 0.001$] were predictive of higher MRS (Table 5), with MRE-quantified LS having a significant correlation with the final MRS ($\rho = 0.50$, $p < 0.0001$).

Discussion

Our study showed that MRE-quantified LS is highly dependent on the ROI method used to measure stiffness values, and may be influenced by the severity of cholestasis and portal hypertension, in a manner that is analogous to VCTE. Importantly, the Trace ROI method utilised to measure LS by MRE had a higher correlation with VCTE and better performance characteristics in distinguishing cirrhosis than Average or Maximum ROI method. Higher MRE-quantified LS values were significantly correlated with higher MRS, suggesting that MRE may be a valuable tool for non-invasive risk stratification. There have been few studies comparing different ROI techniques using MRE, despite the fact that the choice of ROI technique may significantly impact resulting MRE stiffness values [22, 23]. In a study of 49 live liver donors with concomitant liver biopsies, ROI methods that encompassed 70% and 32% of the hepatic area compared to a method that included 16% demonstrated that the former two methods had lower inter-observer variability (ICC, 0.845 vs 0.800 vs 0.416, $p < 0.01$), as well as lower intra-observer variability (ICC, 0.852 vs 0.914 vs 0.238, $p < 0.0001$) [22]. Conversely, another study evaluating MRE ROI methods demonstrated no difference between a single ROI of 70% and three circular ROIs (2.0 cm^2 per slice from four hepatic slices) [23]. However, this study did not stratify patients according to the presence or absence of chronic liver disease, thus it is unclear which of the different ROI methods are more accurate in the presence of fibrosis, particularly when its distribution is heterogeneous (as in PSC).

Our study suggests that the Trace ROI method for calculating MRE-quantified LS has a higher correlation with MRS when compared with VCTE or other ROI methods; and also

Table 2 Patients' characteristics

Parameter	Numbers/values
Number of enrolled patients, total (female, male)	67 (28, 39)
Body mass index, kg/m^2 (IQR)[a]	24.7 (22.7–27.1)
Age at diagnosis, median (IQR)	36 (25–50)
Age at MRE, median (IQR)	46 (31–60)
IBD, no. (%)	53 (79)
Years with PSC, median (IQR)	6 (3–11)
On UDCA at time of MRE, no. (%)	20 (30)
Follow-up (months), median (range)	9.1 (0.3–26)
Outcomes, no. (%)	
Alive without decompensation	62 (93)
Decompensation during follow-up	5 (7)
Dead	0
Transplanted	1 (2)
Laboratory parameters, median (IQR)*	
Alkaline phosphatase, U/L	168 (86–299)
Aspartate aminotransferase, U/L	42 (25–71)
Alanine aminotransferase, U/L	44 (28–85)
Gamma-glutamyltransferase, U/L[b]	171 (85–341)
Bilirubin, mg/dL	0.6 (0.5–1.1)
International normalised ratio[c]	1.00 (0.95–1.05)
Albumin, g/L	4.1 (3.7–4.3)
Platelets, × 10^9/L	240 (1673–311)
IgG4, mg/mL[d]	34 (15–68)

IBD inflammatory bowel disease, *IQR* interquartile range, *MRE* magnetic resonance elastography, *PSC* primary sclerosing cholangitis, *UDCA* ursodeoxycholic acid

[a] Range, 18.4–35.9 kg/m2

[b] 1 missing

[c] 2 patients on warfarin

[d] 9 missing

[*] At time of MRE

shows overall higher sensitivity, specificity, PPV, NPV and accuracy in differentiating between the different stages of fibrosis, particularly cirrhosis. As the Trace ROI method takes into account the largest hepatic area of all ROI methods, it is likely that this method provides a more accurate assessment of overall fibrosis in PSC. Conversely, both the Average and Maximum ROI methods consist of randomly selected ROIs, similar to liver biopsies or VCTE and likely to suffer from similar sampling issues due to the heterogeneous distribution of fibrosis in PSC.

Importantly, LS measurements can be influenced by inflammation, portal hypertension and cholestasis. Specifically, VCTE-quantified LS levels increase with cholestasis [24]. Thus, chronic cholestatic liver diseases [21, 25, 26] appear to require higher thresholds for VCTE results in order to achieve sensitivity and specificity comparable to those in viral hepatitis

Table 3 Operating characteristics for the diagnosis of advanced fibrosis for the ROI methods for MRE derived liver stiffness[a]

ROI method	Test threshold[a]	MRE cut-point (kPa)[b]	c-statistic (95%CI)	SE	SN (95% CI)	SP (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)
Trace	≥F1	≥3.33	0.76 (0.65–0.87)	0.05	68.8 (55.3–78.5)	82.9 (70.5–91.8)	78.6 (63.2–89.7)	74.4 (63.3–82.3)	76.1 (63.2–85.4)
	≥F2	≥3.61	0.82 (0.72–0.92)	0.05	74.1 (59.2–83.3)	90.0 (79.9–96.3)	83.3 (66.6–93.8)	83.7 (74.4–89.5)	83.6 (71.6–91.0)
	≥F4	≥4.02	0.92 (0.83–1.00)	0.05	87.5 (66.9–96.8)	96.1 (89.6–99.0)	87.5 (66.9–96.8)	96.1 (89.6–99.0)	94.0 (84.2–98.5)
Average	≥F1	≥4.79	0.73 (0.62–0.84)	0.05	68.8 (55.0–79.5)	77.2 (64.5–87.0)	73.3 (58.6–84.8)	73.0 (61.1–82.3)	73.1 (60.0–83.4)
	≥F2	≥4.89	0.76 (0.65–0.86)	0.05	74.1 (58.1–86.1)	77.5 (66.7–85.6)	69.0 (54.1–80.1)	81.6 (70.2–90.1)	76.1 (63.3–85.8)
	≥F4	≥6.03	0.95 (0.88–1.00)	0.05	93.8 (74.0–99.7)	96.1 (89.9–97.9)	88.2 (69.6–93.8)	98.0 (91.7–99.9)	95.5 (86.1–98.3)
Maximum	≥F1	≥6.30	0.69 (0.58–0.80)	0.06	75.0 (61.1–86.2)	62.9 (50.1–73.1)	64.9 (52.8–74.5)	73.3 (58.5–85.2)	68.7 (55.4–79.3)
	≥F2	≥7.16	0.69 (0.57–0.80)	0.06	70.4 (54.0–83.6)	67.5 (56.5–76.4)	59.4 (45.6–70.5)	77.1 (64.5–87.3)	68.7 (55.5–79.3)
	≥F4	≥8.68	0.80 (0.68–0.91)	0.06	81.3 (57.4–94.8)	78.4 (71.0–82.7)	54.2 (38.3–63.2)	93.0 (84.2–98.0)	79.1 (67.7–85.6)

MRE magnetic resonance elastography, *NPV* negative predictive value, *PPV* positive predictive value, *ROI* region of interest, *SE* standard error, *SN* sensitivity, *SP* specificity, *VCTE* vibration-controlled transient elastography

[a] Based on VCTE as a reference test, with ≥F1, ≥F2 and ≥F4 defined as 7.4, 8.6 and 14.4 kPa (Corpechot et al [21])

[b] MRE stiffness cut-points as determined by the point of maximal combined sensitivity and specificity on the area under the receiver operator curve analysis

Table 4 Estimated cut-points and operating characteristics of MRE for different fibrosis stages following bootstrap resampling

Test threshold[a]	Replications	Original MRE cut-point (kPa)[b]	Bootstrapped MRE cut-point (kPa)[c]	c-statistic (95% CI)	SN (95% CI)	SP (95% CI)
≥ F1	1,000	3.33	3.49 ± 0.35	0.86 ± 0.44	71.0 ± 12.2	86.9 ± 10.5
≥ F2	1,000	3.61	3.68 ± 0.25	0.90 ± 0.04	74.6 ± 9.6	92.9 ± 6.8
≥ F3	1,000	3.71	3.84 ± 0.21	0.93 ± 0.03	78.3 ± 9.7	96.5 ± 3.5
≥ F4	1,000	4.02	4.11 ± 0.21	0.96 ± 0.03	87.9 ± 8.4	97.1 ± 2.8

AUROC area under the receiver operating curve, *MRE* magnetic resonance elastography, *SN* sensitivity, *SP* specificity

[a] Based on transient elastography (Corpechot et al [21])

[b] Based on cut-points generated by the standard AUROC curve, targeting highest combined sensitivity and specificity

[c] Based on MRE stiffness values generated by the Trace ROI method

[27, 28] and non-alcoholic liver disease [29]. Likewise, MRE may be influenced by cholestasis, as demonstrated a recent study by Eaton et al [30], where subjects with dominant strictures had a higher median LS according to MRE, with stiffness values falling with therapeutic endoscopic treatment. Our study also demonstrates that cholestasis (abnormal ALP and bilirubin), as well as portal hypertension, are associated with higher MRE values. Advancing fibrosis itself leads to worsening of all three biochemical parameters, making it likely that MRE stiffness values will be relatively higher in cholestatic liver disease than other chronic liver diseases for corresponding stages of fibrosis. Thus, it may be necessary for patients with obstructive cholestasis to undergo a repeat MRE to assess for disease severity after the obstruction has resolved.

While our study includes a small number of patients with moderate to advanced disease, internal validation with bootstrap resampling demonstrated similar cut-points for MRE LS as the non-bootstrapped values. Furthermore, the exclusion of patients with VCTEs with an IQR >30% did not change the performance characteristics of MRE. As the majority of patients in this study had early disease, it was not possible to assess the association between MRE stiffness and liver outcomes such as death, transplantation or decompensation. However, MRE stiffness was independently associated with

Table 5 Univariable and multivariable analysis of independent predictors of Mayo Risk Score (MRS) at end of follow-up[a]

Parameters	β	SE	95% CI	*p* value
Univariable analysis				
MRE-quantified LS using Trace ROI[b]	0.43	0.09	0.24–0.61	<0.001
BMI[b]	0.08	0.04	0.008–0.16	0.03
ALP[b]	0.002	0.0007	0.0004–0.003	0.01
Thrombocytopaenia (platelets <150x10^9/L)[b]	0.67	0.37	-0.08–1.41	0.08
GGT	0.001	0.0004	-0.0001–0.002	0.09
ALP 2ULN[c]	0.54	0.35	-0.16–1.24	.13
UDCA use during MRE	0.20	0.32	-0.4–0.83	0.53
Platelets	-0.008	0.001	-0.003–0.001	0.46
Female sex[b]	- 0.04	0.30	-0.64–0.56	0.90
Multivariable analysis				
MRE-quantified LS using Trace ROI	0.44	0.09	0.27–0.61	<0.001
BMI	0.09	0.88	0.03–0.16	<0.001

Identical results using both backwards and forwards selection, with variables in univariable and multivariable analysis listed in order of magnitude of effect and significance

ALP alkaline phosphatase, *β* beta coefficient, *BMI* body mass index, *CI* confidence interval, *GGT* gamma-glutamyltransferase, *LS* liver stiffness, *MRE* magnetic resonance elastography, *ROI* region of interest; *SE* standard error, *UDCA* ursodeoxycholic acid, *ULN* upper limit of normal

[a] 59 patients with available MRS

[b] Included in the multivariable model

[c] Upper limits variable depending on year and laboratory used

the final MRS, supporting previous evidence that MRE stiffness may provide a means for non-invasive prognostication [30]. The study by Eaton et al demonstrated that MRE thresholds of <4.5 kPa, 4.5-6.0 kPa and >6.0 kPa could stratify patients to a low, medium and high risk of decompensation.

There is evidence that MRE has superior performance to VCTE in assessing the severity of liver fibrosis/cirrhosis from a variety of aetiologies; therefore, validating MRE based on the diagnostic findings provided by VCTE is a genuine limitation. However, it is a pragmatic approach since liver biopsy is not performed routinely and cannot be a reference standard. Given the lack of risk stratification and prognostic tools in PSC, additional prospective studies validating the longitudinal reliability and utility of MRE would be highly valuable. In addition, it would also be essential to evaluate the inter- and intra-observer variability of different ROI methods, while also capitalising on the nature of MRE as an imaging technique, by focusing on the development of imaging algorithms that may isolate the impact of intra- and extra-hepatic cholestasis and fibrosis. Further, other non-invasive markers such as the ELF test, which has recently been shown to be predictive of liver transplant-free survival in patients with PSC [31], would provide additional reference standards for comparison.

In conclusion, our study demonstrates that the MRE has a higher correlation with MRS than VCTE, and provides the best performance characteristics for distinguishing early and moderate fibrosis from cirrhosis. Furthermore, although MRE-quantified LS values are likely influenced by cholestasis and portal hypertension, yet they are independently associated with liver transplant-free survival in PSC.

Funding This study has received peer-reviewed research grant funding by "The Physicians Services Incorporated Foundation, Toronto, Canada".

There has been no financial support for this work to the authors that could have influenced its outcome.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Dr. Kartik Jhaveri.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Ravi Menezes kindly provided statistical advice for this manuscript. Dr. Angela Cheung, one of the co-authors, also has significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- prospective
- diagnostic study
- performed at one institution

References

1. Kim WR, Therneau TM, Wiesner RH et al (2000) A revised natural history model for primary sclerosing cholangitis. Mayo Clin Proc 75:688–694
2. Abraham SC, Kamath PS, Eghtesad B, Demetris AJ, Krasinskas AM (2006) Liver transplantation in precirrhotic biliary tract disease: Portal hypertension is frequently associated with nodular regenerative hyperplasia and obliterative portal venopathy. Am J Surg Pathol 30:1454–1461
3. Carbone M, Neuberger J (2011) Liver transplantation in PBC and PSC: indications and disease recurrence. Clin Res Hepatol Gastroenterol 35:446–454
4. Okolicsanyi L, Fabris L, Viaggi S, Carulli N, Podda M, Ricci G (1996) Primary sclerosing cholangitis: clinical presentation, natural history and prognostic variables: an Italian multicentre study. The Italian PSC Study Group. Eur J Gastroenterol Hepatol 8:685–691
5. Ponsioen CY, Vrouenraets SM, Prawirodirdjo W et al (2002) Natural history of primary sclerosing cholangitis and prognostic value of cholangiography in a Dutch population. Gut 51:562–566
6. Bambha K, Kim WR, Talwalkar J et al (2003) Incidence, clinical spectrum, and outcomes of primary sclerosing cholangitis in a United States community. Gastroenterology 125:1364–1369
7. Burak KW, Angulo P, Lindor KD (2003) Is there a role for liver biopsy in primary sclerosing cholangitis? Am J Gastroenterol 98:1155–1158
8. Myers RP, Ramji A, Bilodeau M, Wong S, Feld JJ (2012) An update on the management of hepatitis C: consensus guidelines from the Canadian Association for the Study of the Liver. Can J Gastroenterol 26:359–375
9. Coffin CS, Fung SK, Ma MM, Canadian Association for the Study of the Liver (2012) Management of chronic hepatitis B: Canadian Association for the Study of the Liver consensus guidelines. Can J Gastroenterol 26:917–938
10. Castera L (2012) Noninvasive methods to assess liver disease in patients with hepatitis B or C. Gastroenterology 142:1293–1302 e1294
11. European Association for the Study of the Liver (2011) EASL Clinical Practice Guidelines: management of hepatitis C virus infection. J Hepatol 55:245–264
12. Ratziu V, Cadranel JF, Serfaty L et al (2012) A survey of patterns of practice and perception of NAFLD in a large sample of practicing gastroenterologists in France. J Hepatol 57:376–383
13. Low G, Kruse SA, Lomas DJ (2016) General review of magnetic resonance elastography. World J Radiol 8:59–72
14. Venkatesh SK, Yin M, Ehman RL (2013) Magnetic resonance elastography of liver: technique, analysis, and clinical applications. J Magn Reson Imaging 37:544–555
15. Sandrin L, Fourquet B, Hasquenoph JM et al (2003) Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 29:1705–1713
16. Faria SC, Ganesan K, Mwangi I et al (2009) MR imaging of liver fibrosis: current state of the art. Radiographics 29:1615–1635
17. Carrasco-Avino G, Schiano TD, Ward SC, Thung SN, Fiel MI (2015) Primary sclerosing cholangitis: detailed histologic assessment and integration using bioinformatics highlights arterial fibrointimal hyperplasia as a novel feature. Am J Clin Pathol 143:505–513

18. Chapman R, Fevery J, Kalloo A et al (2010) Diagnosis and management of primary sclerosing cholangitis. Hepatology 51:660–678
19. Armstrong MJ, Corbett C, Hodson J et al (2013) Operator training requirements and diagnostic accuracy of Fibroscan in routine clinical practice. Postgrad Med J 89:685–692
20. Giannini EG (2006) Review article: thrombocytopenia in chronic liver disease and pharmacologic treatment options. Aliment Pharmacol Ther 23:1055–1065
21. Corpechot C, Gaouar F, El Naggar A et al (2014) Baseline values and changes in liver stiffness measured by transient elastography are associated with severity of fibrosis and outcomes of patients with primary sclerosing cholangitis. Gastroenterology 146:970–979 quiz e915-976
22. Lee DH, Lee JM, Han JK, Choi BI (2013) MR elastography of healthy liver parenchyma: Normal value and reliability of the liver stiffness value measurement. J Magn Reson Imaging 38:1215–1223
23. Silva AM, Grimm RC, Glaser KJ et al (2015) Magnetic resonance elastography: evaluation of new inversion algorithm and quantitative analysis method. Abdom Imaging 40:810–817
24. Millonig G, Reimann FM, Friedrich S et al (2008) Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. Hepatology 48:1718–1723
25. Corpechot C, El Naggar A, Poujol-Robert A et al (2006) Assessment of biliary fibrosis by transient elastography in patients with PBC and PSC. Hepatology 43:1118–1124
26. Corpechot C, Carrat F, Poujol-Robert A et al (2012) Noninvasive elastography-based assessment of liver fibrosis progression and prognosis in primary biliary cirrhosis. Hepatology 56:198–208
27. Marcellin P, Ziol M, Bedossa P et al (2009) Non-invasive assessment of liver fibrosis by stiffness measurement in patients with chronic hepatitis B. Liver Int 29:242–247
28. Ziol M, Handra-Luca A, Kettaneh A et al (2005) Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. Hepatology 41:48–54
29. Wong VW, Vergniol J, Wong GL et al (2010) Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 51:454–462
30. Eaton JE, Dzyubak B, Venkatesh SK et al (2016) Performance of magnetic resonance elastography in primary sclerosing cholangitis. J Gastroenterol Hepatol 31:1184–1190
31. Vesterhus M, Hov JR, Holm A et al (2015) Enhanced liver fibrosis score predicts transplant-free survival in primary sclerosing cholangitis. Hepatology 62:188–197