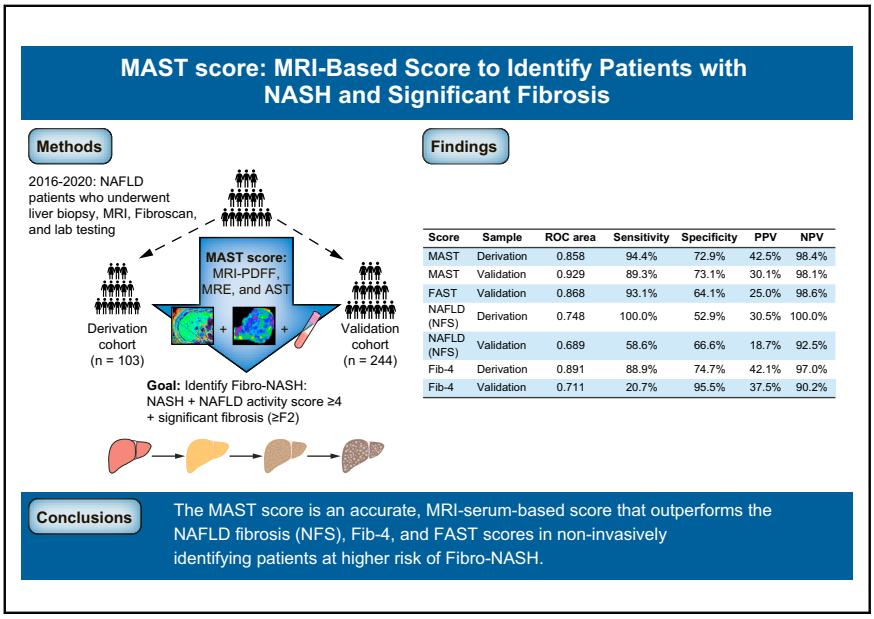
Research Article
NAFLD and Alcohol-Related Liver Diseases

JOURNAL
OF HEPATOLOGY

MRI-based (MAST) score accurately identifies patients with NASH and significant fibrosis

Graphical abstract



Score	Sample	ROC area	Sensitivity	Specificity	PPV	NPV
MAST	Derivation	0.858	94.4%	72.9%	42.5%	98.4%
MAST	Validation	0.929	89.3%	73.1%	30.1%	98.1%
FAST	Validation	0.868	93.1%	64.1%	25.0%	98.6%
NAFLD (NFS)	Derivation	0.748	100.0%	52.9%	30.5%	100.0%
NAFLD (NFS)	Validation	0.689	58.6%	66.6%	18.7%	92.5%
Fib-4	Derivation	0.891	88.9%	74.7%	42.1%	97.0%
Fib-4	Validation	0.711	20.7%	95.5%	37.5%	90.2%

Highlights

- Patients with NASH, NAS ≥4, and F≥2 (Fibro-NASH) are at the highest risk of disease progression.
- Patients with Fibro-NASH are targeted for proof-of-concept NASH trials.
- MRI-PDFF and MR elastography are the most common primary and secondary endpoints in NASH trials, respectively.
- To identify Fibro-NASH, the MRI-AST (MAST) score was created based on these MRI techniques.
- The MAST score was proven to outperform the FIB-4, NAFLD fibrosis, and FAST scores.

Authors

Mazen Noureddin, Emily Truong, Jeffrey A. Gornbein, …, Ju Dong Yang, Stephen A. Harrison, Naim Alkhouri

Correspondence

Mazen.Noureddin@cshs.org (M. Noureddin).

Lay summary

Identifying patients with non-alcoholic steatohepatitis and significant fibrosis – who need treatment and are at risk of clinical liver-related outcomes – is a clinical priority. We developed a more accurate score using MRI-based technologies and a laboratory blood test (aspartate aminotransferase) that outperforms previous non-invasive scores for the identification of patients at higher risk of liver disease progression.

https://doi.org/10.1016/j.jhep.2021.11.012
 J. Hepatol. 2022, 76, 781–787

Research Article
NAFLD and Alcohol-Related Liver Diseases

JOURNAL OF HEPATOLOGY

MRI-based (MAST) score accurately identifies patients with NASH and significant fibrosis

Mazen Noureddin[1,2,3,4,*], Emily Truong[2,4], Jeffrey A. Gornbein[5], Rola Saouaf[4], Maha Guindi[4], Tsuyoshi Todo[3,4], Nabil Noureddin[6], Ju Dong Yang[1,2,3,4], Stephen A. Harrison[7], Naim Alkhouri[8]

[1]*Karsh Division of Gastroenterology and Hepatology, Cedars-Sinai Medical Center, Los Angeles, California, United States;* [2]*Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California, United States;* [3]*Comprehensive Transplant Center, Cedars-Sinai Medical Center, Los Angeles, California, United States;* [4]*Cedars-Sinai Medical Center, Los Angeles, California, United States;* [5]*Department of Medicine, University of California Los Angeles, Los Angeles, California, United States;* [6]*Division of Gastroenterology and Hepatology, University of California San Diego, San Diego, California, United States;* [7]*Pinnacle Research Center, Oxford University, Oxford, United Kingdom;* [8]*Arizona Liver Health, Phoenix, Arizona, United States*

Background & Aims: Among the large population of patients with non-alcoholic fatty liver disease (NAFLD), identifying those with fibrotic non-alcoholic steatohepatitis (Fibro-NASH) is a clinical priority, as these patients are at the highest risk of disease progression and will benefit most from pharmacologic treatment. MRI-based proton density fat fraction (MRI-PDFF) and MR elastography (MRE) can risk-stratify patients with NAFLD by assessing steatosis and fibrosis, respectively. We developed a highly specific MRI-based score to identify patients with Fibro-NASH.

Methods: This analysis included derivation (n = 103) and validation (n = 244) cohorts of patients who underwent MRI, liver biopsy, transient elastography, and laboratory testing for NAFLD from 2016-2020 in 2 tertiary care centers. To identify Fibro-NASH, a formula was developed based on MRI-PDFF, MRE, and a third variable with highest balanced accuracy per logistic regression. The MRI-aspartate aminotransferase (MAST) score was created and compared to NAFLD fibrosis (NFS), Fibrosis-4 (FIB-4), and FibroScan-aspartate aminotransferase (FAST) scores.

Results: The MAST score demonstrated high performance and discrimination in the validation cohort (AUC 0.93; 95% CI 0.88-0.97). In the validation cohorts, the 90% specificity cut-off of 0.242 corresponded to a sensitivity of 75.0%, positive predictive value (PPV) of 50.0% and negative predictive value (NPV) of 96.5%, whereas the 90% sensitivity cut-off of 0.165 corresponded to a specificity of 72.2%, PPV of 29.4%, and NPV of 98.1%. Compared to NFS and FIB-4, MAST resulted in fewer patients having indeterminate scores and an overall higher AUC. Compared to FAST, MAST exhibited a higher AUC and overall better discrimination.

Conclusion: The MAST score is an accurate, MRI-serum-based score that outperforms previous scores in non-invasively identifying patients at higher risk of Fibro-NASH.

Lay summary: Identifying patients with non-alcoholic steatohepatitis and significant fibrosis – who need treatment and are at risk of clinical liver-related outcomes – is a clinical priority. We developed a more accurate score using MRI-based technologies and a laboratory blood test (aspartate aminotransferase) that outperforms previous non-invasive scores for the identification of patients at higher risk of liver disease progression.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease worldwide, involving 25% of the general population and rising in prevalence to 70% of patients with comorbid obesity and type 2 diabetes.[1] NAFLD can not only progress to non-alcoholic steatohepatitis (NASH) and even cirrhosis, but is also one of the leading indications for liver transplantation in the US and Europe.[2–5] NAFLD's clinical silence and lack of symptomatology often prevents clinicians from appropriately diagnosing and intervening in a timely manner.[6] A challenge is identifying patients at higher risk of disease progression to fibrosis and cirrhosis, who may benefit from targeted novel pharmacotherapies.

To date, methods to determine risk for liver-related outcomes include non-invasive risk stratification of liver fibrosis or percutaneous liver biopsy. Non-invasive risk stratification strategies encompass serum biomarkers, imaging, and diagnostic algorithms.[7–9] NASH activity is a histological assessment of the disease activity in clinical trials and summarized by the NAFLD activity score (NAS), which is a composite sum of the histological scores for steatosis, hepatocellular injury as evidenced by ballooning, and lobular inflammation.[10] NAS is usually used in combination with fibrosis stage to determine who may benefit from therapies, as patients with NASH and fibrosis stage ≥2 are at increased risk of liver-related outcomes and disease progression towards cirrhosis.[11–13] Because the combination of NASH activity and fibrosis stage accurately characterizes disease state, patients with fibrotic NASH (Fibro-NASH) defined as elevated NASH activity (NAS ≥4) and significant fibrosis (F ≥2) are targeted for inclusion in clinical trials.[14,15]

Keywords: Fatty Liver; NAFLD; NASH treatment; MRI-PDFF; MRE.
Received 4 February 2021; received in revised form 19 October 2021; accepted 8 November 2021; available online 17 November 2021
* Corresponding author. Address: Cedars-Sinai Fatty Liver Program, Division of Digestive and Liver Diseases, Comprehensive Transplant Center, Department of Medicine, Cedars-Sinai Medical Center, 8900 Beverly Blvd., Los Angeles, CA 90048, United States; Tel.: +1-310-423-1971.
E-mail address: Mazen.Noureddin@cshs.org (M. Noureddin).
https://doi.org/10.1016/j.jhep.2021.11.012

Recent efforts to non-invasively identify this subgroup of patients include the NIS4 and FibroScan-aspartate aminotransferase (FAST) scores.[16,17] The NIS4 score is a blood-based multivariate index designed to rule-in and rule-out at-risk NASH with (NAS ≥4), and significant fibrosis (F≥2) (Fibro-NASH).[16] Similarly, the FAST score identifies patients with Fibro-NASH through the combination of liver stiffness measurement (LSM) by vibration-controlled transient elastography (VCTE), controlled attenuation parameter (CAP), and aspartate aminotransferase (AST).[17] However, studies considering magnetic resonance techniques in identifying this subgroup of patients are scarce, especially considering that MRI-based assessment is the primary inclusion criteria and its results are the primary end points of phase IIA studies. Though one study has investigated a non-invasive MRI biomarker, iron-corrected T1 mapping, as a diagnostic biomarker for NASH, it is MRI-based proton density fat fraction (MRI-PDFF) analysis that has shown more accuracy than VCTE-based CAP in identifying all grades of steatosis in patients with NAFLD.[18,19] Compared to liver biopsy as the standard, MR elastography (MRE) is more accurate than VCTE in detecting liver fibrosis (stage 1 or more).[19]

We hypothesized that an MRI-driven score would outperform a VCTE-driven score in identifying patients with Fibro-NASH. We aimed to develop and validate a highly specific score using MRI-PDFF and MRE that identifies patients with Fibro-NASH for therapeutic pharmacotherapies and for clinical trial eligibility.

Patients and methods

Cohorts

This is a retrospective analysis from ongoing prospective real-world studies that included a derivation (n = 103) and validation (n = 244) cohort. These data were collected between 2016 and 2020 from 2 tertiary care centers in California and Texas in patients with suspected NAFLD who had liver biopsy, MRE/MRI-PDFF, transient elastography, and laboratory tests within a 6-month period. The cohort from California had the following exclusion criteria: i) excessive alcohol intake (defined as >21 standard drinks of alcohol/week for men and >14 standard drinks of alcohol/week for women over a 2-year time frame); ii) history of liver disease including chronic hepatitis B or C, hemochromatosis, Wilson's disease, autoimmune hepatitis, primary sclerosing cholangitis, primary biliary cholangitis, α1-antitrypsin deficiency, or other causes of chronic liver disease including medications that can cause fatty liver; iii) contraindications to MRI; and iv) history of human immunodeficiency virus. Patients underwent detailed medical history, physical examination, and laboratory assessment prior to undergoing transient elastography (FibroScan®) for measurement of CAP and LSM to assess steatosis and fibrosis, respectively, and MRI-PDFF and MRE to assess hepatic steatosis and fibrosis, respectively. Patients were further selected for liver biopsy based on abnormalities on imaging including: MRI-PDFF ≥5% steatosis, LSM >7 kPa on Fibroscan®, or MRE ≥2.5 kPa. The cohort from Texas (validation) underwent a similar selection process, except that the cohort also underwent a cT1 multiparametric scan. Detailed characteristics of this cohort were published recently.[20]

MRIs and VCTE

MRI-PDFF and MRE were performed using Siemens 1.5 or 3 T scanner in the derivative cohort and Siemens 1.5 T scanner in the validation cohort. Patients fasted for at least 4 hours prior to the exam. An acoustic passive driver was positioned on the patient's body in accordance with the location of the liver. For MRE shear waves from active to passive driver of 60 Hz, liver vibrations were measured. Special software were used to process images from the entire liver, assess fat fraction (via measuring 3 regions of interest), and calculate an elastogram. Radiologists from the sites were blinded to clinical and histological data.

VCTE was used to detect the speed of a mechanically generated shear wave across the liver in order to derive an LSM that correlates with hepatic fibrosis burden. The measured attenuation of ultrasound through the liver was used to derive CAP. Technicians, nurses, or physicians who were trained and certified by the manufacturer performed these tests and were blinded to the laboratory and histological data. Exams were performed after at least 3 hours of fasting prior to the exam. Probe selection was made using the automatic probe selection tool (L or XL) provided by the device software. Patients were placed in the supine position with their right arm fully abducted, and measurements were done by scanning the right liver lobe through an intercostal space. At least 10 reliable measurements with at least a 60% success rate and an interquartile range of ≤30% of the median value were obtained to deem the exam valid as indicated by the manufacturer. The CAP is an average estimate of ultrasound attenuation at 3.5 MHz and is expressed in dB/m, while LSM is an average estimate of stiffness at a shear wave frequency of 50 Hz and is expressed as kPa. The FAST score was calculated per the previously published formula.[21]

This study's primary objective was to develop a score combining MRE, MRI-PDFF, and serum biomarkers to identify Fibro-NASH.

Liver histology

Patients with the aforementioned clinical indications for liver biopsy underwent a liver biopsy. Histological assessments for the derivation cohort were made by an expert liver pathologist who was blinded to both imaging and clinical data. The same procedure was followed in the validation cohort. Patients were given the diagnosis of NASH per Brunt criteria and were systemically assessed using the NAS score per Kleiner criteria.[10,22] Steatosis was scored from 0 to 3, lobular inflammation was scored from 0 to 3, and hepatocellular ballooning was scored from 0 to 2. Steatosis, lobular inflammation, and ballooning scores were pooled to acquire the total NAS ranging from 0 to 8. Fibrosis was staged from 0 to 4 (F0–F4), defined as F0: no fibrosis, F1: either mild-moderate perisinusoidal or periportal fibrosis, F2: both peri-sinusoidal and portal/periportal fibrosis, F3: bridging fibrosis, and F4: cirrhosis.

Outcomes

The main outcome was the diagnosis of NASH with NAS ≥4 (with at least one point each for steatosis, lobular inflammation, and ballooning), and fibrosis stage ≥2 (altogether referred to as NASH+NAS ≥4+F≥2 or Fibro-NASH). Up to 11 predictive variables, including MRE, MRI-PDFF, alanine aminotransferase (ALT), AST, AST/ALT ratio, albumin, platelets, diabetes status, sex, age and BMI were considered. A score formula (MRI-AST [MAST] score) was developed based on MRE, MRI-PDFF and a third variable (AST) selected for highest balanced accuracy out of the other 9 variables using logistic regression.

Statistical analysis

Bivariate – The *p* values for comparing continuous variables in positive *vs.* negative that follow the normal distribution were computed using *t* tests. The *p* values for comparing continuous variables that did not follow the normal distribution were computed using the Wilcoxon rank sum test. R 3.5.2 software was used for all analysis and $p < 0.05$ was considered statistically significant.

Receiver-operating characteristic (ROC) – The non-parametric ROC method was used to determine the best cut-off (one value cut-off) of a continuous variable, including logit scores, that maximized balanced accuracy. One measure of balanced accuracy is defined as the average of sensitivity (proportion of true positives correctly classified) and specificity (proportion of true negatives correctly classified). In addition, the area under the ROC curve (AUC), another measure of accuracy, is reported. The one value cut-off is the value where the balanced accuracy (average of sensitivity and specificity) is maximized. Two value cut-offs corresponding to 90% sensitivity (cut-off A) and 90% specificity (cut-off B) are also reported. Values between these 2 cut-offs (A-B) are sometimes referred to as indeterminate or "grey zone" values.

Multivariable – Scores (logit scores) for creating weighted combinations of (log) liver stiffness, PDFF percent, and a third variable (ALT, AST, AST/ALT ratio, albumin, platelets, diabetes status, sex, age, or BMI) were computed using logistic regression. A likelihood ratio (LR) test for interactions was carried out to determine if the score should be an additive combination of the 3 variables (results not shown). A statistically significant LR statistic indicates a lack of additivity. The log scale was used for some variables since a linear relationship with the log odds (logit) was achieved only on this scale. The logistic model with the highest AUC in the validation set and the lowest Akaike information criterion value was selected as the best model. One and two value score values for the final multivariable models are reported.

Results

Baseline characteristics

Based on data availability, after excluding patients with missing data (Fig. S1), 103 of 119 originally eligible participants were included in constructing the MRI-based score, which was then validated in a validation cohort (n = 244 of 266). As Table 1 demonstrated, both the derivation and validation cohorts shared similar demographic, metabolic, serologic, and histologic features. In the derivation and validation cohorts respectively, the mean age was 53.1 (±11.9)/55.7 (±6.3) years with 55.3%/40.2% females and 61%/23% with type 2 diabetes. The derivation/validation cohorts respectively had median (IQR) AST of 38 (27–60)/22 (18-27) IU/L, ALT of 52 (39–95)/25 (20-38) IU/L, liver stiffness

Table 1. Derivation and validation cohort patient characteristics.

	Derivation cohort	Validation cohort
Demographics		
n	103	244
Age- years - mean (SD)	53.1 (11.9)	55.7 (6.3)
Female - n (%)	57 (55.3%)	98 (40.2%)
Male - n (%)	46 (44.7%)	146 (59.8%)
BMI (kg/m^2) - mean (SD)	33.8 (6.6)	33.1 (5.1)
Metabolic		
Diabetes (type 2) - n (%)	63 (61%)	56 (23%)
Blood		
AST (IU/L) - median (IQR)	38 (27-60)	22 (18-27)
ALT (IU/L) - median (IQR)	52 (39-95)	25 (20-38)
Albumin (g/dl) - median (IQR)	4.4 (4.2-4.6)	4.3 (4.1-4.5)
Platelets count (x10^9/L) - median (IQR)	213 (170-282)	234 (201-285)
Triglyceride (mg/dl) - median (IQR)	142 (112-207)	156 (106-222)
Total cholesterol (mg/dl) - median (IQR)	172 (153-202)	88 (72-101)
HDL cholesterol (mg/dl) - median (IQR)	46 (38-54)	22 (18-27)
Fibrosis scores		
FIB-4 - median (IQR)	1.40 (0.80-2.20)	0.98 (0.81-1.31)
NFS - median (IQR)	-0.69 (-2.26 - 0.37)	-1.50 (-2.13- -0.66)
MRI		
Liver stiffness (kPA) - median (IQR)	3.0 (2.3-4.3)	2.3 (2.0-2.6)
MRI-PDFF (%) - median (IQR)	14.2 (9.3-20.2)	7.6 (4.6 - 13.5)
Histology		
NAS score ≥4 - n (%)	50 (48.5%)	80 (32.8%)
NASH (%)		
F0	21 (20.4%)	131 (53.7%)
F1	37 (35.9%)	76 (31.1%)
F2	5 (4.9%)	27 (11.1%)
F3	27 (26.2%)	10 (4.1%)
F4	13 (12.6%)	0 (0%)
NASH + NAS ≥4 + ≥F2	18 (17.5%)	28 (11.5%)

Unless otherwise specified, data are portrayed in n, n (%), mean (SD), or median (IQR). The NAS and Kleiner scoring systems are described in the appendix. Data on ballooning grade and lobular inflammation grade are missing in the derivation cohort. F0=fibrosis stage 0. F1=fibrosis stage 1. F2=fibrosis stage 2. F3=fibrosis stage 3. F4=fibrosis stage 4. ALT, alanine aminotransferase; AST, aspartate aminotransferase;. FIB-4, fibrosis-4; MRI-PDFF, magnetic resonance imaging-derived proton density fat fraction; NASH, non-alcoholic steatohepatitis; NAS, NAFLD activity score. NFS, NAFLD fibrosis score.

in kPA of 3.0 (2.3-4.3)/2.3 (2.0-2.6), and MRI-PDFF of 14.2 (9.3-20.2)/7.6 (4.6-13.5).

In the derivation cohort, Fibro-NASH was reported in 18 (17.5%) of 103 patients. In the validation cohort, Fibro-NASH was reported in 28 (11.5%) of 244 patients (Table 1).

The MAST formula

Models combining MRE, MRI-PDFF, and a third variable selected from ALT, AST, albumin, platelets, diabetes status, age, sex, or BMI were compared. AST (MRI-MRE, MRI-PDFF plus AST) was found to be the best model with highest AUC in the derivation data for predicting Fibro-NASH when combined with steatosis measured with MRI-PDFF and stiffness measured with MRE, resulting in the MAST score (data not shown for other models).

The MAST score is defined as: MAST = -12.17 + 7.07 log MRE + 0.037 PDFF + 3.55 log AST.

ROC plots of sensitivity *vs.* 1-specificity for all possible one score cut-offs for the MAST, FAST, Fibrosis-4 (FIB-4), or NAFLD fibrosis (NFS) scores in either the derivation or validation cohort are shown in Fig. 1. Table S1 shows the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the one score cut-off corresponding to the maximum balanced accuracy in the derivation cohort and applied to both the derivation and validation cohorts. Two score cut-offs that corresponded to 90% specificity and 90% sensitivity in the validation cohorts were computed (Table 2).

MAST in comparison with NFS and FIB-4

MAST's performance in identifying patients with Fibro-NASH was compared to that of the NFS and FIB-4 score using a subgroup of patients in the validation cohort with all the data necessary to determine both the NFS and FIB-4 scores (Table S1 and Fig. 1). Overall, MAST exhibited a higher AUC (Fig. 1).

MAST demonstrated similar PPV and specificity and overall higher sensitivity and NPV at the 90% sensitivity cut-off, except for rare exceptions (Table 2). At the 90% specificity cut-off, MAST exhibited similar specificity except for the NFS and FIB-4 validation cohorts, as well as higher sensitivity, PPV, and NPV in nearly all cohorts (Table 2).

MAST in comparison with FAST

MAST's AUC was higher in both the derivation and validation cohorts compared to the AUC of the FAST score (Fig. 1). To meet a 90% rule-out sensitivity cut-off and 90% rule-in specificity cut-off, respectively, the previously published cut-offs of 0.35 and 0.67 were applied for the FAST score in our cohort. Comparing MAST's validation cohort to FAST's, MAST exhibited higher sensitivity and NPV at the 90% sensitivity cut-off (Table 2). At the 90% specificity cut-off, MAST's validation cohort demonstrated higher sensitivity and NPV (Table 2).

Discussion

This study describes the formulation and validation of a simple, novel MAST score that non-invasively identifies high-risk patients with Fibro-NASH for whom therapeutic pharmacotherapies are indicated. The MAST score included 2 score cut-offs corresponding to 90% specificity and 90% sensitivity in the derivation and validation cohorts that maintained high balanced accuracy while performing better than other non-invasive assessments, specifically the NFS, FIB-4, and FAST scores.

Though selection of patients with NAFLD who are the optimal targets for monitoring and therapeutic interventions has been widely debated, prior data have demonstrated that patients with NASH+NAS≥4+F≥2 are at higher risk of severe liver disease progression and are thus the focus of this study.[11–13] In accordance with current practice, the MAST score was developed with rule-in and rule-out cut-offs that exhibited good performance with a negative LR of 0.15 (rule-out cut-off) and a positive LR of 7.73 (rule-in cut-off) in the validation cohort. The MAST score can thus significantly impact patient care through non-invasive selection of patients for new clinical trials or pharmacotherapies, thereby reducing the need for liver biopsies. This is particularly important in selection of patients in phase II studies where MRI-PDFF is the primary outcome. Using the MAST score in phase II studies may further identify those patients that resemble patients who are eligible for phase III trials.

Although a MAST score of more than 0.242 or less than 0.165 classified more than 80% of patients in both derivation and validation cohorts, clinicians will ultimately need to decide on

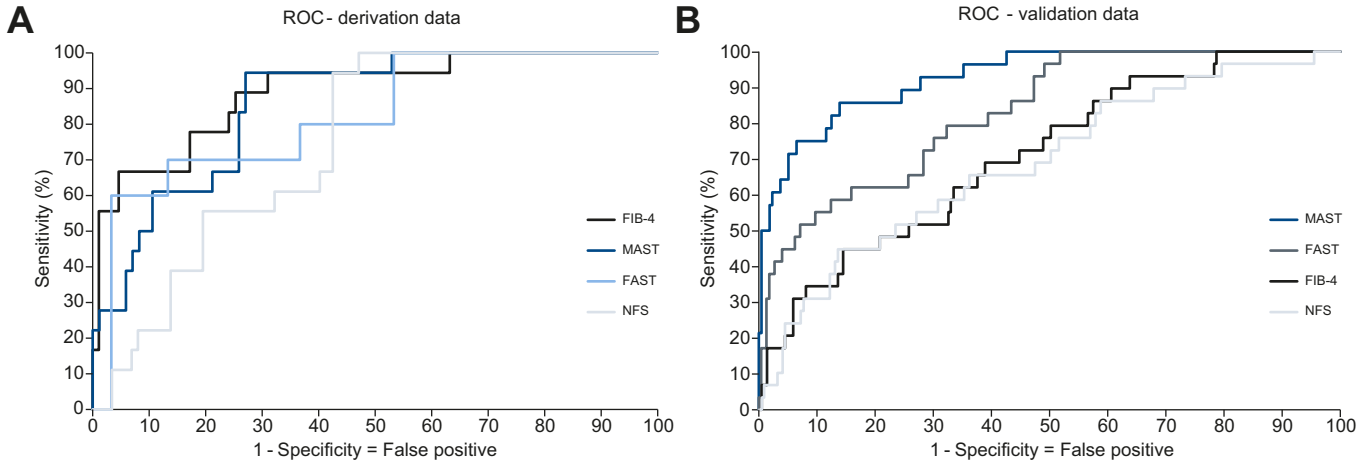


Fig. 1. Diagnostic performance of the MAST, FAST, FIB-4, and NFS scores for the diagnosis of Fibro-NASH. (A) All possible values of sensitivity *vs.* 1-specificity for the MAST, FAST, FIB-4, and NFS scores in the derivation data. (B) All possible values of sensitivity *vs.* 1-specificity for the MAST, FAST, FIB-4, and NFS scores in the validation data. FAST, FibroScan-aspartate aminotransferase; FIB-4, Fibrosis-4 index; Fibro-NASH, fibrotic non-alcoholic steatohepatitis; MAST, MRI-aspartate aminotransferase; NFS, NAFLD fibrosis score; NASH, non-alcoholic steatohepatitis; ROC, receiver operating curve.

Table 2. Diagnostic performance of different scores using 2 cut-offs for the diagnosis of Fibro-NASH.

	MAST		FAST	NAFLD (NFS)		FIB-4	
Sample	Derivation	Validation	Validation	Derivation	Validation	Derivation	Validation
AUC (95% CI)	0.858 (0.776-0.928)	0.929 (0.880-0.971)	0.868 (0.806-0.922)	0.748 (0.644-0.840)	0.689 (0.581-0.788)	0.891 (0.806-0.959)	0.711 (0.613-0.803)
Cut-off A	0.165	0.165	0.35	-0.678	-0.678	1.63	1.63
n	63	159	206	52	182	61	232
Sensitivity	94.40%	89.30%	65.5%	88.90%	51.70%	94.40%	20.70%
Specificity	72.90%	72.20%	86.7%	57.50%	76.00%	69.00%	94.60%
PPV	42.50%	29.40%	38.8%	30.20%	22.10%	38.60%	33.30%
NPV	98.40%	98.10%	95.1%%	96.20%	92.30%	98.40%	90.10%
Cut-off B	0.242	0.242	0.67	1.17	1.17	2.78	2.78
n	20	42	8	12	3	20	0
Sensitivity	61.10%	75.00%	20.7%	22.20%	3.40%	66.67%	0%
Specificity	89.40%	90.30%	99.1%	90.00%	99.10%	90.80%	100.00%
PPV	55.00%	50.00%	74.7%	31.50%	33.30%	60.00%	–
NPV	91.60%	96.50%	90.7%	84.80%	88.70%	92.90%	88.40%
Grey zone A-B, n (%)	20 (19.42%)	43 (17.62%)	41 (16.0%)	41 (39.05%)	65 (26.00%)	24 (22.86%)	18 (7.20% poorly calibrated)

The performances of the MAST, FAST, NAFLD, and FIB-4 scores using 2 cut-offs are assessed with the cut-offs calculated in the derivation cohort and applied in the validation cohort, except for FAST. FAST was applied in the validation cohort only as the derivation cohort had >50% missing CAP data since the probe was not available in the California center until later. For the 2 score cut-offs, the lower threshold comprises a rule-out cut-off corresponding to nearly 90% sensitivity in the derivation cohort, whereas the higher threshold comprises a rule-in cut-off corresponding to nearly 90% specificity in the derivation cohort. Individuals with a score in between these rule-out and rule-in cut-offs are in the grey zone.

AUC, area under the receiver-operating characteristic curve; FAST, FibroScan-aspartate aminotransferase; FIB-4, Fibrosis-4 index; MAST, MRI-aspartate aminotransferase; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; NPV, negative predictive value; PPV, positive predictive value.

further steps for patients in the grey zone. Such steps may include repeat testing after an appropriate time period or liver biopsy for definitive diagnosis. Clinical decision making should consider proximity to cut-offs, individual characteristics (*e.g.*, demographics, serologic results, and comorbidities), and patient choice. Importantly, the MAST score resulted in a narrower grey zone compared to other scores.

The major strengths of this study were its score's balanced accuracy and utilization of accurate, non-invasive MRI. Compared to the NFS and FIB-4, MAST had fewer patients in the grey zone and higher AUC. MAST's PPVs were not unexpected given that PPV is dependent on prevalence, which is on the lower end in our MAST's real-world cohort compared to other studies that focused on clinical trials, yet it likely represents the true prevalence in the general population. Few cohorts yet exist with paired MRI-PDFF, MRE and liver biopsy data, whereas the NIS4 and FAST scores consist of VCTE and histological samples that may be obtained with greater ease.[16,17] In contrast to NIS4 and FAST, the MAST score remains necessary due to its clinical applicability in the setting of clinical phase IIA trials, where MRI-based assessment and its results respectively constitute the primary inclusion criteria and primary endpoints.

Additionally, MRI has demonstrated a higher degree of balanced accuracy than VCTE in identifying both steatosis and fibrosis, yet this study is the first to our knowledge that utilizes MRI-PDFF and MRE in a screening algorithm for patients with Fibro-NASH.[19,23] Comparing MAST to FAST, MAST had similar or higher AUC and overall higher discrimination. Additionally, MAST demonstrated higher sensitivity, specificity, and PPV with some exceptions at the 90% sensitivity cut-off and at the 90% specificity cut-off. In addition to its accurate discrimination, the MAST score proposes score cut-offs that may be applied to select patients at risk of disease progression, thus demonstrating high clinical utility.

This study has limitations. First, this study is retrospective in nature. Nevertheless, both cohorts are part of natural history studies in tertiary care centers with standardized criteria for performing liver biopsy and non-invasive testing, as explained in the Patients and methods section. Second, derivation and validation of the MAST score were implemented in cohorts with large sample sizes but low disease (Fibro-NASH subpopulation) prevalence. However, these cohorts are representative of the large spectrum of patients with NAFLD, and remain similar to reported patient registries in Europe and the US.[11,24] Indeed, compared to those derived from clinical trial settings (which skew disease prevalence), our cohorts represent the real-world experience and thus increase MAST's generalizability and applicability in the clinical setting. It was also encouraging that MAST successfully identified most patients with Fibro-NASH, demonstrating consistency and balanced accuracy with an AUC of 92.9%. Despite low disease prevalence, MAST accurately identified more patients with Fibro-NASH and demonstrated better discrimination than FAST. Moreover, despite the difficulties of accessing paired MRI-PDFF, MRE and AST data, MAST's derivation and validation cohorts are 2 of the largest datasets that currently exist. Third, the cost of MRI can limit uptake in clinical practice, necessitating investment in devices and trained operating personnel, though similar models have been successfully introduced in other instances.[25] Fourth, histological readings were not done centrally, though performed by experienced hepatopathologists at each center. Finally, we have not explored

the differential costs of misclassification in this study due to lack of data availability for statistical analysis. Our study has many strengths as well. MAST has demonstrated high calibration in low prevalence populations, which represent true clinical settings. The introduction of the MAST score is of importance given its superior performance with increased accuracy, precise categorization, and high applicability in clinical trials. Because MRI-based assessment and its results respectively comprise the primary inclusion criteria and endpoints of phase IIA studies, the MAST score will better identify patients at higher risk of disease progression to fibrosis and cirrhosis and enrich the inclusion criteria of clinical trials that may revolutionize therapeutic pharmacotherapies for NAFLD.[14]

In summary, the MAST score non-invasively identifies patients at higher risk of disease progression for clinical trial eligibility and pharmacotherapies.

Abbreviations

ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the ROC curve; CAP, controlled attenuation parameters; FAST, FibroScan-aspartate aminotransferase; FIB-4, fibrosis-4 index; Fibro-NASH, fibrotic non-alcoholic steatohepatitis; LR, likelihood ratio; LSM, liver stiffness measurement; MAST, MRI-aspartate aminotransferase; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-derived proton density fat fraction; NAFLD, non-alcoholic fatty liver disease; NAS, NAFLD activity score; NASH, non-alcoholic steatohepatitis; NFS, NAFLD fibrosis score; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver-operating characteristic; VCTE, vibration-controlled transient elastography.

Financial support

The authors received no financial support to produce this manuscript.

Conflict of interest

MN has been on the advisory board for 89BIO, Gilead, Intercept, Pfizer, Novartis, Novo Nordisk, Allergan, Blade, EchoSens, Fractyl, Terns, OWL, Siemens, Roche diagnostic and Abbott; MN has received research support from Allergan, BMS, Gilead, Galmed, Galectin, Genfit, Conatus, Enanta, Madrigal, Novartis, Shire, Viking and Zydus; MN is a minor shareholder or has stocks in Anaetos and Viking. NA has been on the advisory board for Gilead and Allergan. NA is on the speaker bureau for Intercept and Gilead. SH: advisory boards/consultant for Akero Therapeutics, Alentis Therapeutics, Alimentiv, Altimmune Inc., Arrowhead Pharmaceuticals, Axcella Health, Corcept Therapeutics, Chronwell, Echosens North America, Enyo Pharma, Foresite Labs (MetreaBiosciences), Galectin Therapeutics, Genfit, Gilead Sciences, Hepion Pharmaceuticals, Hightide Therapeutics, Histoindex, Intercept, Madrigal Pharmaceuticals, Medpace, NGM Biopharmaceuticals, Northsea Therapeutics, Novartis, Novo Nordisk, PathAI, Poxel, Sagimet Biosciences, Sonic Incytes, Terns, Theratechnologies, and Viking. Stock options/minor shareholder: Akero, Chronwell, Cirius, Galectin, Genfit, Hepion, HistoIndex, Metacrine, NGM Bio, Northsea. Grant/Research support: Akero, Axcella, BMS, Cirius, CiVi Biopharma, Conatus, Cymabay, Enyo, Galectin, Genentech, Genfit, Gilead, Hepion, Hightide, Intercept, Madrigal, Metacrine, NGM Bio, Novartis, Novo Nordisk, Northsea, Pfizer, Sagimet, Viking, 89 Bio. NA: advisory board for Echosens, Gilead, Intercept, Novo Nordisk, Perspectum, Pfizer, and Zydus; grant/research support from 89Bio, Akero, Bristol Myers Squibb, Enanta, Enyo, Galectin, Genentech, Gilead, Intercept, Inventiva, Madrigal, Metacrine, NGM Bio, Novo Nordisk, Pfizer, Poxel, Viking, and Zydus; speaker bureau for AbbVie, Alexion, Echosens, Gilead, and Intercept.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

MN participated in the design of the study, supervised the study, interpreted the data, and drafted the manuscript. ET interpreted the data and drafted the manuscript. JG performed the statistical analysis. RS, MG, TT, NN, JDY, SH interpreted the data and critically revised the manuscript for important intellectual content. NA interpreted the data and critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Data availability statement

This study is a retrospective in nature with no therapeutic intervention. The study data are confidential. Drs. Mazen Noureddin, Naim Alkouri and Stephen A. Harrison have access to the data. De-identified data might be provided upon request in compliance with patient privacy rules.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.11.012.

References

Author names in bold designate shared co-first authorship

[1] Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2018;15(1):11–20.
[2] Charlton MR, Burns JM, Pedersen RA, Watt KD, Heimbach JK, Dierkhising RA. Frequency and outcomes of liver transplantation for nonalcoholic steatohepatitis in the United States. Gastroenterology 2011;141(4):1249–1253.
[3] Haldar D, Kern B, Hodson J, Armstrong MJ, Adam R, Berlakovich G, et al. Outcomes of liver transplantation for non-alcoholic steatohepatitis: a European Liver Transplant Registry study. J Hepatol 2019;71(2):313–322.
[4] **Noureddin M**, **Vipani A**, Bresee C, Todo T, Kim IK, Alkhouri N, et al. NASH leading cause of liver transplant in women: updated analysis of indications for liver transplant and ethnic and gender variances. Am J Gastroenterol 2018;113(11):1649–1659.
[5] Setiawan VW, Stram DO, Porcel J, Lu SC, Le Marchand L, Noureddin M. Prevalence of chronic liver disease and cirrhosis by underlying cause in understudied ethnic groups: the multiethnic cohort. Hepatology 2016;64(6):1969–1977.
[6] Geier A, Rinella M, Balp MM, McKenna SJ, Brass C, Przybysz R, et al. Real-world burden of nonalcoholic steatohepatitis. Clin Gastroenterol Hepatol 2020:30928–30929. S1542-3565(20).
[7] Guha IN, Parkes J, Roderick P, Chattopadhyay D, Cross R, Harris S, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology 2008;47(2):455–460.
[8] Sanyal AJ, Friedman SL, McCullough AJ, Dimick-Santos L. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American association for the study of liver diseases-U.S. Food and drug administration joint workshop. Hepatology 2015;61(4):1392–1405.
[9] Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007;45(4):846–854.

[10] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41(6):1313–1321.
[11] Kleiner DE, Brunt EM, Wilson LA, Behling C, Guy C, Contos M, et al. Association of histologic disease activity with progression of nonalcoholic fatty liver disease. JAMA Netw Open 2019;2(10):e1912565.
[12] Taylor RS, Taylor RJ, Bayliss S, Hagström H, Nasr P, Schattenberg JM, et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. Gastroenterology 2020;158(6):1611–1625 e12.
[13] Sanyal AJ, Harrison SA, Ratziu V, Abdelmalek MF, Diehl AM, Caldwell S, et al. The natural history of advanced fibrosis due to nonalcoholic steatohepatitis: data from the simtuzumab trials. Hepatology 2019;70(6):1913–1927.
[14] Noureddin N, Schattenberg JM, Alkhouri N, Noureddin M. Noninvasive testing using magnetic resonance imaging techniques as outcomes in nonalcoholic steatohepatitis clinical trials: how full is the glass? Hepatol Commun 2020;4(2):141–144.
[15] Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. Hepatology (Baltimore, Md.) 2012;56(5):1751–1759.
[16] Harrison SA, Ratziu V, Boursier J, Francque S, Bedossa P, Majd Z, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5(11):970–985.
[17] Newsome PN, Sasso M, Deeks JJ, Paredes A, Boursier J, Chan WK, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5(4):362–373.
[18] Dennis A, Mouchti S, Kelly M, Fallowfield JA, Hirschfield G, Pavlides M, et al. A composite biomarker using multiparametric magnetic resonance imaging and blood analytes accurately identifies patients with non-alcoholic steatohepatitis and significant fibrosis. Sci Rep 2020;10:15308.
[19] Park CC, Nguyen P, Hernandez C, Bettencourt R, Ramirez K, Fortney L, 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.
[20] Harrison SA, Gawrieh S, Roberts K, et al. Prospective evaluation of the prevalence of non-alcoholic fatty liver disease and steatohepatitis in a large middle-aged US cohort. J Hepatol 2021. https://doi.org/10.1016/j.jhep.2021.02.034.
[21] Newsome PN, Sasso M, Deeks JJ, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol 2020;5(4):362–373. https://doi.org/10.1016/S2468-1253(19)30383-8.
[22] Brunt EM, Janney CG, di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Off J Am Coll Gastroenterol | ACG 1999;94(9). https://journals.lww.com/ajg/Fulltext/1999/09000/Nonalcoholic_Steatohepatitis__A_Proposal_for.31.aspx.
[23] Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, quantitative assessment of liver fat by MRI-PDFF as an endpoint in NASH trials. Hepatology 2018;68(2):763–772.
[24] McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. J Hepatol 2015;62(5):1148–1155.
[25] Srivastava A, Gailer R, Tanwar S, Trembling P, Parkes J, Rodger A, et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. J Hepatol 2019;71(2):371–378.