Invited Review

Role of noninvasive tests in the prognostication of metabolic dysfunction-associated steatotic liver disease

Short title: Prognostication of MASLD

Yue Wang¹,²*, Sherlot Juan Song¹,²*, Yichong Jiang¹,²*, Jimmy Che-To Lai¹,², Grace Lai-Hung Wong¹,², Vincent Wai-Sun Wong¹,², Terry Cheuk-Fung Yip¹,²

¹Medical Data Analytic Center, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China
²State Key Laboratory of Digestive Disease, Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China

*These authors contributed equally to this work.

Correspondence:
Terry Cheuk-Fung Yip, PhD
Medical Data Analytic Center, Department of Medicine and Therapeutics, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong, China
Email: tcfyp@cuhk.edu.hk

Vincent Wai-Sun Wong, MD
Department of Medicine and Therapeutics, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong, China
Email: wongv@cuhk.edu.hk

Keywords: MASLD; metabolic-associated fatty liver disease; nonalcoholic fatty liver disease; cirrhosis; hepatocellular carcinoma; hepatic decompensation
**Funding:** Vincent Wong’s research in steatotic liver disease is supported in part by the General Research Fund from the Hong Kong SAR Government (project reference 14106923).

**Disclosures:** Jimmy Lai has served as an advisory committee member for Gilead Sciences and Boehringer Ingelheim, and a speaker for Gilead Sciences. Grace Wong has served as an advisory committee member for Gilead Sciences and Janssen, as a speaker for Abbott, Abbvie, Bristol-Myers Squibb, Echosens, Furui, Gilead Sciences, Janssen and Roche, and received research grant from Gilead Sciences. Vincent Wong has served as a consultant or advisory board member for AbbVie, Boehringer Ingelheim, Echosens, Gilead Sciences, Intercept, Inventiva, Novo Nordisk, Pfizer, Sagimet Biosciences, TARGET PharmaSolutions, and Visirna; and a speaker for Abbott, AbbVie, Echosens, Gilead Sciences, Novo Nordisk, and Unilab. He has received a research grant from Gilead Sciences, and is a co-founder of Illuminatio Medical Technology. Terry Yip has served as a speaker and an advisory committee member for Gilead Sciences. The other authors declare that they have no competing interests.

**Author Contributions:** All authors were responsible for the writing plan, content, drafting and critical revision of the manuscript for important intellectual content.
In managing metabolic dysfunction-associated steatotic liver disease, which affects over 30% of the general population, effective noninvasive biomarkers for assessing disease severity, monitoring disease progression, predicting the development of liver-related complications, and assessing treatment response are crucial. The advantage of simple fibrosis scores lies in their widespread accessibility through routinely performed blood tests and extensive validation in different clinical settings. They have shown reasonable accuracy in diagnosing advanced fibrosis and good performance in excluding the majority of patients with a low risk of liver-related complications. Among patients with elevated serum fibrosis scores, a more specific fibrosis and imaging biomarker has proved useful to accurately identify patients at risk of liver-related complications. Among specific fibrosis blood biomarkers, enhanced liver fibrosis is the most widely utilized and has been approved in the United States as a prognostic biomarker. For imaging biomarkers, the availability of vibration-controlled transient elastography has been largely improved over the past years, enabling the use of liver stiffness measurement (LSM) for accurate assessment of significant and advanced fibrosis, and cirrhosis. Combining LSM with other routinely available blood tests enhances the ability to diagnose at-risk metabolic dysfunction-associated steatohepatitis; and predict liver-related complications, some reaching an accuracy comparable to that of liver biopsy. Magnetic resonance imaging-based modalities provide the most accurate quantification of liver fibrosis, though the current utilization is limited to research settings. Expanding their future use in clinical practice depends on factors such as cost and facility availability.
INTRODUCTION

Metabolic dysfunction-associated steatotic liver disease (MASLD), previously known as nonalcoholic fatty liver disease (NAFLD) or metabolic-associated fatty liver disease, affects over 30% of the general population and is emerging as a leading cause of cirrhosis and hepatocellular carcinoma (HCC) (1). Its active form, metabolic dysfunction-associated steatohepatitis (MASH), is associated with accelerated fibrosis progression and a higher risk of liver-related complications (2). Traditionally, the diagnosis of MASH and liver fibrosis required a liver biopsy. However, liver biopsy is an invasive procedure with a small risk of bleeding. Its accuracy in detecting MASH and fibrosis is also limited by sampling variability, intraobserver and interobserver variability (3). Besides, although serial liver biopsies are often required to demonstrate MASH resolution and fibrosis improvement in clinical trials, this approach is unrealistic in real-world practice. It is thus important to develop noninvasive tests to assess the severity of MASLD.

To qualify as a surrogate endpoint, a biomarker has to demonstrate the ability to diagnose (i.e., MASLD, MASH or fibrosis), prognosticate (i.e., predicting future development of liver-related events [LREs] such as HCC, hepatic decompensation and liver-related death), monitor (i.e., detecting patients who have disease progression over time), or determine treatment response (i.e., detecting patients who improve over time after treatment) (4). Emerging data support that some existing noninvasive tests may already fulfill some of these requirements. Among various biomarkers, noninvasive tests of fibrosis have been most extensively evaluated and adopted in clinical practice, especially as liver fibrosis has the strongest correlation with LREs among various histological parameters (5). Some of these tests also demonstrate the ability to predict the
presence of clinically significant portal hypertension (CSPH) and high-risk varices. In this article, we review the diagnosis, stratification and prognostication of MASLD by noninvasive tests and discuss the implications on clinical practice and future research. In particular, according to the Food and Drug Administration-National Institute of Health Biomarker Working Group, a prognostic biomarker is “a biomarker used to identify the likelihood of a clinical event, disease recurrence, or progression in patients who have the disease or medical condition of interest” (4).

Notably, most studies on noninvasive tests were published in the NAFLD era, whereas the diagnosis of MASLD requires the presence of at least one cardiometabolic risk factor (6). However, because the vast majority of patients with NAFLD would have cardiometabolic risk factors anyway, studies have consistently shown over 95% overlap between the NAFLD and MASLD populations (7). Not surprisingly, the two populations showed similar disease characteristics and natural history (8-10). For the purpose of this review, data from NAFLD are considered applicable to the MASLD population unless otherwise specified.

**SIMPLE FIBROSIS SCORES**

Examples of conventional simple fibrosis scores developed on widely available serum parameters include aspartate aminotransferase (AST)-to-platelet ratio index (APRI), Fibrosis-4 index (FIB-4), and NAFLD fibrosis score (NFS). All three scores have good diagnostic accuracy for liver fibrosis and particularly for advanced fibrosis among MASLD with the moderate area under the receiver-operating characteristic curve (AUROC)(4, 11, 12) (Table 1). FIB-4 is the most validated and useful test among the three. It is recommended as the first-line assessment for liver fibrosis due to its
superior diagnostic accuracy compared to other non-invasive tests (11, 13, 14). All three panels have negative predictive values (NPV) exceeding 90%, thus effectively ruling out advanced fibrosis (11). Consequently, they can be utilized as first-tier screening tools to select patients from low-risk populations requiring further investigations. However, these scores are less accurate in detecting milder degrees of fibrosis (15, 16) (Table 1). Also, data suggested that APRI has inferior performance to FIB-4 and NFS in detecting advanced fibrosis (15).

**Prediction of cirrhotic complications and HCC**

The three fibrosis scores showed consistently good prediction for LREs including cirrhotic complications and HCC among patients with MASLD (Table 1)(17). Based on validated cutoffs of FIB-4, NFS and APRI, patients stratified into intermediate- or high-risk groups, compared with the low-risk group, had a significantly higher risk of LREs (18, 19). APRI has the lowest accuracy among the three scores in predicting LRE and overall mortality (17, 20, 21).

FIB-4 is the most studied and validated tool for accurately risk stratifying patients with MASLD (17, 22). Despite evidence showing that vibration-controlled transient elastography (VCTE) and histological assessment outperformed FIB-4 in their accuracy of predicting LREs, FIB-4 remains an independent predictor of LREs, and the FIB4-VCTE stepwise algorithm accurately identifies high-risk individuals with MASLD for liver-related complications (22). Moreover, an individual participant data meta-analysis suggests that FIB-4 and NFS perform similarly to liver biopsy in predicting LREs (23). This is further proved among patients with MASH (24), indicating these
panels could be considered as alternatives to unnecessary invasive methods for monitoring the risk of LREs.

Elevated FIB-4 also represents an important predictor of HCC in European and Asian MASLD cohorts (25, 26). A high NFS is also strongly associated with HCC development. As such, the American Gastroenterological Association recommended that patients with non-invasive scores suggestive of cirrhosis should be screened for HCC (27).

In contrast, several studies suggest that FIB-4 might not predict overall and cardiovascular mortality (28-31). This could be related to the study setting and patient composition.

The steatosis-associated fibrosis estimator (SAFE) score and LiverRisk score are examples of fibrosis scores that have been developed more recently, both utilizing simple serum biomarkers (32, 33). SAFE score was designed to aid in triaging patients at the lowest risk of clinically significant fibrosis (F0-1 vs. ≥F2)(34), while LiverRisk score was developed to predict individual levels of liver stiffness for the general population without chronic liver disease. Both scores outperformed conventional blood-based algorithms in detecting liver fibrosis and identifying high-risk individuals for liver-related outcomes (32, 33).

Fibrosis stage is a strong predictor for the development of adverse liver outcomes in MASLD (35). Advanced fibrosis as determined by simple fibrosis scores is associated with higher mortality (36). Data suggest that patients who develop advanced fibrosis have greater changes in NFS and FIB-
4 (37). However, whether repeated measures of these panels over time would carry the prognostic importance for predicting adverse outcomes remains to be demonstrated.

Although the predictive ability of FIB-4, NFS, and APRI for fibrosis progression was found to be suboptimal with AUROC levels ranging from 0.65 to 0.73 (17), it is important to note that a significant association was found between a one-unit change in all these tests and the corresponding change in fibrosis stage (13), which underscores the relevance of minor score changes in these biomarkers and liver progression. However, Siddiqui et al. found a significant increase in all three scores only among individuals who showed progression in fibrosis, but not among those who experienced regression (13). This suggests that improvement in FIB-4 can be primarily driven by a falsely high score at the first test and subsequent regression towards the mean (38). These findings highlight the importance of exercising caution about false positives when interpreting changes in repeated testing to identify high-risk individuals.

A Swedish population-based cohort study found that a transition from low- to intermediate- or high-risk group by repeated FIB-4 testing within five years is associated with increased severity of liver disease (38). This finding was also supported by a recent UK population-based cohort study that changes in FIB-4 levels over an ever shorter interval of 12 months predicted the risk of subsequent liver events (18). Another large cohort study showed a strong association between longitudinal changes in FIB-4 and subsequent development of both HCC and LREs by a three-year landmark analysis (39). Furthermore, changes in FIB-4 are found to be associated with cardiovascular events and all-cause mortality (18). This is of particular significance given the high
prevalence of metabolic diseases observed within the FIB-4 elevated group in the population without known liver disease (40), suggesting the longitudinal utility of these panels in both liver- and non-liver-related adverse outcomes. Additionally, a Phase 2b clinical trial found significant correlations between reductions in FIB-4, APRI, and NFS and improvement in fibrosis after the treatment with obeticholic acid (41). This correlation along with the confirmed obeticholic acid-related benefits for fibrosis indicates the prognostic utility of these measures in predicting treatment response in future clinical trials.

Utilization of serum biomarkers for fibrosis scoring has practical advantages of wide applicability and easy availability (42). However, the tests are less accurate than patented markers such as VCTE with modest diagnostic value (positive predictive value [PPV] around 70%) and can thus lead to unnecessary referral to hepatology clinics (42, 43). More specific assessments with higher diagnostic accuracy are required to rule in positive cases more efficiently.

Furthermore, these tests utilize non-liver-specific biomarkers, and challenges are present with the reproducibility. The results are also influenced by spectrum bias, e.g., acute hepatitis can lead to false positive results of FIB-4 and APRI due to reactive surge in aminotransferase levels; both younger and older age (<35 y; >65 y) are suggested to limit the accuracy of FIB-4 and NFS (4, 42). The interpretation of these scores thus requires additional caution to avoid unnecessary examination and undue stress for healthy individuals and healthcare systems.
FIB-4 and APRI were derived from smaller cohorts of patients with chronic hepatitis C with a high prevalence of liver fibrosis (42). NFS was derived from the population with MASLD for separation of those with and without advanced fibrosis (44). These scores were developed in the population with existing liver disease and thus limited in their generalizability and role as a screening tool. There is a clear need for the development of scoring systems in primary care settings for early risk detection and tailored interventions to reduce overall liver disease burden. However, we should not deny the constant usefulness and prognostic value of these scoring systems.

**SPECIFIC FIBROSIS BIOMARKERS**

Specific fibrosis biomarkers for MASLD may serve as the first- or second-line detection methods. Examples include the enhanced liver fibrosis (ELF) score, PRO-C3, FibroTest, FibroMeter, and Mac2-binding protein (Table 2). These markers reflect the dynamic changes of inflammatory cytokines, fibrogenesis or fibrinolysis, vascular, and neuroendocrine signals in the body. Among them, ELF is more widely used and has been approved by the US Food and Drug Administration as a prognostic biomarker. ELF combines three serum biomarkers, namely hyaluronic acid, tissue inhibitor of metalloproteinase 1, and N-terminal propeptide of type III procollagen, providing a more reliable basis for assessing and predicting fibrosis and cirrhosis compared to individual indicators (45). A meta-analysis indicated that when using a cut-off of 7.7, ELF has a high sensitivity of 0.93 and a specificity of 0.34 for excluding fibrosis. In populations with disease prevalence <50%, adopting a low cut-off value for excluding advanced fibrosis (F≥3) yields a sensitivity of 0.90, with an NPV ranging from 0.82 to 0.99 (46). In a study conducted in Denmark, the false positive rate of ELF testing for screening the general population for liver disease was
lower than that of FIB-4, making it an effective tool for screening high-risk populations for liver fibrosis and reducing the occurrence of ineffective referral pathways (47).

PRO-C3 has been validated as an independent factor associated with the fibrotic stage and related to the NAFLD activity score, with an AUROC of 0.81 for detecting significant fibrosis and an AUROC of 0.79 for advanced fibrosis (48). In a cohort study involving patients with diabetes, PRO-C3 has been confirmed to identify rapid disease progression and patients who respond to antimetabolic treatments, and potentially accelerate the development and validation of antifibrotic medications (49). Additionally, the ADAPT score, which was developed in the MASLD population, incorporating PRO-C3 together with platelet count, diabetes, and age, outperforms APRI and FIB-4 in identifying patients with advanced fibrosis, suggesting its potential as a routine blood biochemical marker for screening (50).

FibroTest utilizes age and gender information, α-2-macroglobulin, haptoglobin, apolipoprotein A1, GGT, cholinesterase, and total bilirubin to assess liver fibrosis. Table 2 reveals that the FibroTest applies a 90% sensitivity exclusion threshold for MASLD at a cut-off score of 0.3, and a 90% sensitivity inclusion criterion at a score of 0.7. Additionally, the FibroTest demonstrated a 72% sensitivity and an 85% specificity for accurately detecting the significant fibrosis with a cut-off range from 0.30 to 0.48 (51).
FibroMeter includes liver function indicators, coagulation indicators, urea, and hyaluronic acid. It has been used to evaluate liver fibrosis in different chronic liver diseases. A cross-sectional study involving 452 patients with MASLD demonstrated that FibroMeter exhibited comparable accuracy to liver stiffness measurement (LSM) in detecting fibrosis stage, with greater predictive ability in cohorts with advanced fibrosis (52). Furthermore, FibroMeter has better discriminatory ability than other blood markers (Obuchowski index 0.798 ± 0.016), with a cutoff of 0.31 recommended for patients with advanced fibrosis to achieve maximum sensitivity and specificity.

Mac2-binding protein is a secretory protein composed of a carbohydrate-rich N-terminal and a galectin-like domain. Elevated levels of wisteria floribunda agglutinin-positive mac-2 binding protein in the blood reflect the progression of fibrosis. Moreover, serum mac-2 binding protein glycosylation isomer (M2BPGi) has been clinically validated as a glyco-biomarker for predicting cirrhosis in patients with MASLD (53, 54). Additionally, in a Japanese cohort, high levels of M2BPGi have been found to be independently associated with an increased incidence of HCC (55). Finally, although created in patients with chronic hepatitis C, hepscore and MMP/MP3 have also been identified as specific biomarkers with predictive value for the degree of liver fibrosis, but further evidence related to MASLD is still needed.

In a retrospective cohort study recruiting 453 patients with chronic hepatitis B in Hong Kong, a two-step scoring system combining ELF score (cutoff 9.8) and LSM (cutoff 20 kPa) effectively predicted HCC, with a sensitivity of 86.7% and NPV of 95.3% (56). This model allows for HCC
management based on different ELF risk levels, although further data for MASLD cohorts are needed. Another prospective cohort study in Denmark involving 462 patients followed for an average of 4.1 years confirmed the predictive role of ELF and FibroTest for LREs, with C-indexes of 0.859 and 0.794, respectively. The time-dependent AUROC for the prediction of 5-year LRE reached 0.87, higher than FIB-4 (0.83) and NFS (0.78) (57). Additionally, advanced fibrosis increases the burden of LREs, with serological markers showing high predictive value in MASLD patient populations, particularly for advanced fibrosis stages. Moreover, this study suggests that an ELF score with a cut-off of 9.1 for advanced fibrosis achieves better diagnostic accuracy than the manufacturer-recommended value of 7.7 (58). Another meta-analysis mentioned that FibroMeter testing showed acceptable sensitivity (83.1%) and specificity (84.4%) in detecting late-stage fibrosis in MASLD patients. When FibroMeter was integrated with VCTE, the sensitivity increased to 83.5%, while specificity reached 91.1% (59).

Nevertheless, serological testing for predicting MASLD in the general population has limitations, including high economic costs and its inability to be used independently for early diagnosis and patient referral. Consequently, its utilization as a screening tool is less frequent compared to APRI or FIB-4 (60). Moreover, the accuracy of ELF in detecting fibrosis varies depending on the etiology of the disease, with ELF showing lower accuracy for staging in MASLD compared to alcohol-associated liver disease and chronic hepatitis C. Therefore, it is recommended to discuss further the more specific cut-off values for ELF in diagnosing different fibrosis stage in MASLD (61). Furthermore, not every biomarker was initially designed for MASLD prognostication. For example,
ELF score and FibroTest were developed in patients with chronic hepatitis C. Additionally, PRO-C3 is not a liver-specific marker; it is also increased in pulmonary fibrosis (48).

**IMAGING BIOMARKERS**

In the past decades, imaging biomarkers have been extensively investigated in the MASLD population. VCTE is currently widely used for disease screening and prognosis assessment. In addition, there are emerging technologies and scoring systems that have been developed and examined, and have been shown to be of significant importance in the assessment of liver fibrosis (Table 3).

**Vibration-controlled transient elastography**

VCTE is the most commonly employed diagnostic technique. VCTE utilizes pulse-echo ultrasound to evaluate LSM and controlled attenuation parameter (CAP)(62, 63). Steatosis, advanced fibrosis and cirrhosis can be detected with high sensitivity and specificity (64, 65). In a prospective study of MASLD population, the AUROC of LSM by VCTE was 0.89 when evaluating advanced fibrosis (66). LSM has shown promising results in predicting the presence and severity of hepatic decompensation and is also associated with the risk of HCC (52, 67-69). According to a prospective study of 2251 MASLD patients with a median follow-up of 27 months, VCTE could accurately predict LREs, with a C-index of 0.911 (95% CI 0.816-0.993, p = 0.002)(68). Similarly, in a study among 594 participants with MASLD, VCTE demonstrated comparable accuracy to histologically assessed fibrosis in predicting clinical outcomes, outperforming FIB-4 in the whole study cohort (70). In another multicenter study of 16603 patients with MASLD, VCTE
demonstrated superior performance in predicting hepatic decompensation and HCC (71). Importantly, a change in VCTE-LSM also appears to be prognostic. In patients with LSM >15 kPa at baseline, the incidence of liver-related events was 7.8 per 1000 person-years in patients whose LSM decreased to <10 kPa during follow-up, compared with 38.7 per 1000 person-years if LSM remained >15 kPa. Despite being a blind technique, VCTE is widely utilized for monitoring treatment efficacy, as well as long-term follow-up in diverse chronic liver diseases. Changes in LSM by VCTE were found to be correlated with transitions in histological fibrosis stage, liver biochemistry, and other metabolic-related clinical parameters in phase III clinical trials (72). Although VCTE is recommended by guidelines as a robust and consistent tool for fibrosis assessment, screening high-risk patients and progression monitoring (73, 74), it has the limitation of being affected by technical factors and patient-related conditions like obesity, food or alcohol consumption and acute inflammation (75).

Scoring systems combining VCTE with other parameters have been developed to identify patients at high risk (76). FibroScan-AST (FAST) score, comprised of LSM, CAP and serum AST, is a screening test for fibrotic MASH. Patients with FAST score ≥0.67 are more likely to have at-risk MASH, and FAST score ≤0.35 suggests low risk (77). The numerical variation of FAST score can also reflect changes in histological fibrosis stage (78). FAST score has good accuracy (C-index 0.80, 95% CI 0.76-0.85) and excellent sensitivity for the 0.35 cut-off and specificity for the 0.67 cut-off (77). However, its PPV is relatively low, and the 0.67 cut-off is associated with a high case missing rate of 51.7% (77). Agile 3+ and Agile 4 are novel scores based on LSM by VCTE, routinely available clinical parameters and demographic features to identify advanced fibrosis or cirrhosis (79).
cohort studies verifying its performance, the Agile scores were superior to FIB-4 and LSM in terms of AUROC, percentage of patients with indeterminate outcomes, and PPV to rule-in cirrhosis or advanced fibrosis (79). For predicting cirrhosis, Agile 4 achieved optimal combinations with 85% sensitivity for ruling out and 95% specificity for ruling in. In the case of predicting advanced fibrosis, Agile 3+ demonstrated rule-out sensitivity of 85% and rule-in specificity of 90% (80). Agile scores also showed excellent capability in LRE prediction (80, 81). The potential use of VCTE-based scores to monitor disease progression and their use to predict clinical outcomes needs to be further investigated.

**Point shear wave elastography and two-dimensional shear wave elastography**

Shear wave elastography (SWE) is another imaging technique utilizing an acoustic radiation force impulse to generate shear waves within liver tissue visualized, which can reduce measurement failure (82). Point shear wave elastography (pSWE) shows an excellent prediction result in advanced fibrosis and cirrhosis in patients with MASLD (83, 84), regardless of the presence of ascites (85-88). Previous studies mainly focused on predicting liver fibrosis and associated complications in populations with chronic viral hepatitis (89). In a recent systematic review, only MRE and pSWE met the ideal criteria of larger than 80% sensitivity and specificity for the diagnosis of advanced fibrosis (84). However, the performance of pSWE varied in different studies (90-92). Compared to other imaging biomarkers, steatosis significantly impacts the results of pSWE (92). In addition, the narrow range of values for pSWE restricts the accurate thresholds for the fibrosis stage, which leads to potential variations in the results (93).
Two-dimensional shear wave elastography (2D-SWE) can measure a larger area in the liver and enables the generation of real-time and quantitative elasticity maps of liver tissue, providing visual differentiation of stiff tissues (87). Unlike VCTE, inflammation does not affect the results of 2D-SWE in MASLD patients (94). Recent studies developed deep learning radiomics of elastography (DLRE) which applies the radiomic methodology to perform quantitative analysis of the heterogeneity observed in 2D-SWE images, showing excellent diagnostic performance in assessing liver fibrosis in patients with chronic hepatitis B (95). Among patients with MASLD, a prospective study reported an AUROC of 0.82 for predicting fibrosis risk and 0.90 for diagnosing advanced fibrosis when utilizing 2D-SWE (96). Currently, 2D-SWE is primarily utilized in research due to insufficient evidence supporting their efficacy in monitoring disease progression and diagnosing associated complications (62).

**Magnetic resonance elastography and Iron-corrected T1**

Magnetic resonance elastography (MRE) is based on an improved phase-contrast method to image the propagation of shear waves within liver tissue, providing a more quantitative assessment of significant fibrosis, advanced fibrosis, cirrhosis and MASH than ultrasound-based elastography (15, 97, 98). As reported by Xiao et al. (99), MRE performed better than other NITs for detecting advanced fibrosis, with an AUROC of 0.96, sensitivity of 0.84, and specificity of 0.90. MRE is also a useful tool for predicting and monitoring cirrhotic complications. In a systematic review of 1707 MASLD patients, individuals whose LSM by MRE were from 5 to 8 kPa had a hazard ratio of 11.0 (95% CI 7.03-17.1, P < 0.001) for liver-related outcome, and 15.9 (95% CI 9.32-27.2, P < 0.001) for those whose result ≥ 8 kPa, which indicates a significantly higher risk compared to
participants with MRE measurements <5 kPa (100). Other cohort studies also showed that elevated liver stiffness by MRE was correlated with the presence of ascites, hepatic encephalopathy, variceal bleeding, and overall mortality (98, 101). In contrast, reduction in MRE reflects fibrosis stage improvement with an AUROC of 0.62 (95% CI: 0.46–0.78) (102). The combination of MRE and FIB-4 is developed as MEFIB (MRE-FIB-4) for assessing fibrosis stage. MEFIB showed associations with disease progression like decompensation and HCC (76), and the NPV was over 95% in ruling out liver disease progression (103, 104). Due to its diagnostic accuracy for liver fibrosis, MRE is recommended as a promising surrogate measure for monitoring liver disease progression and evaluating therapeutic endpoints (105).

Iron-corrected T1 (cT1) by magnetic resonance imaging (MRI) was used in the UK Biobank population health study, and was designed to evaluate a combination of liver inflammation and fibrosis (106-108). A meta-analysis examined the diagnostic ability of cT1 and reported its high sensitivity but relatively low specificity (84). In a study of 50 patients with biopsy-proven MASLD, cT1 differentiated simple steatosis from MASH with an AUROC of 0.69 (109). In a general hepatology outpatient setting, cT1 showed predictive ability for various liver-related outcomes (110). In longitudinal disease monitoring, cT1 had better prediction ability of LREs in patients with cirrhosis when taking into account technical failure and unreliability (HR 9.9, 95% CI 1.29–76.4; P=0.007) (111). Based on MRI technique, MRI-proton density fat fraction (MRI-PDFF) has been examined as a useful tool for treatment response evaluation, providing more accurate results than CAP in detecting all grades of steatosis in MASLD patients (97). By combining MRI-PDFF,
MRE result and serum AST level, MRI-AST (MAST) score was designed for fibrotic MASH diagnosis. It had an impressive AUROC of 0.93, which outperformed FAST score, FIB-4, and NFS (112).

MRI-based techniques possess the benefit of being accurate, operator-independent and less influenced by patient-related factors (113), while their utilization is restricted in clinical trials due to the high cost and unavailability of facilities for daily clinical practice(76).

**PORTAL HYPERTENSION IN MASLD**

MASLD ranges from simple steatosis and MASH with or without liver fibrosis, to cirrhosis. With the progressive distortion in architecture of the liver parenchyma and hepatic microcirculatory dysfunction leading to increased intrahepatic vascular resistance, portal hypertension ensues and is an important component in the natural history of cirrhosis (114, 115).

As the gold standard, portal pressure is assessed by hepatic venous pressure gradient (HVPG) in which a balloon-tipped catheter is inserted into the hepatic vein under fluoroscopic guidance to acquire the wedged hepatic venous pressure (WHVP) and free hepatic venous pressure (FHVP), and HVPG is determined by subtracting the FHVP from WHVP (116-118). A HVPG of ≥5mmHg signifies portal hypertension and that of ≥10mmHg suggests CSPH.(118) CSPH is an important hallmark as it determines a significantly higher risk of hepatic decompensation (119, 120). However, HVPG measurement is constrained by its invasiveness and potential complications, limiting it for academic purposes most of the time. As well, several observational studies suggested that HVPG tends to underestimate portal pressure gradient in MASH-related cirrhosis.
due to an underestimation of WHVP in the actual portal vein pressure, with the development of
decompenating events while HVPG is <10mmHg (121-124). On the other hand, portal
hypertension may precede liver fibrosis in MASLD, suggesting a bidirectional effect in these
pathological processes (125). Given the increasing burden of MASLD worldwide, the number of
patients with MASH-related cirrhosis and thus portal hypertension will surge (126). It is thus
unrealistic to perform HVPG in a huge number of patients, particularly with the uncertainty in its
diagnostic accuracy for MASH-related portal hypertension.

**Application of Baveno VII consensus and its limitation in MASLD**

Numerous validation studies have been conducted over the past two decades to show high
accuracy in liver fibrosis assessment using LSM measured by VCTE (127-130). With the broader
use of LSM as a noninvasive tool to assess liver fibrosis, the term “compensated advanced chronic
liver disease” (cACLD) depicts a spectrum consisting of advanced liver fibrosis to compensated
cirrhosis.(118, 131) Moreover, LSM is shown to correlate with not only the degree of fibrosis
but also the severity of portal hypertension (132, 133). From the ANTICIPATE model which
studied 518 patients with cACLD, the combination of LSM and platelet count had an AUROC of
0.85 in identifying CSPH by HVPG, and that could also predict <5% of varices needing treatment
(VNT) using LSM of 20kPa and platelet count ≥150 x 10⁹/L (134). On this basis, the updated
Baveno VII consensus endorsed the noninvasive modality in the determination of cACLD and
CSPH using LSM and platelet count (118). For instance, LSM ≥10kPa is suggestive of cACLD. LSM
<15kPa with platelet count ≥150 x 10⁹/L rules out CSPH whereas that of ≥25kPa depicts CSPH in
which non-selective beta-blockers should be initiated to prevent hepatic decompensation (118).
Nonetheless, a significant proportion of cACLD patients (around 40%) have LSM 15-24.9kPa and/or platelet count <150 \times 10^9/L, a category known as the gray zone (134). Patients in gray zone still carry a substantial risk of decompensation compared to non-cACLD patients (135, 136).

Despite the adoption of Baveno VII consensus in diagnosing cACLD patients, there is a pitfall in its use in obese patients with MASLD as it is difficult for the shear wave from VCTE to penetrate through the thick subcutaneous and prehepatic fat, causing inaccurate LSM. Studies reported a body mass index (BMI) of 28-30kg/m^2 as a factor associated with failed or unreliable LSM (137). A study with 248 patients with MASH confirmed suboptimal performance of the ANTICIPATE model in detecting CSPH among obese (i.e. BMI \geq 30kg/m^2) patients with MASH across different LSM categories, and that LSM \geq 25kPa only had a PPV of 62.8% in identifying CSPH (134). Thus, the ANTICIPATE model, as well as the Baveno VII consensus, cannot be applied to obese patients with MASH. However, the ANTICIPATE-NASH model, which was developed from the same cohort, considered the phenomenon of lower HVPG in higher BMIs leading to inaccuracy in the original ANTICIPATE model (134). With the addition of BMI to LSM and platelet count, the ANTICIPATE-NASH model serves as a new tool for prediction of CSPH in obese patients with MASH, but further validation is needed.
Combining spleen stiffness measurement into Baveno VII consensus

Alongside LSM, spleen stiffness measurement (SSM) can be obtained by the same VCTE machine, and it appears to be a reliable adjunct to solve the limitation of Baveno VII criteria. Regardless of the liver disease etiology, by combining SSM (40kPa as cutoff) with LSM and platelet count, a significant proportion of patients in the gray zone can be categorized into having CSPH and non-CSPH, reducing the proportion in gray zone to less than 20%, while maintaining a high discrimination in the risk prediction of hepatic decompensation between the CSPH and non-CSPH groups (135, 138). Second, SSM, as a direct surrogate of portal pressure, enhances the performance of VCTE in MASLD and hence aids its prognostication (139). In a study involving 154 patients with MASLD, SSM <40kPa was validated with 100% NPV for high-risk varices and that >40kPa had 100% sensitivity in predicting esophageal varices. SSM ≥21kPa had a high sensitivity and PPV of 96% and 88% for cirrhosis, respectively (140). In another study, SSM was also shown to have good correlation with HVPG and hence was accurate, with AUROC of 0.95, in ruling in (using SSM ≥50kPa) and out (using SSM <40.9kPa) CSPH in patients with MASLD (141). Although there are no unified cutoff values for SSM in MASLD, its validation supports the utility in assisting the diagnosis of portal hypertension in patients with MASLD along with the use of Baveno VII consensus, especially in obese patients. Figure 1 summarizes the role of Baveno VII consensus and SSM in MASLD.

Apart from LSM/SSM-based assessments, the von Willebrand factor antigen-to-platelet ratio (VITRO) appears to have similar prognostic performance as HVPG and the ANTICIPATE±NASH models in patients with cACLD (142). In a study of 420 patients, the incidence of hepatic
decompensation at 1 year was 0% in patients with VITRO <2.5, compared with 10.4% in those with VITRO ≥2.5.

**CONCLUSION AND FUTURE PERSPECTIVES**

Various noninvasive tests, especially those for liver fibrosis, have demonstrated the ability to predict HCC and hepatic decompensation at a level that is non-inferior to that of liver biopsy, thus raising the question of whether liver biopsy should remain the only acceptable surrogate marker for prognostication and drug approval. In recent years, a number of regional and international guidelines have converged towards recommending FIB-4 as the initial assessment in patients with MASLD due to its low cost and wide availability, followed by LSM or other specific fibrosis biomarkers in case FIB-4 falls into the gray zone (i.e., between 1.3 and 2.67) (14, 143, 144). However, the algorithm has much room for improvement. Despite the high NPV, FIB-4 still misses a significant proportion of patients with advanced fibrosis (145). In contrast, at the low cutoff of 1.3, FIB-4 has a rather low PPV, meaning that the majority of patients referred for LSM or hepatology consultation are not expected to have advanced fibrosis. To have a sustainable care model, it is imperative that a second fibrosis test is made available to primary care physicians and colleagues in other specialties (146). Additionally, as colleagues outside the hepatology field need to attend to many other medical conditions, it is difficult for them to prioritize MASLD assessment even in a high-risk population. To achieve this, we have to make the process part of the routine or automatic. For example, many groups have demonstrated the possibility of including fibrosis assessment as part of a comprehensive screening for diabetic complications (147, 148). Some medical centers have also started automatic FIB-4 calculation based on existing
laboratory values, and flagging of abnormal test results in the electronic health record can increase the identification of patients for further evaluation (149).

Case selection is another issue for consideration. Current guidelines recommend MASLD and fibrosis assessment in patients with type 2 diabetes, obesity and other cardiometabolic risk factors (144, 150). Among them, type 2 diabetes has the strongest association with advanced liver disease and liver-related complications (151). While the high pretest probability of advanced liver disease makes it logical to consider fibrosis assessment in this high-risk group, some studies suggest that the performance of noninvasive tests might be worse in patients with type 2 diabetes (152). Further refinement of the diagnostic process is necessary.

Although this article focuses on liver fibrosis, there is increasing interest in diagnosing “at-risk” or “high-risk” MASH, which represents patients with steatohepatitis and F2-F3 fibrosis. This is in keeping with the patient population in most clinical trials in MASH and is indeed the target population for treatment once a drug comes to the market, as exemplified by the recent conditional approval of resmetirom. The relative roles of MASH and fibrosis biomarkers in patient management and clinical care pathways should be determined.

In addition, most of the prognostic data on specific fibrosis biomarkers came from secondary and tertiary centers, while the prognostic performance of simple fibrosis scores such as FIB-4 has been extensively evaluated in the general population or primary care settings due to the possibility of retrospective score calculation based on existing laboratory results. Determining
the performance of specific fibrosis biomarkers in primary care should be a research priority, especially as most patients with MASLD are seen there (146). Adoption is also key. Table 4 summarizes the regional guidelines/guidance on the use of noninvasive tests in MASLD.

In the past 16 years, a number of gene polymorphisms (e.g., PNPLA3, TM6SF2, MBOAT7, HSD17B13) were identified to be associated with MASLD, its histological severity, and future risk of HCC and hepatic decompensation (153). Although polygenic risk scores have been proposed for the prediction of HCC using these genetic markers (154), the added value is marginal over existing clinical prediction models. Nonetheless, new genetic markers are being identified in recent years using better powered cohorts (155). Besides, the identification of epigenetic markers and multiomic approaches will likely refine noninvasive assessments further. Currently, such an approach is limited by low availability and prohibitive costs of the tests. However, the field should keep an open mind as technological advances and innovative approaches may drive down the cost and improve test availability. Notably, molecular tests are already routinely used in oncology practice. With concerted effort, we can produce noninvasive tests and clinical care pathways that can make precision medicine in MASLD a reality.
REFERENCES


73. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64:1388-1402.


113. Venkatesh SK, Talwalkar JA. When and how to use magnetic resonance elastography for patients with liver disease in clinical practice. Am J Gastroenterol 2018;113:923-926.


Reported Outcomes in Patients With Advanced Fibrosis Due to Nonalcoholic Steatohepatitis. Gastroenterology 2021;160:1608-1619 e1613.


Table 1. Prognostication of metabolic dysfunction-associated steatotic liver disease (MASLD) by simple fibrosis scores.

<table>
<thead>
<tr>
<th>Simple fibrosis scores</th>
<th>Common cut-offs (15)</th>
<th>Sensitivity (%) for F3,(15) LRE,(71) MALO(65)</th>
<th>Specificity (%) for F3,(15) LRE,(71) MALO(65)</th>
<th>Diagnostic performance for detection of F2, F3, F4,(15) and prediction of HD, HCC, LRE,(71) and MALO(65) in 5 years AUROC (95% CI)</th>
<th>Risk levels for prognostication</th>
<th>5-year cumulative risk of event (95% CI) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FIB-4</strong></td>
<td>1.30</td>
<td>F3: 77.8 LRE: 88 MALO: 82.6</td>
<td>F3: 71.2 MALO: 54.5</td>
<td>F2: 0.75 (0.70-0.79) F3: 0.80 (0.77-0.84) F4: 0.85 (0.81-0.89) HD: 0.88 (0.85-0.92) HCC: 0.77 (0.71-0.83) LRE: 0.84 (0.81-0.88) MALO: 0.74 (0.64-0.82)</td>
<td>&lt;1.30</td>
<td>0.4 (0.3-0.6)</td>
</tr>
<tr>
<td></td>
<td>2.67</td>
<td>F3: 31.9 LRE: 91 MALO: 41.3</td>
<td>F3: 95.7 LRE: 91 MALO: 87.7</td>
<td></td>
<td>1.30 - 2.67</td>
<td>1.8 (1.4-2.4)</td>
</tr>
<tr>
<td></td>
<td>3.25</td>
<td>F3: 37.3</td>
<td>F3: 95.8</td>
<td></td>
<td>&gt;2.67</td>
<td>12.0 (9.8-14.3)</td>
</tr>
<tr>
<td><strong>NFS</strong></td>
<td>-1.455</td>
<td>F3: 72.9 LRE: 88 MALO: 78.9</td>
<td>F3: 73.8 MALO: 46.5</td>
<td>F2: 0.72 (0.65-0.79) F3: 0.78 (0.75-0.81) F4: 0.83 (0.76-0.89) HD: 0.89 (0.86-0.93) HCC: 0.77 (0.70-0.83) LRE: 0.84 (0.80-0.88) MALO: 0.70 (0.63-0.80)</td>
<td>&lt; -1.455</td>
<td>0.4 (0.2-0.6)</td>
</tr>
<tr>
<td></td>
<td>0.676</td>
<td>F3: 43.1 MALO: 31.6</td>
<td>F3: 88.4 LRE: 94 MALO: 84.6</td>
<td></td>
<td>-1.455 - 0.676</td>
<td>2.5 (1.9-3.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&gt;0.676</td>
<td>12.8 (10.3-15.8)</td>
</tr>
<tr>
<td><strong>APRI</strong></td>
<td>0.5</td>
<td>F3: 72.9 LRE: 75</td>
<td>F3: 67.7</td>
<td>F2: 0.70 (0.64-0.76) F3: 0.75 (0.72-0.77) F4: 0.75 (0.70-0.80) HD: 0.84 (0.79-0.88) HCC: 0.72 (0.65-0.78) LRE: 0.79 (0.75-0.83)</td>
<td>&lt;0.5</td>
<td>0.7 (0.5-0.9)</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>F3: 43.2</td>
<td>F3: 86.1</td>
<td></td>
<td>0.5 - 1.5</td>
<td>4.8 (3.9-5.8)</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>F3: 32.9</td>
<td>F3: 90.5 LRE: 97</td>
<td></td>
<td>&gt;1.5</td>
<td>11.4 (7.6-16.0)</td>
</tr>
<tr>
<td><strong>BARD</strong></td>
<td>2</td>
<td>F3: 75.2 LRE: 86</td>
<td>F3: 61.6 LRE: 37</td>
<td>F2: 0.64 (0.53-0.75) F3: 0.73 (0.71-0.75) F4: 0.70 (0.63-0.77) HD: 0.73 (0.68-0.77) HCC: 0.63 (0.56-0.69) LRE: 0.69 (0.65-0.73)</td>
<td>&lt;2</td>
<td>0.9 (0.5-1.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>≥2</td>
<td>2.7 (2.2-3.1)</td>
</tr>
</tbody>
</table>

* At a median follow-up of 57 (25th - 75th percentile 33-91) months.

LRE was defined as a composite endpoint of HCC, HD, liver transplantation, and liver-related death. MALO was defined as a composite endpoint that included all-cause mortality, liver transplantation, HCC, and cirrhosis decompensation. (variceal bleeding, ascites, hepatic encephalopathy), and increase in MELD score to 15 or higher.

Abbreviations: APRI = aspartate aminotransferase to platelet ratio, AUROC = area under the receiver operating characteristic curve, CI = confidence interval, F2 = significant fibrosis, F3 = advanced fibrosis, F4 = cirrhosis, FIB-4 = fibrosis 4, HCC = hepatocellular carcinoma, HD = hepatic decompensation, LRE = liver-related events, MALO = major adverse liver outcomes, N.A. = not available, NFS = non-alcoholic fatty liver disease fibrosis score.
Table 2. Prognostication of metabolic dysfunction-associated steatotic liver disease (MASLD) by specific fibrosis biomarkers.

<table>
<thead>
<tr>
<th>Simple scores fibrosis</th>
<th>Common cut-offs</th>
<th>Sensitivity [Youden] (%) (95% CI) for F2, F3, LRE</th>
<th>Specificity [Youden] (%) (95% CI) for F2, F3, LRE</th>
<th>Diagnostic performance for detection of F2, F3, F4; and prediction of liver-related outcomes</th>
<th>Cumulative risk of event (95% CI) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUROC (95% CI)</td>
<td>HCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LRE</td>
</tr>
<tr>
<td>ELF(46, 156)</td>
<td>F2: 8.8 [90% sen rule out] 10.0 [90% spe rule in] F3: 7.7 F3: 9.8 F3: 10.51 F3: 11.30 F4: 9.7 [90% sen rule out] 10.9 [90% spe rule in]</td>
<td>F2: 72 F3: 93 (82-98) F3: 65 (49-77) F3: 51 (31-70) F3: 36 (15-63) F4: 82</td>
<td>F2: 82 F3: 34 (13-65) F3: 86 (77-92) F3: 93 (85-96) F3: 96 (90-99) F4: 73</td>
<td>F2: 0.81 (0.66-0.89) F3: 0.83 (0.71-0.90) F4: 0.86 (0.82-0.89)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Median follow-up: 4 years(157, 158) 9.8 – 11.3: 15.0% ≥11.3: 30-50%</td>
</tr>
<tr>
<td>PRO-C3(48, 156) (ng/mL)</td>
<td>F2: 9.7-20.9 [Youden] 12.8 [90% sen rule out] 20.1 [90% spe rule in] F3: 12.7-21.3 [Youden] 13.6 [90% sen rule out] 25.0 [90% spe rule in] F4: 15.1 [90% sen rule out] 30.6 [90% spe rule in]</td>
<td>F2: 68 (50-82) F3: 72 (62-81) F4: 66</td>
<td>F2: 79 (71-86) F3: 73 (65-80) F4: 69</td>
<td>F2: 0.81 (0.77-0.84) F3: 0.79 (0.73-0.82) F4: 0.73 (0.69-0.77)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N.A.</td>
</tr>
<tr>
<td>FibroTest(51, 162)</td>
<td>F2: 0.322-0.48 [Youden] F3: 0.316-0.506 [Youden] F4: 0.57-0.75 [Youden] 0.48 <a href="163">EASL rule out</a> 0.3 [90% sen rule out] 0.7 [90% spe rule in]</td>
<td>F2: 72 (28-94) F3: 70 (15-72) F4: 40-77 [range]</td>
<td>F2: 85 (45-98) F3: 93 (73-99) F4: 90-95 [range]</td>
<td>F2: 0.86 F3: 0.78 F4: 0.92 10y LRD: 0.94 (0.91-0.98)(67)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10-year LRD:(67) &lt;0.74: 28% &gt;0.74: 44%</td>
</tr>
<tr>
<td>FibroMeter(51, 156) (NAFLD, V2G, V3G, VCTE)</td>
<td>F2: 0.31-0.38 [Youden] 0.2 [90% sen rule out] 0.6 [90% spe rule in] F3: 0.311-0.589 [Youden] 0.45 <a href="163">EASL rule out</a> 0.25-0.385 [90% sen rule out] 0.585-0.831 [90% spe rule in] F4: 0.7 [90% sen rule out] 0.9 [90% spe rule in]</td>
<td>F2: 68 (48-82) F3: 74 (68-79) F4: 94</td>
<td>F2: 89 (80-95) F3: 82 (76-87) F4: 70</td>
<td>F2: 0.88 F3: 0.84 F4: 0.90 (0.84-0.95) 10y LRD: 0.84 (0.75-0.93)(52)</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10-year LRD:(52) &lt;0.168: 2% 0.168-0.373: 15% 0.374-0.499: 15% 0.5-0.754: 51% 0.755-0.969: 64% 0.970-1: 80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N.A.</td>
</tr>
</tbody>
</table>

Abbreviations: AUROC = area under the receiver operating characteristic curve, CI = confidence interval, COI = cut-off index, ELF = enhanced liver fibrosis, F2 = significant fibrosis, F3 = advanced fibrosis, F4 = cirrhosis, HCC = hepatocellular carcinoma, LRD = liver related death, LRE = liver-related events, N.A. = not available, NAFLD = non-alcoholic fatty liver disease, VCTE = vibration-controlled transient elastography.
Table 3. Prognostication of metabolic dysfunction-associated steatotic liver disease (MASLD) by imaging biomarkers.

<table>
<thead>
<tr>
<th>Simple fibrosis scores</th>
<th>Common cut-offs</th>
<th>Sensitivity (%) (95% CI) for F2, fibrotic MASH, F3, F4, MALO</th>
<th>Specificity (%) (95% CI) for F2, fibrotic MASH, F3, F4, MALO</th>
<th>Diagnostic performance for detection of F2, fibrotic MASH, MASH, F3, and F4; and prediction of HD, HCC, LRE, and MALO in 5 years(172)</th>
<th>5-year cumulative risk of event (95% CI) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LSM by VCTE(65, 84) (kPa)</strong></td>
<td>F2: 3.8-10.2 [range] F3: 6.8-12.9 [range] F4: 6.9-19.4 [range] MALO: 10</td>
<td>F2: 80 (76-83) F3: 80 (77-83) F4: 76 (70-82) MALO: 71 (62-79) MALO: 29 (19-40)</td>
<td>F2: 73 (68-77) F3: 77 (74-80) F4: 88 (85-91) MALO: 66 (64-69) MALO: 92 (90-93)</td>
<td>F2: 0.83 (0.80-0.87) F3: 0.85 (0.83-0.87) F4: 0.89 (0.84-0.93) HCC: 0.76 (0.69-0.82) HD: 0.90 (0.88-0.94) LRE: 0.85 (0.82-0.88) MALO: 0.76 (0.70-0.83)</td>
<td>Median FU: 27 months [IQR 25-38] &lt;8: 0.4 (0.3-0.6) 8-12: 1.2 (0.8-1.8) &gt;12: 10.4 (8.9-12.1)</td>
</tr>
<tr>
<td><strong>Agile 3+</strong></td>
<td>F3: 0.451 [rule out] F3: 0.679 [rule in]</td>
<td>F3 [rule out]: 88 (81-93) F3 [rule in]: 68 (57-78)</td>
<td>F3 [rule out]: 65 (54-75) F3 [rule in]: 87 (80-92)</td>
<td>F3:(173) 0.86 (0.82-0.89) HCC: 0.80 (0.73-0.87) HD: 0.94 (0.91-0.96) LRE: 0.80 (0.85-0.92)</td>
<td>Median FU: 2.7 years [IQR 0-12.5] Annual incidence(174) &lt;0.679: 0 ≥0.679: 1.61</td>
</tr>
<tr>
<td><strong>Agile 4</strong></td>
<td>F4: 0.251 [rule out] F4: 0.565 [rule in]</td>
<td>F4 [rule out]: 71-88 [range]</td>
<td>F4 [rule out]: 71-88 [range] F4 [rule in]: 93-97 [range]</td>
<td>F4:(79, 175, 176) 0.85-0.93 [range] HCC: 0.80 (0.73-0.86) HD: 0.93 (0.90-0.96) LRE: 0.88 (0.85-0.91)</td>
<td>Median FU: 2.7 years [IQR 0-12.5] Annual incidence(174) &lt;0.565: 0.08 ≥0.565: 4.04</td>
</tr>
<tr>
<td><strong>FAST</strong></td>
<td>Fibrotic MASH: 0.35 [rule out] Fibrotic MASH: 0.67 [rule in]</td>
<td>F2 [rule out]: 77</td>
<td>F2 [rule out]: 61 Fibrotic MASH [rule out]: 80 F2 [rule in]: 94 Fibrotic MASH [rule in]: 88</td>
<td>F2:(177) 0.73 (0.68-0.77) HCC: 0.74 (0.67-0.81) HD: 0.81 (0.75-0.86) LRE: 0.78 (0.74-0.83)</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>pSWE (m/s)(84)</strong></td>
<td>F2: 1.18-1.81 [range] F3: 1.34-4.24 [range] F4: 1.36-2.54 [range]</td>
<td>F2: 69 (59-77) F3: 80 (70-88) F4: 76 (59-87)</td>
<td>F2: 85 (80-88) F3: 86 (82-92) F4: 88 (82-92)</td>
<td>F2: 0.86 (0.78-0.90) F3: 0.89 (0.83-0.95) F4: 0.90 (0.82-0.95) HCC:(179) 0.94</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>2D-SWE (kPa)(84)</strong></td>
<td>F2: 8.3-11.6 [range] F3: 9.3-13.1 [range] F4: 14.4-15.7 [range]</td>
<td>F2: 71 (56-83) F3: 72 (65-78) F4: 78 (50-93)</td>
<td>F2: 67 (43-84) F3: 72 (52-86) F4: 84 (74-90)</td>
<td>F2: 0.75 (0.58-0.87) F3: 0.72 (0.60-0.84) F4: 0.88 (0.81-0.91)</td>
<td>N.A.</td>
</tr>
<tr>
<td><strong>MRE (kPa)(84)</strong></td>
<td>F2: 2.86-4.14 [range] F3: 3.29-4.80 [range] F4: 3.35-6.70 [range] MASH: 2.53-3.26 [range]</td>
<td>F2: 78 (67-85) F3: 83 (77-88) F4: 81 (66-90) MASH: 65 (46-80)</td>
<td>F2: 89 (83-94) F3: 89 (86-92) F4: 90 (85-94) MASH: 83 (69-91)</td>
<td>F2: 0.91 (0.80-0.97) F3: 0.92 (0.88-0.95) F4: 0.90 (0.81-0.95) MASH: 0.83 (0.69-0.91)</td>
<td>3-year risk(100) &lt;5: 0.35% 5-8: 2.5% ≥8: 5.66% 3-year death(100) &lt;5: 1.64% 5-8: 16.87% ≥8: 19.14% 3-year risk(100) &lt;5: 4.5% 5-8: 10.18% ≥8: 20.19%</td>
</tr>
<tr>
<td>ct1(180) (ms)</td>
<td>MRI-PDFF (%) (105, 180)</td>
<td>MEFIB(177, 181)</td>
<td>MAST(112, 177, 182, 183)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrotic MASH: 800 825 875 900 925</td>
<td>MASH: 5 [Youden] G1: 6.4-8.9 G2: 16.3-17.4 G3: 22.1-25</td>
<td>Positive [rule in]: MRE ≥3.3 kPa and FIB-4 ≥1.6 Negative [rule out]: MRE &lt;3.3 kPa or FIB-4 &lt;1.6</td>
<td>Fibrotic MASH: 0.165 0.242</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrotic MASH: 86 78 59 48 39</td>
<td>MASH: 92</td>
<td>F2 [rule out]: 94 Fibrotic MASH [rule out]: 93 F2 [rule in]: 69 Fibrotic MASH [rule in]: 64</td>
<td>F2 [rule out]: 62 Fibrotic MASH [rule out]: 65 F2 [rule in]: 47 Fibrotic MASH [rule in]: 50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibrotic MASH: 56 67 81 86 90</td>
<td>MASH: 40</td>
<td>F2 [rule out]: 73 Fibrotic MASH [rule out]: 43 F2 [rule in]: 94 Fibrotic MASH [rule in]: 63</td>
<td>F2 [rule out]: 93 Fibrotic MASH [rule out]: 79 F2 [rule in]: 95 Fibrotic MASH [rule in]: 84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MASH: 0.78 (0.74-0.82) Fibrotic MASH: 0.73 (0.68-0.78)</td>
<td>G1: 0.989 G2: 0.825-0.950 [range] G3: 0.893 MASH: 0.78 (0.74-0.82) Fibrotic MASH: 0.69 (0.63-0.75)</td>
<td>F2: 0.90 (0.86-0.94) Fibrotic MASH: 0.69 (0.63-0.75)</td>
<td>F2: 0.77 (0.73-0.81) Fibrotic MASH: 0.72-0.93 [range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*At a median follow-up of 57 (25th - 75th percentile 33-91) months.*

LRE was defined as a composite endpoint of HCC, HD, liver transplantation, and liver-related death. MALO was defined as a composite endpoint that included all-cause mortality, liver transplantation, HCC, and cirrhosis decompensation, (variceal bleeding, ascites, hepatic encephalopathy), and increase in MELD score to 15 or higher.

Abbreviations: AST = aspartate aminotransferase, AUROC = area under the receiver operating characteristic curve, CI = confidence interval, ct1 = corrected T1, FAST = FibroScan-AST, FU = follow-up, IQR=interquartile range, HCC = hepatocellular carcinoma, LRE = liver-related events, LSM = liver stiffness measurement, MALO = major adverse liver outcomes, MAST = MRI-AST, MEFIB = MRE plus FIB-4, MRE = magnetic resonance elastography, MRI=magnetic resonance imaging, MRI-PDFF = MRI proton density fat fraction, N.A. = not available, SWE = shear wave elastography, VCTE = vibration-controlled transient elastography.
Table 4. Comparison of selected regional guidelines/guidance on the use of noninvasive tests in MASLD

<table>
<thead>
<tr>
<th>First-line test</th>
<th>AASLD (144)</th>
<th>AGA (14)</th>
<th>EASL-EASD-EASO (150)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interpretation of first-line test</td>
<td>FIB-4 &lt;1.3 = low risk 1.3-2.67 = second-line test &gt;2.67 = referral to hepatologist</td>
<td>FIB-4 &lt;1.3 = low risk 1.3-2.67 = second-line test &gt;2.67 = referral to hepatologist</td>
<td>FIB-4 &lt;1.3 = low risk 1.3-2.67 = second-line test or reassess after intensified management of comorbidities &gt;2.67 = referral to hepatologist</td>
</tr>
<tr>
<td>Second-line test</td>
<td>LSM, ELF or alternative test</td>
<td>LSM or alternative test</td>
<td>LSM or alternative test</td>
</tr>
<tr>
<td>Interpretation of second-line test</td>
<td>LSM &lt;8 kPa = low risk 8-12 kPa = intermediate risk and referral to hepatologist &gt;12 kPa = high risk and referral to hepatologist</td>
<td>ELF &lt;7.7 = low risk 7.7-9.8 = intermediate risk and referral to hepatologist &gt;9.8 = high risk and referral to hepatologist</td>
<td>ELF &lt;8 kPa = low risk 8-12 kPa = indeterminate risk &gt;12 kPa = high risk</td>
</tr>
</tbody>
</table>

AASLD = American Association for the Study of Liver Diseases, AGA = American Gastroenterological Association, APASL = Asian Pacific Association for the Study of the Liver, EASL-EASD-EASO = European Association for the Study of the Liver, European Association for the Study of Diabetes, and European Association for the Study of Obesity, ELF = Enhanced Liver Fibrosis score, FIB-4 = Fibrosis-4 index, LSM = liver stiffness measurement by vibration-controlled transient elastography
FIGURE LEGENDS

Figure 1. The role of Baveno VII consensus and spleen stiffness measurement (SSM) in metabolic-dysfunction associated steatotic liver disease (MASLD). cACLD, compensated advanced chronic liver disease; CSPH, clinically significant portal hypertension; MASLD, metabolic-dysfunction associated steatotic liver disease; SSM, spleen stiffness measurement.

* = Liver stiffness measurement by transient elastography