Gastroenterology 2023;165:1080–1088

CLINICAL PRACTICE UPDATE

AGA Clinical Practice Update on the Role of Noninvasive Biomarkers in the Evaluation and Management of Nonalcoholic Fatty Liver Disease: Expert Review

Julia J. Wattacheril,[1,2,*] **Manal F. Abdelmalek,**[3,*] Joseph K. Lim,[4] and Arun J. Sanyal[5]

[1]*Division of Digestive and Liver Diseases, Columbia University Vagelos College of Physicians and Surgeons, New York, New York;* [2]*Center for Liver Disease and Transplantation, Columbia University Irving Medical Center, New York Presbyterian Hospital, New York, New York;* [3]*Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota;* [4]*Section of Digestive Diseases, Yale Liver Center, Yale University School of Medicine, New Haven, Connecticut; and* [5]*Division of Gastroenterology and Hepatology, Virginia Commonwealth University School of Medicine, Richmond, Virginia*

DESCRIPTION: The purpose of this American Gastroenterological Association (AGA) Clinical Practice Update Expert Review is to provide clinicians with guidance on the use of noninvasive tests (NITs) in the evaluation and management of patients with nonalcoholic fatty liver disease (NAFLD). NAFLD affects nearly 30% of the global population and is a growing cause of end-stage liver disease and liver-related health care resource utilization. However, only a minority of all patients with NAFLD experience a liver-related outcome. It is therefore critically important for clinicians to assess prognosis and identify those with increased risk of disease progression and negative clinical outcomes at the time of initial assessment. It is equally important to assess disease trajectory over time, particularly in response to currently available therapeutic approaches. The reference standard for assessment of prognosis and disease monitoring is histologic examination of liver biopsy specimens. There are, however, many limitations of liver biopsies and their reading that have limited their use in routine practice. The utilization of NITs facilitates risk stratification of patients and longitudinal assessment of disease progression for patients with NAFLD. This clinical update provides best practice advice based on a review of the literature on the utilization of NITs in the management of NAFLD for clinicians. Accordingly, a combination of available evidence and consensus-based expert opinion, without formal rating of the strength and quality of the evidence, was used to develop these best practice advice statements. **METHODS:** This Expert Review was commissioned and approved by the AGA Institute Clinical Practice Updates Committee and the AGA Governing Board to provide timely guidance on a topic of high clinical importance to the AGA membership and underwent internal peer review by the Clinical Practice Updates Committee and external peer review through standard procedures of *Gastroenterology*. These best practice advice statements were drawn from a review of the published literature and from expert opinion. Because systematic reviews were not performed, these best practice advice statements do not carry formal ratings of the quality of evidence or strength of the presented considerations.

BEST PRACTICE ADVICE STATEMENTS

BEST PRACTICE ADVICE 1: NITs can be used for risk stratification in the diagnostic evaluation of patients with NAFLD. **BEST PRACTICE ADVICE 2:** A Fibrosis 4 Index score <1.3 is associated with strong negative predictive value for advanced hepatic fibrosis and may be useful for exclusion of advanced hepatic fibrosis in patients with NAFLD. **BEST PRACTICE ADVICE 3:** A combination of 2 or more NITs combining serum biomarkers and/or imaging-based biomarkers is preferred for staging and risk stratification of patients with NAFLD whose Fibrosis 4 Index score is >1.3. **BEST PRACTICE ADVICE 4:** Use of NITs in accordance with manufacturer's specifications (eg, not in patients with ascites or pacemakers) can minimize risk of discordant results and adverse events. **BEST PRACTICE ADVICE 5:** NITs should be interpreted with context and consideration of pertinent clinical data (eg, physical examination, biochemical, radiographic, and endoscopic) to optimize positive predictive value in the identification of patients with advanced fibrosis. **BEST PRACTICE ADVICE 6:** Liver biopsy should be considered for patients with NIT results that are indeterminate or discordant; conflict with other clinical, laboratory, or radiologic findings; or when alternative etiologies for liver disease are suspected. **BEST PRACTICE ADVICE 7:** Serial longitudinal monitoring using NITs for assessment of disease progression or regression may inform clinical management (ie, response to lifestyle modification or therapeutic intervention). **BEST PRACTICE ADVICE 8:** Patients with NAFLD and NITs results suggestive of advanced fibrosis (F3) or cirrhosis (F4) should be considered for surveillance of liver complications (eg, hepatocellular carcinoma screening and variceal screening per Baveno criteria). Patients with NAFLD and NITs suggestive of advanced hepatic fibrosis (F3) or (F4), should be monitored with serial liver stiffness measurement; vibration controlled transient elastography; or magnetic resonance elastography, given its correlation with clinically significant portal hypertension and clinical decompensation.

Keywords: Risk Stratification; Nonalcoholic Steatohepatitis; Fibrosis; Cirrhosis.

Nonalcoholic fatty liver disease (NAFLD) is an emerging global public health crisis. It affects approximately 30% of the worldwide population.[1] As a

*Authors share co-first authorship.

Abbreviations used in this paper: **AF, advanced fibrosis; AGA, America Gastroenterological Association; APRI, aspartate aminotransferase to platelet ratio index; BPA, best practice advice; CSPH, clinically significant portal hypertension; ELF, enhanced liver fibrosis; FAST, FibroScan-aspartate aminotransferase; FIB-4, Fibrosis 4 Index; HCC, hepatocellular carcinoma; LSM, liver stiffness measurement; MEFIB, magnetic resonance elastography with Fibrosis 4 score; NAFLD, nonalcoholic fatty liver disease; NAFLD-FS, nonalcoholic fatty liver disease fibrosis score; NASH, nonalcoholic steatohepatitis; NIT, noninvasive test; NPV, negative predictive value; PPV, positive predictive value; SWE, shear wave elastography; VCTE, vibration controlled transient elastography.**

Most current article

© 2023 The Author(s). Published by Elsevier Inc. on behalf of the AGA Institute. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

0016-5085
https://doi.org/10.1053/j.gastro.2023.06.013

metabolic disease representing the hepatic manifestation of a systemic metabolic disorder, NAFLD is associated with significant morbidity and mortality, as well as substantial health care resource utilization.[2,3] The traditional approach to defining disease severity in patients with NAFLD has been to perform a liver biopsy for histologic grading of necroinflammation and staging of hepatic fibrosis—2 key features of disease severity. *Disease activity* refers to the biological processes leading to hepatic injury and inflammation, whereas *fibrosis stage* refers to the amount of scarring and thus proximity to cirrhosis. Fibrosis stage is the strongest predictor of future outcomes and thus provides the greatest prognostic information. Natural history studies have found that assessment of fibrosis, specifically cirrhosis, serves as a meaningful surrogate for liver-related outcomes, such as hepatocellular carcinoma (HCC), liver decompensation (eg, variceal hemorrhage, ascites, and hepatic encephalopathy), liver transplantation, and death.[3,4]

Liver biopsies, however, are invasive; variable in sampling[5]; subject to intra- and interobserver variability; and, rarely, may be associated with severe and/or fatal procedural complications. It is thus impractical to use biopsy-based risk stratification for a prevalent and chronic disease such as NAFLD. Noninvasive tests (NITs) have emerged as validated tools to address the problem of early risk stratification in NAFLD. Guidance for clinicians on the use of NITs in the care of patients with NAFLD is limited. We reviewed the use of NITs as a noninvasive surrogate approach to risk stratification of advanced fibrosis (AF) in the longitudinal care of patients with NAFLD to predict liver-related outcomes and guide responses to therapies.

NITs can be subcategorized into serum-based and imaging-based biomarkers. Multiple noninvasive models using biochemical serum and clinical measurements have been proposed to detect liver fibrosis. The focus of this review will be on readily available, point-of-care, cost-effective testing strategies to risk stratify patients with NAFLD for AF.

This review was designed to provide best practice advice (BPA) and guidance on several key clinical issues pertaining to NAFLD management using NITs. We have developed BPA statements to address 8 key clinical issues. This Expert Review was commissioned and approved by the American Gastroenterological Association (AGA) Institute Clinical Practice Updates Committee and the AGA Governing Board to provide timely guidance on a topic of high clinical importance to the AGA membership and underwent internal peer review by the Clinical Practice Updates Committee and external peer review through standard procedures of the journal *Gastroenterology*. These BPA statements represent the current literature in studied populations. However, risk stratification should be undertaken only upon confirmation of diagnosis of NAFLD, exclusion of competing diagnoses, and evaluation of presenting signs and symptoms in accordance with good clinical practice. Furthermore, the final approach used for a given patient must take into consideration the unique clinical context for that patient and having discussed the "pros and cons" of alternative approaches.

Best Practice Advice 1: NITs can be used for risk stratification in the diagnostic evaluation of patients with NAFLD.

Among patients with NAFLD, those with nonalcoholic steatohepatitis (NASH) are at risk of fibrosis progression. Most experts believe that patients with NAFLD should be monitored for development of chronic liver injury (ie, NASH and hepatic fibrosis). Within the context of a gastroenterology practice, surveillance strategies for monitoring liver-related outcomes are implemented upon the knowledge of the presence of advanced hepatic fibrosis or cirrhosis. A patient's access and engagement with the health care system, ability to follow lifestyle recommendations, and effective extrahepatic disease management may also improve clinical outcomes in patients with NAFLD.

A key predictor for liver-related outcomes is AF, defined as presence of stage 3 (bridging fibrosis) or 4 (cirrhosis) on liver biopsy.[6] Because fibrosis stage is not synchronized across the liver, stages 3–4 represent a continuum of fibrosis. Initial NITs were developed and applied to patients with viral hepatitis,[7] but have subsequently been validated in other chronic liver diseases, specifically NAFLD.[8] Several have been developed (eg, Fibrosis 4 Index [FIB-4] score, NAFLD fibrosis score [NAFLD-FS], and aspartate aminotransferase to platelet ratio index [APRI]; however, FIB-4 score is the most validated. FIB-4 score is calculated using a simple algorithm based on age, alanine aminotransferase, aspartate aminotransferase, and platelet count[7] and outperforms other calculations in its ability to identify patients with a low probability of AF. High values of FIB-4 and other NITs have also been associated with all-cause and liver-related outcomes in population-based studies.[9] Although FIB-4 score does not outperform other proprietary fibrosis biomarkers (eg, FibroTest/FibroSure [eviCore Healthcare], FIBROSpect NASH [Prometheus Laboratories], Hepamet Fibrosis Score, a Pro-C3 based score [ADAPT], FibroMeter [ARUP Laboratories], and Hepascore), FIB-4 is recommended as a first-line assessment for practitioners based on its simplicity and low cost.[10–12] The Enhanced Liver Fibrosis (ELF; Siemens Healthineers USA) test, a proprietary blood test consisting of 3 elements involved in matrix turnover, and NIS2+™ (property of GENFIT; Loos, France), as an optimization of NIS4® (property of GENFIT; Loos, France) technology, are validated blood-based serum biomarkers for detection of advanced hepatic fibrosis and "at-risk" NASH with F2 or higher, respectively. An ELF score $\geq$9.8 reliably identifies patients with NAFLD at increased risk of progression to cirrhosis and liver-related clinical events.[13,14] Such serum-based fibrosis tests may be good options as secondary risk assessments when elastography is not available.

Imaging-based biomarkers, such as vibration controlled transient elastography (VCTE), shear wave elastography (SWE), and magnetic resonance elastography (MRE), are used most frequently and have been validated. Ultrasound-based 3-dimensional elastography (Velacur) and iron-

corrected T1 magnetic resonance imaging, although used less frequently, are emerging technologies.

Currently, there is no minimum cutoff established for diagnostic accuracy for AF using NITs. Using liver histology as the reference standard, a meta-analysis of 10 studies evaluated the performance of serum based and imaging NITs in NAFLD.[15] The performance of FIB-4 for absence of AF with a cutoff value in the 1.24–1.45 range in 2759 patients demonstrated a mean sensitivity of 77.8% (range, 63.0%–90.0%), specificity of 71.2% (range, 55.5%–88.0%), positive predictive value (PPV) of 40.3% (range, 24.0%–50.6%), and negative predictive value (NPV) of 92.7% (range, 88.0%–98.0%). An FIB-4 score threshold of <1.3 was otherwise adapted from several studies as being an ideal threshold for clinical practice. NAFLD-FS performance for AF using a cutoff of −1.455 in 10 studies with 3057 patients had a sensitivity of 72.9% (range, 22.7%–96.0%), specificity of 73.8% (range, 42.9%–100%), PPV of 50.4% (range, 24.0%–100%), and NPV of 91.8% (range, 81.3%–98.1%).[15] An APRI cutoff of 1.00 for AF in 1101 patients had a sensitivity of 43.2% (range, 27.0%–67.0%), specificity of 86.1% (range, 81.0%–89.0%), PPV of 33.5% (range, 26.0%–40.0%), and NPV of 89.8% (range, 84.0%–95.0%). Although there are several scores (eg, FIB-4, NAFLD-FS, and APRI), FIB-4 is the most validated and outperforms other calculations in its ability to identify patients with a low probability of AF.

Performance of serum-based NITs in cirrhosis improved in NPV. Using FIB-4 score, a cutoff of 1.92–2.48 in 439 patients had a sensitivity of 76.4% (range, 72.7%–80.0%), specificity of 82.4% (range, 76.0%–88.7%), PPV of 39.0% (range, 37.5%–40.4%), and NPV of 96.2% (range, 95.5%–96.9%). NAFLD-FS performance for a cutoff of −0.014 in 197 patients was sensitivity of 80%, specificity of 80.8%, PPV of 42.8%, and NPV of 95.7%. For a cutoff of 0.54–2.00 for APRI in 790 patients, sensitivity was 56.2% (range, 20.5%–77.3%), specificity was 83.6% (range, 56.3%–100%), PPV was 37.8% (range, 16.7%–100%), and NPV was 91.7% (range, 83.0%–96.7%). Despite some variability in the diagnostic profile of these serum-based tests, the consistent concordance in NPV for AF provides reassurance for their utilization in risk stratification[16] (Table 1).

Evaluation of VCTE M-probe performance for AF within the same meta-analysis in 1540 patients from 9 studies using a cutoff of 7.6–8 kPa had a sensitivity of 88.9% (range, 65.0%–100.0%), specificity of 77.2% (range, 65.9%–90.2%), PPV of 43.4% (range, 27.0%–52.0%), and NPV of 95.5% (range, 86.0%–100%). VCTE XL-probe performance at a cutoff of 5.7–9.3 from 3 studies of 579 patients had a sensitivity of 75.3% (range, 57.0%–91.0%), specificity of 74.0% (range, 54.0%–90.0%), PPV of 58.7% (range, 45.0%–71.0%), and NPV of 88.7% (range, 84.0%–93.0%). In cirrhosis, test performance similarly improved; a VCTE cutoff of 10.3–11.3 in 1362 patients had a sensitivity of 87.7% (range, 78.0%–100%), specificity of 86.3% (range, 82.0%–90.0%), PPV of 46.8% (range, 33.0%–75.0%), and NPV of 98.0% (range, 94.0%–100%). For the XL-probe, a broader cutoff of 7.2–16 in 654 patients was associated with a sensitivity of 87.8% (71.0%–100%), specificity of 82.0% (70.0%–91.0%), PPV of 39.8% (range, 31.0%–53.0%), and NPV of 97.8% (range, 95.0%–100%).

Table 1. Noninvasive Tests and Accuracy for Advanced Fibrosis (F3–4) in Nonalcoholic Fatty Liver Disease

Noninvasive test	Recommended cutoff to rule in advanced hepatic fibrosis	AF (F3–4) by biopsy, AUROC (95% CI)
Serum		
FIB-4 score	>2.67	0.83 (0.79–0.86)
NAFLD-FS	>0.676	0.75 (0.71–0.79)
APRI	>0.84	0.76 (0.74–0.79)
ELF	>9.8	0.81 (0.77–0.85)
Imaging		
VCTE, *kPa*	>12.0	0.93 (0.89–0.96)
SWE, *kPa*	>8.0	0.89 (0.80–0.98)
MRE, *kPa*	>3.6	0.93 (0.90–0.96)

AUROC, area under receiver operating curve.

SWE and MRE also have excellent diagnostic performance in the identification of AF. For SWE detection of AF, cutoffs of 3.02–10.6 among 429 patients had a sensitivity of 89.9% (range, 88.2%–91.5%), specificity of 91.8% (range, 90.0%–94.0%), PPV of 88.2% (range, 83.3%–93.1%), and NPV of 93.4% (range, 92.6%–94.2%). Using a cutoff of 3.36 for the detection of cirrhosis in 181 patients, the sensitivity was 100%, specificity was 85.6%, PPV was 55.2%, and NPV was 100%. MRE detection of AF with a cutoff of 3.62–4.8 in 628 patients had a sensitivity of 85.7% (range, 74.5%–92.2%), specificity of 90.8% (range, 86.9%–93.3%), PPV of 71.0% (range, 67.9%–74.5%), and NPV 93.4% (range, 81.0%–98.1%). A liver stiffness measurement (LSM) cutoff of 4.15–6.7 for cirrhosis in 384 patients had a sensitivity of 86.6% (range, 80.0%–90.9%), specificity of 93.4% (range, 91.4%–94.5%), PPV of 53.4% (range, 44.4%–58.8%), and NPV of 98.8% (range, 98.1%–99.2%).[15] Although MRE is not recommended as a first-line approach to risk stratification in a patient with NAFLD,[10–12] it can be an important tool if clinical uncertainty exists, if there is a need for concomitant cross-sectional imaging, or when other elastography techniques are unavailable.

Best Practice Advice 2: An FIB-4 score <1.3 is associated with strong NPV for advanced hepatic fibrosis and may be useful for exclusion of advanced hepatic fibrosis in patients with NAFLD.

The goal of testing algorithms is to establish or exclude a diagnosis with confidence. Indeterminate results require additional testing. For NITs, diagnostic performance is reported using area under receiver operating curve, sensitivity, and specificity. The clinical utility of guidance cutoffs is dependent on the prevalence of NAFLD in the target population. Primary care populations, endocrinology clinics, and gastroenterology practices are likely to differ in the prevalence of AF. Shah et al[17] compared FIB-4 score with 7 other serum-based NITs in a national database of predominantly Caucasian subjects with histologically confirmed NAFLD. Statistically significant differences between groups

included female gender and nondiabetic status in the earlier stage (F1–2) vs AF (F3–4) cohorts. The sensitivity and specificity for AF using a low cutoff of <1.30 in this population were 74% and 71%, respectively. More importantly, the NPV was 90% in the studied population. When applied predictively to a variable prevalence of AF, the NPV increased to 96% with a prevalence of 10% and dropped to 73% with 50% prevalence. In the comparison of NITs that follows, it is important to remember than the NPV and PPV reported are functions of disease prevalence, hence, the context of use within specialty clinics vs the general population is an important distinction. One notable limitation for FIB-4 score is the proportion of scores in the indeterminate range in some populations.[18] Despite this, given the general population prevalence of NAFLD at 30%, it is reasonable to choose a cutoff of <1.3 to augment other clinical measurements to exclude most individuals with AF.f

Furthermore, FIB-4 score is influenced by age and therefore performs poorly in patients younger than 35 years or older than 65 years. When used alone, FIB-4 score has poor diagnostic accuracy in younger patients and low specificity in older adults.[19,20]

Best Practice Advice 3: A combination of 2 or more NITs, combining serum biomarkers and/or imaging-based biomarkers, is preferred over a solitary NIT for staging and risk stratification of patients with NAFLD whose FIB-4 score is >1.3.

Given the heterogeneity of NAFLD in diverse populations, overreliance on 1 NIT may reduce sensitivity in identifying individuals with AF and cirrhosis in both primary care and specialist contexts. When imaging biomarkers are not readily available in clinical practice, a second serum biomarker, such as ELF, may be considered.[12] A primary care referral pathway in the United Kingdom using FIB-4 score <1.30 as a determination for AF with comprehensive review revealed the benefit of sequential testing.[21] Of 1452 patients over 2 years, 1022 (71.3%) had an FIB-4 score <1.3, 43 (3.0%) had an FIB-4 score >3.25, and 387 had an indeterminate FIB-4 score and proceeded to another biomarker (ELF) for further risk stratification, of whom 155 (40%) had an ELF < 9.5 (low risk of AF). These indeterminate findings were confirmed by means of liver biopsy, VCTE, or radiologic features of cirrhosis.[21] Subsequent hepatologist review revealed detection of 29.6% of AF and 14.5% cirrhosis with the risk-stratification pathway (compared with 7.7% and 3.6%, respectively, before pathway utilization), thereby supporting a stepwise risk-stratification approach.

Within the context of clinical trials, sequential testing of NITs reduces indeterminate risk stratification for AF by 20% over single NITs alone,[22] improves the PPV in the detection of moderate fibrosis,[23,24] and optimizes the identification of cirrhosis across prevalence groups,[25] and thereby may reduce the need for liver biopsy. A recent individual patient data meta-analysis of 37 primary studies (n = 5735; 30% AF) evaluating VCTE against liver histology demonstrated that the sequential combination of NITs with a lower cutoff (FIB-4 score <1.3 and VCTE <8.0) to rule out AF and a higher cutoff (FIB-4 score ≥3.48 and VCTE ≥20 kPa) to rule in cirrhosis can reduce the need for liver biopsies from 33% to 19%.[26]

Patients with NAFLD with significant hepatic fibrosis (stage ≥2) are at increased risk of liver-related morbidity and are candidates for pharmacologic therapies. The diagnostic accuracy of the combination of LSM by VCTE or MRE with another biomarker such as FIB-4 score or aspartate aminotransferase may further optimize diagnostic accuracy for detection of the "at-risk" patient with NASH. FibroScan-asparate aminotransferase (FAST)[23] and MRE combined with FIB-4 index (MEFIB)[27] have been developed to optimize predictive values for assessment of NASH with fibrosis stage ≥2. In a prospective study of patients with biopsy-proven NAFLD (n = 563) undergoing contemporaneous MRE, magnetic resonance imaging proton density fat fraction and VCTE from 2 prospective cohorts, MEFIB outperformed FAST (both, $P < .001$); areas under the curve for MEFIB and FAST were 0.901 (95% CI, 0.875–0.928) and 0.725 (95% CI, 0.683–0.767), respectively.[28] The PPV of MEFIB (95.3%) was significantly higher than that of FAST (83.5%; $P = .001$).[28] The NPV of MEFIB (90.1%) was significantly higher than that of FAST (71.8%) ($P < .001$). Furthermore, to diagnose at-risk NASH, defined as NAFLD activity score ≥4 and fibrosis stage ≥2, MEFIB outperformed FAST ($P < .05$); areas under the curve for MEFIB and FAST were 0.768 (95% CI, 0.728–0.808) and 0.687 (95% CI, 0.640–0.733), respectively.[28] Although these combination NITs are of great interest for identifying the at-risk patients with NASH who may be eligible for pharmacologic therapies in clinical trials, pending a US Food and Drug Administration–approved therapy for the treatment of NASH, their role in clinical practice is not yet defined.

Best Practice Advice 4: Use of NITs in accordance with manufacturer's specifications (eg, not in patients with ascites or pacemakers) can minimize risk of discordant results and adverse events.

VCTE is rapid, safe, and has been validated as a reproducible measure of LSM. In addition to adhering to the manufacturer's recommendations regarding probe calibration and selection (M-prove vs XL-probe), quality measures require at least 10 validated measurements with an interquartile range <30% of the median value. Validity also depends on the success rate (the ratio of successful measurements to total number of acquisitions), which should be ≥60%.[29] VCTE performance can vary based on operator experience, high body mass index, ethnic subgroups, nonfasting state, and significantly elevated alanine aminotransferase.[30,31] Due to false-positive rates of VCTE when performed within 3 hours of oral intake, an abnormal LSM should undergo repeat VCTE measures during a fasting state in order to validate an elevated LSM.[32,33] Furthermore, validating elevated LSM on repeat VCTE markedly increased sensitivity of AF among patients with NAFLD.[34,35]

Given the limitations of VCTE, MRE has been studied in patients with severe and morbid obesity. In this population, MRE performance may be limited by waist circumference,[30] therefore, cautious interpretation of LSM in patients with morbid obesity and increased skin-to-capsule distance is advisable.[36] Although MRE is limited by cost, availability,

and time to perform examinations, it does image the entire liver and captures histologic severity.[37] Cost-effectiveness analyses reveal that a combination of FIB-4 score and MRE has a marginally greater cost than FIB-4 score and VCTE per correct diagnosis.[38]

Best Practice Advice 5: NITs should be interpreted with context and consideration of pertinent clinical data (eg, physical examination, biochemical, radiographic, and endoscopic) to optimize PPV in the identification of patients with AF.

Once a patient has been risk stratified as at risk for AF using single or combination NITs, thorough review of their clinical information can improve diagnostic accuracy. Studies examining prospective referral pathways using initial NIT stratification followed by thorough specialist case review for the identification of clinical findings may improve PPV in the identification of AF and cirrhosis.[21] Although FIB-4 score <1.3 has a high NPV for AF, patients initially excluded for AF may be reassessed periodically for evidence of interval progression. Available data are inadequate to establish an optimal frequency for follow-up, although, based on rates of disease progression noted in retrospective and prospective series,[39,40] it is reasonable to repeat an FIB-4 annually in high-risk populations (eg, type 2 diabetes) and every 2 years in low-risk populations (young individuals with few risk factors). In patients with NAFLD, FIB-4 score >2.67 and FIB-4 score >3.25 may be associated with AF and/or cirrhosis, respectively,[17] although associated with high specificity and low sensitivity and, consequently, high PPV but poor NPV. The addition of VCTE may further improve PPV, with LSM >12 kPa representing the high-sensitivity threshold for cirrhosis, although PPV is low (range, 0.34–0.71).[41,42]

Among patients with intermediate FIB-4 scores between 1.3 and 2.67, stage 2 fibrosis may be present, which may be independently associated with increased risk of liver outcomes.[43] In these individuals, VCTE with an LSM between 8 and 12 kPa may provide additional evidence for such "clinically significant" fibrosis. Furthermore, due to the poor sensitivity of rule-in cutoffs with high specificity for FIB-4 score, many individuals with AF fall in the intermediate FIB-4 category. Sequential application of VCTE should help identify the subset who have AF, as noted above. In those who are morbidly obese precluding VCTE, MRE should be considered. An ELF score >9.8 has also been reported to identify those with AF, and its use has been validated in diabetes clinics and primary care populations.[44,45] More data are needed to establish the utility of SWE and other emerging tools for this purpose. A proposed approach to risk stratification of patients with suspected NAFLD, as reviewed in recent guidance statements,[10–12] is detailed in Figure 1.

Best Practice Advice 6: Liver biopsy should be considered for patients with NIT results that are indeterminate or discordant; conflict with other clinical, laboratory, or radiologic findings; or when alternative etiologies for liver disease are suspected.

In the context of clinical evaluation, some patients fall into categories poorly stratified by NITs. From well-phenotyped patients enrolled in clinical trials, a large percentage of patients (43%–51%) with serum-based biomarkers fall into an indeterminate range.[22] From a real-world perspective, these indeterminate results will often prompt sequential tests. Discordant results from sequential tests, concern for competing or alternative diagnoses, or clinical parameters suggestive of AF, should prompt consideration for liver biopsy to confirm diagnosis and stage severity of hepatic fibrosis.

Best Practice Advice 7: Serial longitudinal disease monitoring using NITs for assessment of disease progression or regression may inform clinical management.

The utility of repeated serum-based NITs during longitudinal monitoring has been associated with histologic improvement in clinical trials. In the phase 3 REGENERATE study of patients with NASH and fibrosis stage F2 or F3 (n = 931) randomized to receive placebo, obeticholic acid 10 mg, or obeticholic acid 25 mg once daily, reductions from baseline in liver aminotransferase levels, as well as in FIB-4 score, FibroTest, FibroMeter, FAST scores, and LSM by VCTE were observed in obeticholic acid–treated vs placebo-treated patients at month 18.[46] Changes in NITs were associated with shifts in histologic fibrosis stage, with the greatest improvements observed in patients with 1-stage or more fibrosis improvement.[46] The dynamic changes in NITs separated histologic responders from nonresponders, suggesting that NITs may be useful alternatives to liver biopsy in assessing the response to therapy of patients with NASH.

From a recent phase 3 trial in subjects with AF, a threshold improvement (defined as an improvement of at least 0.5 unit) in ELF and LSM by VCTE was associated with statistically significant reductions in liver biochemistry, other NITs, and clinical parameters that may represent underlying changes in metabolism not yet associated with histologic change.[47]

In patients with compensated cirrhosis due to NASH (n = 1135) enrolled into the selonsertib and simtuzumab studies, compared with nonregressors, patients with cirrhosis regression had greater reductions during follow-up in hepatic collagen content and α–smooth muscle actin expression by morphometry, machine-learning–based parameters of fibrosis, ELF, and LSM by VCTE.[48] In another study of patients with cirrhosis (n = 198) followed up for 693 patient-years (median 43 months; interquartile range, 26–58 months), clinical outcomes were predicted by liver iron-corrected T1 > 825 ms with a hazard ratio of 9.9 (95% CI, 1.29–76.4; P = .007), and outperformed VCTE in its predictive value after taking into account technical failure and unreliability.[49]

Additional evidence for longitudinal prediction of fibrosis regression and progression and response to intervention (lifestyle and pharmacologic) is needed in trials and real-world clinical practice. Strong evidence-based recommendations cannot be made, given the limitations of reproducible data in large cohorts with adequate follow-up.

Best Practice Advice 8: Patients with NAFLD and NIT results suggestive of AF (F3) or cirrhosis (F4) should be considered for surveillance of liver complications (eg, HCC screening and variceal screening per Baveno criteria).

Treatment pathway studies reveal the benefit of risk stratification and specialty referral from general

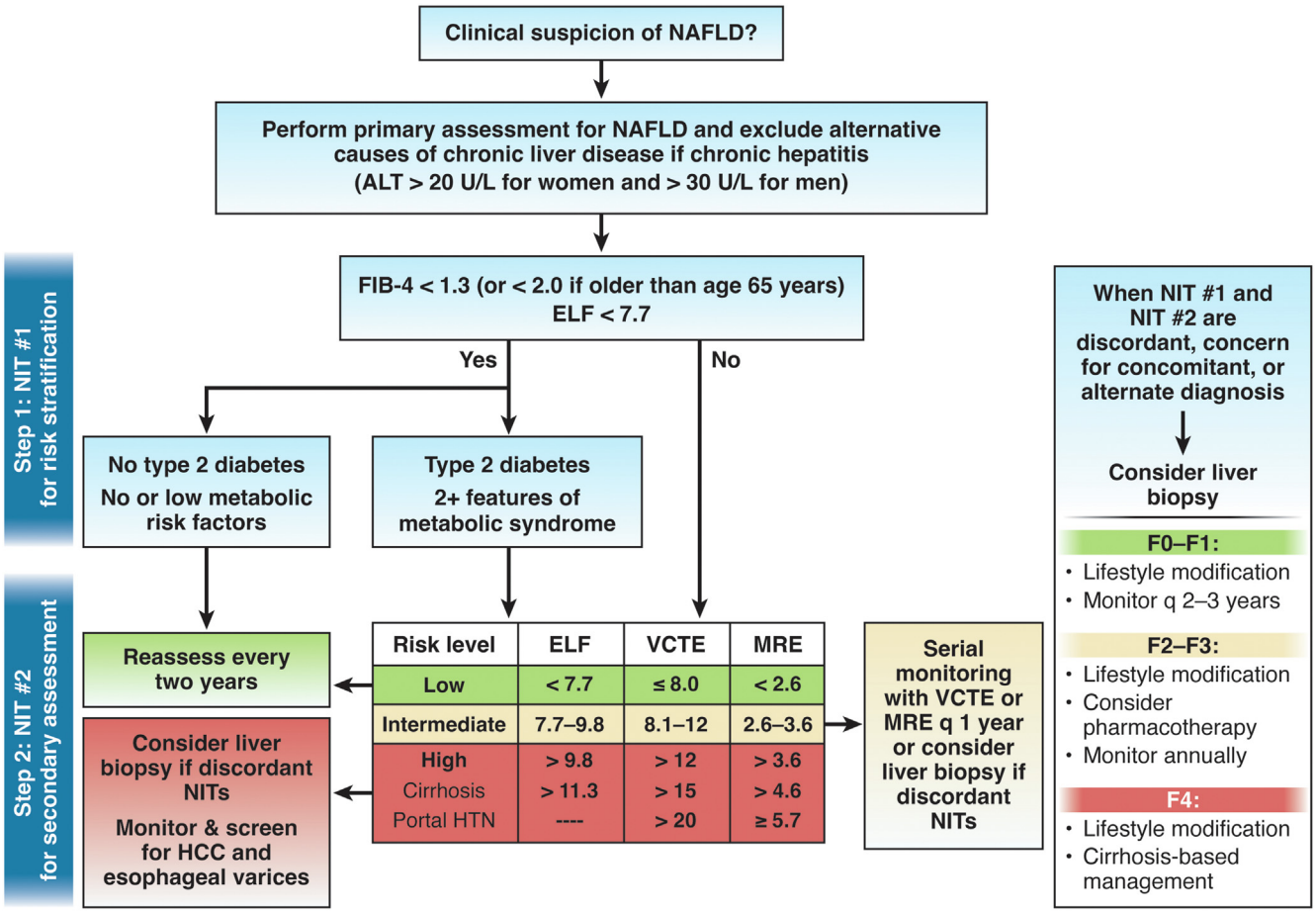


Figure 1. Algorithm for evaluation of patients with suspected NAFLD. Patients with clinical suspicion of NAFLD, such as those with metabolic risk factors, unexplained elevation in liver aminotransferases or who are noted to have hepatic steatosis on abdominal imaging, should undergo further evaluation. Screening and/or serologic evaluation for alternate causes of chronic hepatitis is recommended if liver aminotransferases are elevated (ALT >20 U/L for women and >30 U/L for men) or as clinically indicated. The Fibrosis 4 Index (FIB-4) can be used as a first-line point-of-care test. When there is low risk for advanced hepatic fibrosis (ie, absence of diabetes mellitus and/or features of metabolic syndrome), a FIB-4 <1.3 can, with a high negative predictive value, exclude those with advanced hepatic fibrosis. Such patients can be reassessed every 2 years. In patients with FIB-4 ≥1.3, a secondary assessment (enhanced liver fibrosis [ELF], vibration-controlled elastography [VCTE] or magnetic resonance elastography [MRE]) can be performed as accessible or feasible. Patients with prediabetes/T2DM or 2 or more metabolic risk factors are at higher risk for hepatic fibrosis and should have sequential or consecutive testing with a second noninvasive test (NIT). In patients older than 65 years of age, a FIB-4 cutoff of <2.0 should be used to exclude advanced hepatic fibrosis. When NITs are indeterminant or discordant or there is clinical suspicion of more advanced disease, a liver biopsy may be considered. Identification of cirrhosis should prompt referral to a specialist for cirrhosis-based management (ie, screening for HCC and esophageal varices). Patients at all stages of disease should be counseled on lifestyle modifications. Specific threshold values of NITs are approximations supported by current evidence and are meant to guide clinical management rather than be interpreted in isolation. ALT, alanine aminotransferase; ELF, enhanced liver fibrosis; HCC, hepatocellular carcinoma; MRE, magnetic resonance elastography; NAFLD, nonalcoholic fatty liver disease; NIT, noninvasive test; VCTE, vibration-controlled elastography.

populations from both a cost-saving perspective (decreased referrals to specialists at early stages of disease) and a disease management perspective (avoidance of increased cost from decompensated disease presenting at late stages). Liver-related complications, such as variceal hemorrhage, ascites, hepatic encephalopathy, and HCC have been associated with AF through prospective analysis.[6] Patients with AF may benefit from specialist referrals for dedicated HCC surveillance, variceal screening, clinical trial enrollment, and referral for liver transplantation when appropriate according to Model for End-Stage Liver Disease or decompensation. Although data are discrepant for mortality benefits from HCC surveillance, especially in Western populations, systematic reviews have demonstrated improvement in early tumor detection, receipt of curative treatment, and overall survival.[39] Endoscopic surveillance and risk stratification for varices in accordance with Baveno VI criteria—criteria based on LSM and platelet count—are used for ruling out the presence of varices needing treatment.[50]

Furthermore, patients with NAFLD and NITs suggestive of advanced hepatic fibrosis (F3) or (F4) should be monitored

with serial NITs, preferentially LSM by VCTE or MRE, given its correlation with clinically significant portal hypertension (CSPH) and clinical decompensation. A growing body of evidence suggests that NITs, specifically LSM by both VCTE[51] and MRE,[52,53] can predict CSPH and clinical outcomes in NAFLD. An LSM by VCTE $\geq$25 kPa reliably diagnosed CSPH across multiple etiologies of compensated advanced chronic liver disease.[51] In patients with NASH, a new model (ANTICIPATE-NASH model) considering body mass index, LSM $\leq$15 kPa plus platelets $\geq$150,000/mm^3 ruled out CSPH, although an LSM by VCTE $\geq$25 kPa is sufficient to rule in CSPH in non-obese patients with NASH.[54] CSPH can be diagnosed using the ANTICIPATE model if platelet count is $\leq$150,000/mm^3 and LSM by VCTE $\geq$20 kPa.[54] Furthermore, 2 clinical studies among patients with NAFLD identified LSM as a predictor of liver-related outcomes and survival.[55,56] Change in LSM on repeat VCTE correlated with liver decompensation in addition to all-cause mortality.[56]

As with VCTE, baseline LSM by MRE predicts future hepatic decompensation.[52,53] Approximately 30% of patients with baseline LSM-MRE of $\geq$5.7 kPa developed hepatic decompensation and 66% of patients with baseline LSM-MRE of $\geq$6.8 kPa had either death or transplantation.[53] Among patients with NAFLD with compensated cirrhosis at baseline by LSM via MRE, progression in LSM ($\geq$19% change from baseline) predicted hepatic decompensation or death occurred in 100% of LSM progressors and 19% of nonprogressors ($P < .001$) over a median of 2.5 years of follow-up.[52]

Conclusions

The health care burden of longitudinal management of patients with NAFLD is significant. The emergence and utilization of NITs in gastroenterology practices has the potential to significantly enhance the care of patients with NAFLD by improving detection of patients with AF who are at increased risk for cirrhosis, hepatic decompensation, and HCC, thereby facilitating timely clinical management. In this Expert Review, we have provided clinicians with best practice advice for optimal utilization of NITs in patients with NAFLD.

References

1. Younossi ZM, Golabi P, Paik JM, et al. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology 2023;77:1335–1347.
2. Allen AM, Lazarus JV, Younossi ZM. Healthcare and socioeconomic costs of NAFLD: a global framework to navigate the uncertainties. J Hepatol 2023;79:209–217.
3. Younossi ZM. Non-alcoholic fatty liver disease - a global public health perspective. J Hepatol 2019;70:531–544.
4. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology 2015;149:389–397.e10.
5. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898–1906.
6. Sanyal AJ, Van Natta ML, Clark J, et al. Prospective study of outcomes in adults with nonalcoholic fatty liver disease. N Engl J Med 2021;385:1559–1569.
7. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317–1325.
8. McPherson S, Stewart SF, Henderson E, et al. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut 2010;59:1265–1269.
9. Unalp-Arida A, Ruhl CE. Liver fibrosis scores predict liver disease mortality in the United States population. Hepatology 2017;66:84–95.
10. Kanwal F, Shubrook JH, Adams LA, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. Gastroenterology 2021;161:1657–1669.
11. Cusi K, Isaacs S, Barb D, et al. American Association of Clinical Endocrinology Clinical Practice Guideline for the Diagnosis and Management of Nonalcoholic Fatty Liver Disease in Primary Care and Endocrinology Clinical Settings: Co-Sponsored by the American Association for the Study of Liver Diseases (AASLD). Endocr Pract 2022;28:528–562.
12. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology 2023;77:1797–1835.
13. Miele L, De Michele T, Marrone G, et al. Enhanced Liver Fibrosis test as a reliable tool for assessing fibrosis in nonalcoholic fatty liver disease in a clinical setting. Int J Biol Markers 2017;32:e397–e402.
14. Guha IN, Parkes J, Roderick P, et al. Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: validating the European Liver Fibrosis Panel and exploring simple markers. Hepatology 2008;47:455–460.
15. Xiao G, Zhu S, Xiao X, et al. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: a meta-analysis. Hepatology 2017;66:1486–1501.
16. Morling JR, Fallowfield JA, Guha IN, et al. Using non-invasive biomarkers to identify hepatic fibrosis in people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. J Hepatol 2014;60:384–391.
17. Shah AG, Lydecker A, Murray K, et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2009;7:1104–1112.
18. Sun W, Cui H, Li N, et al. Comparison of FIB-4 index, NAFLD fibrosis score and BARD score for prediction of advanced fibrosis in adult patients with non-alcoholic fatty liver disease: a meta-analysis study. Hepatol Res 2016;46:862–870.
19. Mosca A, Della Volpe L, Alisi A, et al. Non-invasive diagnostic test for advanced fibrosis in adolescents with

non-alcoholic fatty liver disease. Front Pediatr 2022;10:885576.
20. van Kleef LA, Sonneveld MJ, de Man RA, et al. Poor performance of FIB-4 in elderly individuals at risk for chronic liver disease - implications for the clinical utility of the EASL NIT guideline. J Hepatol 2022;76:245–246.
21. Srivastava A, Gailer R, Tanwar S, et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. J Hepatol 2019;71:371–378.
22. Anstee QM, Lawitz EJ, Alkhouri N, et al. Noninvasive tests accurately identify advanced fibrosis due to NASH: baseline data from the STELLAR Trials. Hepatology 2019;70:1521–1530.
23. 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;5:362–373.
24. Jung J, Loomba RR, Imajo K, et al. MRE combined with FIB-4 (MEFIB) index in detection of candidates for pharmacological treatment of NASH-related fibrosis. Gut 2021;70:1946–1953.
25. Majumdar A, Campos S, Gurusamy K, et al. Defining the minimum acceptable diagnostic accuracy of noninvasive fibrosis testing in cirrhosis: a decision analytic modeling study. Hepatology 2020;71:627–642.
26. Mozes FE, Lee JA, Selvaraj EA, et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. Gut 2022;71:1006–1019.
27. Tamaki N, Imajo K, Sharpton S, et al. Magnetic resonance elastography plus Fibrosis-4 versus FibroScan-aspartate aminotransferase in detection of candidates for pharmacological treatment of NASH-related fibrosis. Hepatology 2022;75:661–672.
28. Kim BK, Tamaki N, Imajo K, et al. Head-to-head comparison between MEFIB, MAST, and FAST for detecting stage 2 fibrosis or higher among patients with NAFLD. J Hepatol 2022;77:1482–1490.
29. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. J Hepatol 2008;48:835–847.
30. Chen J, Yin M, Talwalkar JA, et al. Diagnostic performance of MR Elastography and vibration-controlled transient elastography in the detection of hepatic fibrosis in patients with severe to morbid obesity. Radiology 2017;283:418–428.
31. Vuppalanchi R, Siddiqui MS, Van Natta ML, et al. Performance characteristics of vibration-controlled transient elastography for evaluation of nonalcoholic fatty liver disease. Hepatology 2018;67:134–144.
32. Kjaergaard M, Thiele M, Jansen C, et al. High risk of misinterpreting liver and spleen stiffness using 2D shear-wave and transient elastography after a moderate or high calorie meal. PLoS One 2017;12:e0173992.
33. Nascimbeni F, Lebray P, Fedchuk L, et al. Significant variations in elastometry measurements made within short-term in patients with chronic liver diseases. Clin Gastroenterol Hepatol 2015;13:763–771.e1–e6.
34. Chow JC, Wong GL, Chan AW, et al. Repeating measurements by transient elastography in non-alcoholic fatty liver disease patients with high liver stiffness. J Gastroenterol Hepatol 2019;34:241–248.
35. Chuah KH, Lai LL, Vethakkan SR, et al. Liver stiffness measurement in non-alcoholic fatty liver disease: two is better than one. J Gastroenterol Hepatol 2020;35:1404–1411.
36. Imajo K, Honda Y, Kobayashi T, et al. Direct comparison of US and MR elastography for staging liver fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2022;20:908–917.e11.
37. Serai SD, Obuchowski NA, Venkatesh SK, et al. Repeatability of MR elastography of liver: a meta-analysis. Radiology 2017;285:92–100.
38. Vilar-Gomez E, Lou Z, Kong N, et al. Cost effectiveness of different strategies for detecting cirrhosis in patients with nonalcoholic fatty liver disease based on united states health care system. Clin Gastroenterol Hepatol 2020;18:2305–2314.e12.
39. Singal AG, Pillai A, Tiro J. Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis. PLoS Med 2014;11:e1001624.
40. Kleiner DE, Brunt EM, Wilson LA, et al. Association of histologic disease activity with progression of nonalcoholic fatty liver disease. JAMA Netw Open 2019;2:e1912565.
41. Siddiqui MS, Vuppalanchi R, Van Natta ML, et al. Vibration-controlled transient elastography to assess fibrosis and steatosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2019;17:156–163.e2.
42. Hsu C, Caussy C, Imajo K, et al. Magnetic resonance vs transient elastography analysis of patients with nonalcoholic fatty liver disease: a systematic review and pooled analysis of individual participants. Clin Gastroenterol Hepatol 2019;17:630–637.e8.
43. Dulai PS, Singh S, Patel J, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatology 2017;65:1557–1565.
44. Fagan KJ, Pretorius CJ, Horsfall LU, et al. ELF score >/=9.8 indicates advanced hepatic fibrosis and is influenced by age, steatosis and histological activity. Liver Int 2015;35:1673–1681.
45. Patel P, Hossain F, Horsfall LU, et al. A pragmatic approach identifies a high rate of nonalcoholic fatty liver disease with advanced fibrosis in diabetes clinics and at-risk populations in primary care. Hepatol Commun 2018;2:893–905.
46. Rinella ME, Dufour JF, Anstee QM, et al. Non-invasive evaluation of response to obeticholic acid in patients with NASH: results from the REGENERATE study. J Hepatol 2022;76:536–548.
47. Harrison SA, Wong VW, Okanoue T, et al. Selonsertib for patients with bridging fibrosis or compensated cirrhosis due to NASH: results from randomized phase III STELLAR trials. J Hepatol 2020;73:26–39.

48. Sanyal AJ, Anstee QM, Trauner M, et al. Cirrhosis regression is associated with improved clinical outcomes in patients with nonalcoholic steatohepatitis. Hepatology 2022;75:1235–1246.
49. Jayaswal ANA, Levick C, Selvaraj EA, et al. Prognostic value of multiparametric magnetic resonance imaging, transient elastography and blood-based fibrosis markers in patients with chronic liver disease. Liver Int 2020; 40:3071–3082.
50. Petta S, Sebastiani G, Bugianesi E, et al. Non-invasive prediction of esophageal varices by stiffness and platelet in non-alcoholic fatty liver disease cirrhosis. J Hepatol 2018;69:878–885.
51. de Franchis R, Bosch J, Garcia-Tsao G, et al. Baveno VII - renewing consensus in portal hypertension. J Hepatol 2022;76:959–974.
52. Gidener T, Dierkhising RA, Mara KC, et al. Change in serial liver stiffness measurement by magnetic resonance elastography and outcomes in NAFLD. Hepatology 2023;77:268–274.
53. Gidener T, Yin M, Dierkhising RA, et al. Magnetic resonance elastography for prediction of long-term progression and outcome in chronic liver disease: a retrospective study. Hepatology 2022;75:379–390.
54. Pons M, Augustin S, Scheiner B, et al. Noninvasive diagnosis of portal hypertension in patients with compensated advanced chronic liver disease. Am J Gastroenterol 2021;116:723–732.
55. Kamarajah SK, Chan WK, Nik Mustapha NR, et al. Repeated liver stiffness measurement compared with paired liver biopsy in patients with non-alcoholic fatty liver disease. Hepatol Int 2018;12:44–55.
56. Petta S, Sebastiani G, Vigano M, et al. Monitoring occurrence of liver-related events and survival by transient elastography in patients with nonalcoholic fatty liver disease and compensated advanced chronic liver disease. Clin Gastroenterol Hepatol 2021;19:806–815.e5.

Received December 10, 2022. Accepted June 9, 2023.

Correspondence
Address correspondence to: Julia J. Wattacheril, MD, MPH, Nonalcoholic Fatty Liver Disease Program, Center for Liver Disease and Transplantation, Columbia University, New York Presbyterian Hospital, 622 West 168th Street, PH 14 105-D, New York, New York 10032. e-mail: jjw2151@cumc.columbia.edu.

Author Contributions
All authors contributed to drafting of the manuscript and critical revision of the manuscript and approved the final version.

Conflicts of interest
The authors disclose the following: Julia J. Wattacheril: research grants (paid to institution): Intercept, Galectin, GENFIT, and Pfizer; consulting/advisory: AstraZeneca and AMRA Medical. Manal F. Abdelmalek: research grants (paid to institution): Intercept, Allergan, Madrigal, Viking, Celgene, Genentech, Novo Nordisk, BMS, NGM Bio, Poxel, Durect, Galmed, Gilead, Hanmi, Boeringher Ingelheim, and Inventiva; consulting/advisory; Intercept, SonicIncytes, NGM Bio, BMS, Inventiva, Madrigal, Hanmi, Merck, Novo Nordisk, and 89Bio. Joseph K. Lim: research grants (paid to institution): Allergan, Celgene, GENFIT, Intercept, Viking, and Pfizer. Arun J. Sanyal: research grants (paid to institution): Intercept, Lilly, Novo Nordisk, Echosense, Boehringer Ingelhiem, Pfizer, Merck, Bristol Myers Squibb, Hanmi, Madrigal, Galmed, Gilead, Salix, and Malinckrodt; consultant: Intercept, Gilead, Merck, NGM Bio, Terns, Regeneron, Alnylam, Amgen, Genentech, Pfizer, Novo Nordisk, Astra Zeneca, Salix, Malinckrodt, Lilly, Histoindex, Path AI, Rivus, Hemoshear, Northsea, 89Bio, Altimmune, Surrozen, and Poxel; ownership interests: Tiziana, Durect, Exhalenz, GENFIT, Galmed, Northsea, and Hemoshear.