

Non-Invasive Tests in the Management of Metabolic Dysfunction-Associated Steatotic Liver Disease: From Diagnosis to Treatment Monitoring

Han Ah Lee, Young Youn Cho, Hyung Joon Kim

Department of Internal Medicine, Chung-Ang University College of Medicine, Seoul, Korea

Metabolic dysfunction-associated steatotic liver disease (MASLD) affects approximately 30% of the global adult population, and accurate identification of patients with clinically significant fibrosis is critical for risk stratification and therapeutic decision-making. Non-invasive tests (NITs)—encompassing serum-based biomarkers, vibration-controlled transient elastography (VCTE), magnetic resonance elastography, and magnetic resonance imaging-derived proton density fat fraction—have emerged as indispensable tools across the MASLD management continuum. The sequential two-step pathway combining FIB-4 with liver stiffness measurement by VCTE has been validated as the standard framework for fibrosis triage, demonstrating a more than 20-fold gradient in 5-year liver-related event rates between low- and high-risk categories. In the clinical trial setting, NITs have progressively evolved from exploratory secondary endpoints to regulatory co-primary endpoints, culminating in the FDA approvals of resmetirom (2024) and semaglutide (2025) for noncirrhotic metabolic dysfunction-associated steatohepatitis with F2–F3 fibrosis. NIT-based algorithms are now formalized for both patient selection and serial treatment response monitoring, effectively operationalizing a biopsy-sparing treat-and-monitor paradigm in routine clinical practice. Achieving full regulatory acceptance of NIT-only trial designs—requiring prospective validation that NIT-defined response categories predict long-term liver-related outcomes—represents the next critical frontier in MASLD management.

Keywords: Metabolic dysfunction-associated steatotic liver disease; Non-invasive test; Resmetirom; Semaglutide

INTRODUCTION

Metabolic dysfunction-associated steatotic liver disease (MASLD) has emerged as the most prevalent chronic liver disease worldwide, affecting approximately 30% of the global adult population and imposing a substantial burden on public health systems [1]. Within the MASLD spectrum, metabolic dysfunction-associated steatohepatitis (MASH) with significant

fibrosis (stage F2–F3) represents the critical inflection point at which the risk of liver-related complications—including cirrhosis, hepatocellular carcinoma (HCC), and liver-related mortality—increases substantially [2,3]. Identification of patients at this stage is therefore important, both for risk stratification and for guiding therapeutic intervention.

Liver biopsy remains the reference standard for MASLD staging, yet its procedural risks, sampling variability, inter-ob-

Address for Correspondence: Han Ah Lee, M.D., Ph.D.

Department of Internal Medicine, Chung-Ang University College of Medicine, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Korea
E-mail: amelia86@naver.com
ORCID: <https://orcid.org/0000-0002-0430-1607>

Received : 2026. 4. 17.

Revised : 2026. 5. 15.

Accepted : 2026. 5. 17.

server discordance, and impracticality for serial assessments have driven the development of non-invasive alternatives [4]. This need has become especially pressing in the context of clinical drug development, where the reliance on biopsy-based endpoints has historically constrained trial feasibility, increased patient burden, and prolonged the path to regulatory approval [5,6]. Non-invasive tests (NITs) encompass a heterogeneous toolkit of serum-based biomarkers, imaging modalities, and composite scoring systems that can quantify hepatic steatosis, necroinflammation, and fibrosis without tissue sampling [7].

The role of NITs in MASLD management has evolved considerably over the past decade. Sequential application of Fibrosis-4 (FIB-4) followed by liver stiffness measurement (LSM) by vibration-controlled transient elastography (VCTE) has been endorsed by multiple international guidelines as the recommended two-step pathway for fibrosis risk stratification in MASLD [2,3]. Initially deployed as exploratory or secondary endpoints in early-phase trials, NITs have progressively assumed co-primary and, in some cases, surrogate endpoint status in pivotal studies [6]. This paradigm shift was exemplified by the MAESTRO-NASH trial of resmetirom, in which magnetic resonance imaging-derived proton density fat fraction (MRI-PDFF) served as a co-primary endpoint alongside histological assessment, culminating in the first Food and Drug Administration (FDA) approval for a MASH-specific pharmacotherapy in March 2024 [8]. Concurrently, the phase 3 ESSENCE trial of semaglutide incorporated both LSM and MRI-PDFF as co-primary endpoints, further cementing the regulatory legitimacy of NIT-based efficacy assessment [9].

Beyond their diagnostic utility, NITs now occupy a central role in guiding treatment decisions in routine clinical practice [8,9]. Published expert guidance documents for both resmetirom and semaglutide have formalized NIT-based algorithms for patient selection, treatment initiation, and longitudinal monitoring—effectively operationalizing a treat-and-monitor paradigm that circumvents the need for repeat liver biopsy in most patients [10,11]. These frameworks reflect a broader recognition that NITs, when applied sequentially and interpreted in clinical context, can adequately capture the fibrosis trajectory that determines prognosis and therapeutic response.

In this review, we summarize the evolving role of NITs across the MASLD management continuum—from disease staging and clinical trial endpoint validation to treatment guidance—and further discuss future directions in NIT-based risk assessment for MASLD-related complications and extrahepatic outcomes.

REPRESENTATIVE NITS IN MASLD MANAGEMENT

Serum-Based Biomarkers

Serum-based NITs represent the most accessible and widely deployed tools for fibrosis assessment in MASLD, owing to their low cost, broad availability, and ease of calculation in routine clinical settings.

The most widely used traditional serum-based scores for fibrosis assessment in MASLD include the FIB-4 index, the non-alcoholic fatty liver disease (NAFLD) fibrosis score (NFS), and the aspartate aminotransferase (AST)-to-platelet ratio index (APRI) [7]. Among these, FIB-4 has been most extensively validated: values below 1.3 effectively rule out advanced fibrosis ($F \geq 3$), while values above 2.67 raise the probability of its presence [12,13]. Despite its strong negative predictive value (NPV), FIB-4 demonstrates only moderate diagnostic accuracy for confirming advanced fibrosis, reflected by a low positive predictive value (PPV) [12]. Age is an important confounder—the exclusionary cut-off should be raised to 2.0 in patients older than 65 years, and the index shows suboptimal performance in those younger than 35 [13].

The enhanced liver fibrosis (ELF) score, a composite of hyaluronic acid, tissue inhibitor of metalloproteinase 1 (TIMP-1), and the amino-terminal propeptide of type III procollagen (PIIINP), captures direct markers of extracellular matrix turnover. A meta-analysis demonstrated that an ELF score below 7.7 achieves a sensitivity of 93% for excluding advanced fibrosis, while a score exceeding 9.8 yields a specificity of 86% for its confirmation, with an area under the receiver operating characteristic curve (AUROC) of 0.83 (0.71–0.90) for advanced fibrosis [14]. Despite these performance characteristics, the individual components of the ELF score require specialized

laboratory assays that are not universally available, and the PPV is reduced in low-prevalence primary care populations, increasing the risk of false-positive results.

Serum procollagen type III N-terminal peptide (PRO-C3) reflects active hepatic fibrogenesis and serves as a more dynamic marker of fibrosis activity than traditional fibrosis indices. The ADAPT score integrates PRO-C3 with age, pre-existing diabetes, and platelet count, achieving AUROCs of 0.86 and 0.87 in derivation and validation cohorts, respectively, for advanced fibrosis [15]. Among various algorithms evaluated for identifying advanced fibrosis, a stepwise combination of ADAPT followed by Agile 4 yielded the highest overall diagnostic accuracy [16]. A key practical limitation, however, is the dependence on specialized laboratory infrastructure for PRO-C3 quantification, which restricts its availability in resource-limited settings.

Several novel serum-based panels have been developed to improve identification of at-risk MASH—defined as MASH with a NAFLD activity score ≥ 4 and fibrosis $\geq F2$ —which constitutes the primary target population for current pharmacotherapies. The NIS2+ platform, an optimized version of the NIS4 technology, incorporates serum miR-34a-5p and YKL-40, achieving an AUROC of 0.813 in the validation cohort, surpassing both NIS4 (0.792) and FIB-4 (0.653) [17]. A notable advantage is that NIS2+ performance is unaffected by age, sex, body mass index (BMI), or diabetes status, enabling consistent diagnostic performance across patient subgroups [17]. However, measurement of serum miR-34a-5p requires complex laboratory equipment and technical expertise, limiting its deployment outside specialized centers [18].

The MACK-3 score, validated in a multicenter cohort of 1,924 biopsy-proven MASLD patients, integrates AST, homeostasis model assessment of insulin resistance (HOMA-IR), and cytokeratin-18 (CK-18), yielding an AUROC of 0.79 for fibrotic MASH, with high sensitivity (91%) and specificity (85%) [19]. Its adoption is similarly constrained by the limited availability of CK-18 assays in routine clinical laboratories and a relatively high proportion of indeterminate results.

The fibrotic NASH index (FNI), derived from AST, high-density lipoprotein cholesterol, and glycated hemoglobin, offers the practical advantage of relying exclusively on routine

laboratory parameters, making it particularly suitable for primary care settings [20]. AUROCs of FNI in external validation cohorts ranged from 0.80 to 0.95. Nonetheless, FNI carries a low PPV, particularly in low-prevalence populations, and requires further validation in more diverse ethnic groups.

Imaging-Based Tests

Assessment of hepatic fibrosis

Imaging-based NITs offer the advantage of direct, quantitative assessment of hepatic steatosis and fibrosis, overcoming several limitations of serum-based markers that are influenced by systemic metabolic and inflammatory factors.

VCTE, measures LSM by quantifying the velocity of shear waves propagated through hepatic parenchyma. A meta-analysis demonstrated AUROCs of 0.85 for VCTE-derived LSM in identifying advanced fibrosis, outperforming FIB-4 (0.76), NFS (0.73), and APRI (0.70) [21]. In the context of MASLD fibrosis staging, an LSM below 8 kPa effectively excludes advanced fibrosis, values between 8 and 12 kPa suggest significant fibrosis ($F \geq 2$), and values exceeding 12 kPa indicate advanced fibrosis ($F \geq 3$); an LSM of 20 kPa or greater, particularly when combined with a FIB-4 above 3.48, indicates cirrhosis [21]. VCTE also simultaneously generates the controlled attenuation parameter (CAP), which quantifies hepatic steatosis through ultrasound attenuation and provides a non-invasive estimate of liver fat content. Diagnostic accuracy of VCTE is reduced in patients with severe obesity, ascites, or active hepatic inflammation, all of which can spuriously elevate LSM independent of fibrosis stage [22].

Magnetic resonance elastography (MRE) is currently regarded as the most accurate non-invasive imaging modality for liver fibrosis assessment across all fibrosis stages. An individual patient data meta-analysis across eight independent cohorts established cut-off values of 3.14 kPa for significant fibrosis ($F \geq 2$), 3.53 kPa for advanced fibrosis ($F \geq 3$), and 4.45 kPa for cirrhosis ($F \geq 4$), with AUROCs consistently in the range of 0.92–0.94 and corresponding sensitivities and specificities exceeding 85% [23]. A notable limitation, however, is that elevated hepatic inflammatory activity and serum gamma-glutamyl transferase levels may lead to overestimation of early-stage fibrosis ($F0$ –

Table 1. NITs for assessing liver disease severity in MASLD

Category	Tool	Formula/components	Cut-off values	Advantages	Limitations
Serum biomarkers					
Simple fibrosis score	FIB-4 index	Age (yr)×AST (IU/L) Platelet ($\times 10^9/L$)× $\sqrt{ALT}$ (IU/L)	<1.3 (or <2.0 if age ≥65 yr): Low risk—advanced fibrosis unlikely; safe to monitor in primary care >2.67: High risk—refer to hepatology 1.3–2.67: Intermediate—proceed to second-line NIT Note: Not sufficient alone for treatment eligibility	Universally available from routine blood tests; low cost; validated as first-line triage in AGA, AASLD, EASL, and KASL guidelines; high NPV (>95%) for excluding advanced fibrosis in primary care	Performance confounded by age, diabetes, and liver inflammation; low PPV for ruling in fibrosis; requires age-adjusted threshold in elderly; should not be used as standalone treatment eligibility criterion
	APRI	AST/ULN Platelet ($\times 10^9/L$)×100	0.5 (rule out)/1.5 (rule in) MASLD-specific 0.7/1.0	Inexpensive; simple two-variable formula (AST, platelets) using routine labs; broadly accessible at the point of care	Lower accuracy than FIB-4 in MASLD (AUROC ~0.70); originally validated in HCV; confounded by extrahepatic AST elevation and non-hepatic thrombocytopenia
	NFS	$-1.675 + 0.037 \times \text{age} + 0.094 \times \text{BMI} + 1.13 \times \text{IFG/T2DM (yes=1, no=0)} - 0.99 \times \text{AST/ALT} - 0.013 \times \text{Platelet (}\times 10^9/L\text{)} - 0.66 \times \text{Albumin (g/dL)}$	<1.455: Excludes advanced fibrosis (F≥3) >0.676: Suggests advanced fibrosis (F≥3)	Incorporates metabolic parameters (BMI, diabetes); validated in large NAFLD cohorts; no additional laboratory cost beyond routine tests	Reduced specificity in low-prevalence settings; influenced by BMI and metabolic comorbidities independently of fibrosis; less widely endorsed than FIB-4 in current guidelines
Fibrosis-specific biomarker	ELF	$-7.412 + 0.681 \times \ln(\text{HA [ng/mL]}) + 0.775 \times \ln(\text{PIIINP [ng/mL]}) + 0.494 \times \ln(\text{TIMP1 [ng/mL]})$	9.2–10.5: Consistent with F2–F3; recommended range for resmetrom and semaglutide treatment candidacy 10.5–11.3: Individualized decision; exclude cirrhosis before treating >11.3: Consistent with cirrhosis; treatment contraindicated <9.2: Advanced fibrosis unlikely	Directly measures extracellular matrix components; high sensitivity and specificity for significant fibrosis; treatment eligibility thresholds validated in both MAESTRO-NASH and ESSENCE trials; useful where elastography is unavailable	Requires dedicated immunoassay platform; higher cost than simple serum scores; less widely available outside tertiary centers; ELF thresholds for treatment selection are based on trial IQR data rather than AASLD NILDA guideline recommendations
PRO-C3–based scores	ADAPT	$\exp(\log_{10}(\text{age [yr]} \times \text{PRO-C3 [ng/mL]} / \sqrt{\text{Platelet } [\times 10^9/L]} + \text{T2DM (yes=1, no=0)}))$	<0.46: At-risk MASH unlikely ≥0.46: Possible at-risk MASH—further evaluation recommended >1.55: High-risk MASH—specialist referral indicated	Integrates a specific fibrogenesis marker (PRO-C3) with clinical variables; higher diagnostic value than NFS for detecting at-risk MASH; validated in multi-ethnic cohorts	PRO-C3 assay not universally available; requires dedicated laboratory infrastructure; limited head-to-head comparison data with ELF or elastography in treatment selection context

Table 1. Continued

Category	Tool	Formula/components	Cut-off values	Advantages	Limitations
Multi-analyte panel	NIS2+	$e^{\beta_0/(1 + e^{\beta_1})}$, where: $y = \beta_0 + \beta_1 \times \log_{10}(\text{miR-34a-5p})$ $+ \beta_2 \times \log_{10}(\text{YKL-40 [ng/mL]})$ $+ \beta_3 \times \text{sex (F=0, M=1)}$ $+ \beta_4 \times \log_{10}(\text{miR-34a-5p}) \times \text{sex}$	≥0.36: Suggests at-risk MASH <0.15: At-risk MASH unlikely	Incorporates microRNA biomarkers; high sensitivity (91%) and specificity (82%) for at-risk MASH; performance maintained across BMI subgroups	Requires microRNA quantification, which is not yet routinely available in clinical laboratories; limited external validation data; GGT confounding may reduce specificity
	MACK-3	Combination of: AST, miR-34a-IR, cytokeratin-18 (M30)	≥0.18: Suggests at-risk MASH <0.088: At-risk MASH unlikely	Based on widely used laboratory and apoptosis markers; combines inflammation and fibrosis signals; suitable for primary and non-specialist settings	Requires cytokeratin-18 and microRNA assays; limited availability at primary care level; less validated in prospective cohorts for treatment guidance
	FNI	Combination of: AST, HDL-cholesterol, HbA1c	≥0.33: Suggests at-risk MASH	Utilizes routine clinical chemistry parameters; no additional laboratory cost; simultaneously reflects inflammation and metabolic dysfunction	Sensitivity reduced in early fibrosis (F0–F1); limited prognostic validation data; not yet incorporated into major treatment guidance documents
Imaging-based tests					
LSM	VCTE (FibroScan®)	LSM (kPa): propagation velocity of a shear wave generated by a vibrating probe, measured in a cylinder of liver tissue (~1 cm wide, 4 cm long) CAP (dB/m): ultrasonic attenuation—co-acquired steatosis measure (not used for fibrosis staging or treatment monitoring)	<8 kPa: Low-risk; advanced fibrosis unlikely 8–15 kPa: Recommended treatment candidate range (resmetrom and semaglutide) 15–20 kPa: Individualized decision; exclude cirrhosis >20 kPa: Cirrhosis range; treatment contraindicated Response threshold: ≥25% reduction (resmetrom); ≥30% reduction (semaglutide)	Point-of-care availability; rapid, reproducible, and widely validated; endorsed as primary second-line NIT in all major guidelines; QOR used to define treatment eligibility thresholds in MAESTRO-NASH and ESSENCE; applicable to treatment monitoring with same-modality repeat	Reliability reduced with BMI >35 kg/m ² , severe steatosis, or narrow intercostal spaces; operator-dependent; CAP not validated as treatment response marker and should not be used to assess futility of resmetrom therapy; requires ≥2 h fasting; LSM may be elevated by inflammation, congestion, or cholestasis independent of fibrosis
	MRE	LSM (kPa): stiffness of a liver volume mapped from propagation of mechanical shear waves induced by an external driver, acquired during standard MRI protocol	<3.1 kPa: Advanced fibrosis unlikely 3.1–4.4 kPa: Recommended treatment candidate range (resmetrom and semaglutide) 4.4–5.0 kPa: Individualized decision; exclude cirrhosis >5.0 kPa: Cirrhosis range; treatment contraindicated Response threshold: ≥20% reduction (both agents)	Highest diagnostic accuracy for fibrosis staging among all NITs; unaffected by obesity (BMI >35 kg/m ²), ascites, or narrow intercostal space; preferred when VCTE is technically limited; co-acquisition with MRI-PDFF allows simultaneous fibrosis and steatosis assessment	Limited availability (~1,400 sites in the US); higher cost than VCTE; contraindicated with ferromagnetic implants; requires dedicated radiology infrastructure; MRE was not used in ESSENCE—treatment eligibility thresholds are extrapolated from prior studies and AASLD NITDA guidelines

Table 1. Continued

Category	Tool	Formula/components	Cut-off values	Advantages	Limitations
Hepatic steatosis quantification	MRI-PDFF	Proton density fat fraction (%): ratio of mobile proton density of fat to total mobile proton density of the liver, measured by chemical shift-encoded MRI	≥5%: Hepatic steatosis present ≥30% reduction from baseline at 12 months: Associated with histological response to resmetrom (MAESTRO-NASH) ≥30% reduction at 72 weeks: Treatment response threshold for semaglutide <30% reduction: Associated with lack of histological response	Most accurate quantitative measure of hepatic steatosis; validated as treatment response biomarker in MAESTRO-NASH (co-primary endpoint) and ESSENCE; MRI-PDFF reduction correlates with MASH resolution; often co-acquired with MRE without additional contrast	Requires MRI availability; cost-prohibitive for routine monitoring in many settings; not required for treatment initiation; a ≥30% MRI-PDFF reduction does not independently predict histological response and should not be used as a standalone continuation criterion
		CAP	Ultrasonic attenuation (dB/m) measured simultaneously with LSM during VCTE; reflects degree of steatosis based on signal attenuation <248 dB/m: Steatosis grade S0–S1 248–280 dB/m: Steatosis grade S1–S2 ≥280 dB/m: Steatosis grade S2–S3 ≥30% reduction: Proposed threshold for steatosis response in semaglutide trials	Simultaneously acquired with VCTE-LSM; widely available; low additional cost; useful for confirming steatosis in treatment candidate screening; CAP reduction observed in semaglutide-treated patients correlated with MASH resolution in ESSENCE Lower accuracy than MRI-PDFF for steatosis quantification; affected by BMI, probe type (M vs. XL), and operator; not validated as standalone steatosis response marker for resmetrom—changes in CAP were NOT associated with histological response in MAESTRO-NASH and should not be used to assess futility of resmetrom treatment	

AASLD, American Association for the Study of Liver Diseases; AGA, American Gastroenterological Association; ALT, alanine aminotransferase; APRI, aspartate aminotransferase-to-platelet ratio index; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; BMI, body mass index; CAP, controlled attenuation parameter; EASL, European Association for the Study of the Liver; ELF, enhanced liver fibrosis score; FIB-4, Fibrosis-4; FNI, fibrotic NASH index; HA, hyaluronic acid; HbA1c, glycated hemoglobin; HCV, hepatitis C virus; HDL, high density lipoprotein; IFG, impaired fasting glucose; IQR, interquartile range; KASL, Korean Association for the Study of the Liver; LSM, liver stiffness measurement; MASH, metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-derived proton density fat fraction; NAFLD, non-alcoholic fatty liver disease; NFS, NAFLD fibrosis score; NILDA, noninvasive liver disease assessment; NIT, non-invasive test; NPV, negative predictive value; PLINP, N-terminal propeptide of type III procollagen; PPV, positive predictive value; T2DM, type 2 diabetes mellitus; TIMP1, tissue inhibitor of metalloproteinase 1; ULN, upper limit of normal; VCTE, vibration-controlled transient elastography. VCTE and MRE cut-offs presented herein correspond to treatment eligibility thresholds for resmetrom and semaglutide; cut-offs used for general fibrosis stratification in clinical practice are described in the main text (Section Imaging-Based Tests).

F1) [6,7]. The principal barriers to widespread adoption remain high cost, limited scanner availability in many regions, and the requirement for specialized technical infrastructure.

Hepatic steatosis quantification

MRI-PDFF has emerged as the gold standard for non-invasive quantification of hepatic steatosis in both clinical trials and increasingly in clinical practice. By measuring the fraction of mobile proton signal arising from fat relative to total tissue signal, MRI-PDFF provides highly reproducible, scanner-independent estimates of liver fat content [24]. The primary limitation of MRI-PDFF is its restricted availability outside specialist or academic centers and its relatively high cost compared with ultrasound-based alternatives.

CAP, measured concurrently with LSM during VCTE examination, provides a more accessible and cost-effective alternative to MRI-PDFF for hepatic steatosis quantification at the point of care. While CAP demonstrates adequate performance for detecting the presence of steatosis ($S \geq 1$), its accuracy for grading steatosis severity is more variable, and it is susceptible to confounding by obesity and fluctuations in serum liver enzyme levels [25]. In the context of advanced fibrosis, where hepatic fat content may be reduced, the diagnostic utility of CAP is further attenuated. Nevertheless, its co-measurement with LSM during single VCTE session renders CAP a practical tool for simultaneous steatosis and fibrosis assessment in routine clinical workflows. The NITs referenced throughout this algorithm are summarized in Table 1.

THE TWO-STEP NIT PATHWAY: FROM PRIMARY CARE TO SPECIALIST REFERRAL

Rationale and Guideline Recommendations

MASLD affects over 30% of the global adult population, yet only a minority will ultimately develop cirrhosis, hepatic decompensation, or HCC [1]. Because most patients with MASLD are first encountered in primary care rather than hepatology specialist settings, a scalable and cost-effective strategy is need-

ed to identify the subset with clinically significant fibrosis who are at risk of disease progression—without subjecting the entire population to invasive or resource-intensive testing [26]. Applying second-line imaging-based NITs such as VCTE or MRE universally to all patients with MASLD would impose considerable resource burden and is not cost-effective given the large and heterogeneous population at risk. To address this gap, a sequential two-step clinical care pathway was formally proposed by the American Gastroenterological Association (AGA) in 2021, using an inexpensive and widely available first-line test to triage patients before selectively deploying more specific second-line modalities [27]. This approach has since been endorsed with broadly consistent recommendations by international guidelines, reflecting its adoption as the standard framework for MASLD risk stratification across international guidelines [28–30].

In the two-step pathway, the FIB-4 index serves as the first-line screening test given its low cost, universal availability from routine blood tests, and validated diagnostic performance. A FIB-4 below 1.3 (adjusted to 2.0 in patients aged ≥ 65 years) effectively excludes advanced fibrosis with a NPV exceeding 95% in primary care settings, allowing safe monitoring without specialist referral [31]. A FIB-4 above 2.67 indicates high probability of advanced fibrosis and warrants direct hepatology referral. The approximately 24% of patients falling in the intermediate zone (FIB-4 1.3–2.67) proceed to second-line assessment—most commonly LSM by VCTE, with cutoffs of 8 kPa and 12 kPa used to reclassify patients as low, intermediate, or high risk, respectively. Patients with intermediate FIB-4 and LSM below 8 kPa are reclassified as low risk; those with LSM above 12 kPa are classified as high risk and referred to hepatology; and those with LSM between 8 and 12 kPa remain in an intermediate zone warranting specialist evaluation or follow-up reassessment [27,28]. The ELF score and other validated blood-based NITs may serve as alternatives to VCTE in settings where elastography is unavailable (Fig. 1A).

Prognostic Performance of the Two-Step Pathway in Real-World Study

The prognostic validity of the FIB-4/LSM two-step approach

has been rigorously evaluated in the VCTE-prognosis study, a large multicenter longitudinal cohort of 12,950 patients with MASLD across 16 centers in 12 countries, with prospective data collection at 14 sites [32]. Using the two-step algorithm, 81.5%, 4.6%, and 13.9% of patients were classified as low-, intermedi-

ate-, and high-risk, respectively. The 5-year cumulative incidences of liver-related events (LREs) were 0.5%, 1.0%, and 10.8% in the three risk groups—representing a more than 20-fold gradient between the low- and high-risk categories. The sensitivity and NPV of the low-risk classification for excluding LREs at 5 years were 79.0% and 99.5%, respectively, confirming the pathway's effectiveness as a rule-out strategy. Importantly, the study demonstrated that restricting VCTE examination to patients with intermediate FIB-4 (approximately 24% of the cohort) rather than performing LSM on all patients resulted in only marginal loss of prognostic accuracy, with the integrated time-dependent AUC increasing only modestly from 0.840 to 0.866 when VCTE was expanded to the full cohort [9]. This supports the cost-effectiveness rationale of the sequential approach over universal elastography. Furthermore, substituting LSM with composite scores such as Agile 3+, Agile 4, or FAST as the second-line test did not improve prognostication beyond the FIB-4/LSM combination, supporting VCTE as the preferred second-line modality.

Limitations and Unresolved Challenges of the Two-Step Pathway

Despite the overall robustness of the two-step pathway, approximately 30% of patients at tertiary centers exhibit discordant FIB-4 and LSM results, posing interpretive challenges in practice. A secondary analysis of the VCTE-Prognosis cohort categorized patients into four groups based on concordant or discordant FIB-4 (cutoff 1.3) and LSM (cutoff 8 kPa): 49.4% had concordantly low results, 16.7% concordantly high, 10.5% low FIB-4 but elevated LSM, and 23.3% elevated FIB-4 but low LSM [33].

In the biopsy subgroup, patients with low FIB-4 but elevated LSM consistently had more severe histological disease than those with the reverse pattern, suggesting that LSM more frequently reflects true histological severity when the two tests disagree. This was mirrored in clinical outcomes: compared with the concordantly low group, the low FIB-4/high LSM group had approximately 4-fold higher risk for LRE, while elevated FIB-4/low LSM patients had similar risk. These findings

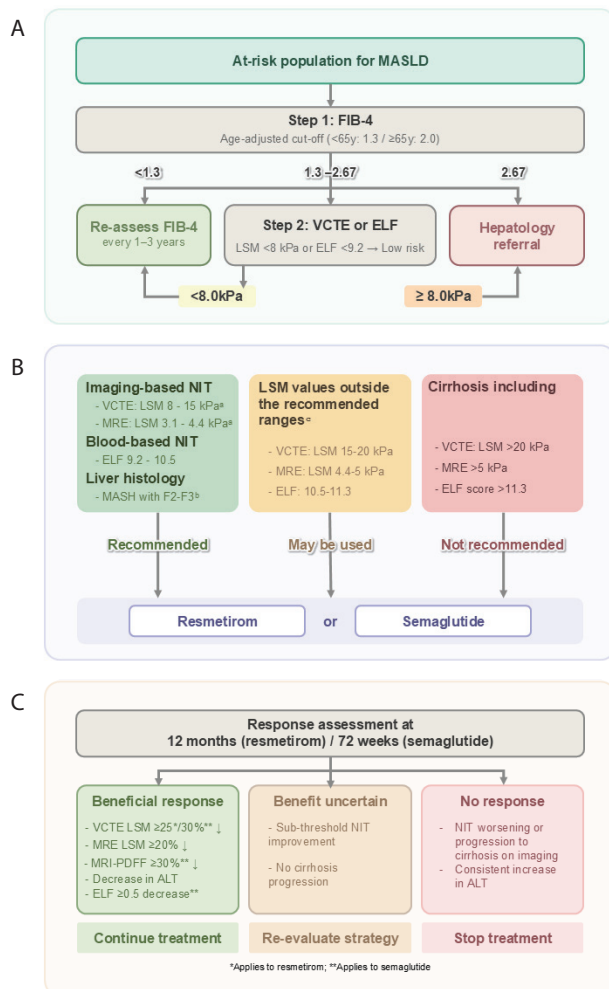


Figure 1. NIT-based clinical algorithm for the management of MASLD. (A) Two-step NIT screening pathway for fibrosis risk stratification in patients with MASLD. FIB-4 serves as the first-line triage test, with VCTE or ELF as second-line confirmation in patients with intermediate FIB-4 values. (B) NIT-based treatment eligibility criteria for resmetirom and semaglutide. Both agents share identical NIT thresholds for patient selection. (C) Drug-specific NIT-based treatment response monitoring timelines and outcome-driven decision algorithm. ALT, alanine aminotransferase; ELF, enhanced liver fibrosis score; FIB-4, Fibrosis-4 index; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; MASH, metabolic dysfunction-associated steatohepatitis; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-derived proton density fat fraction; NIT, non-invasive test; VCTE, vibration-controlled transient elastography.

highlight a key limitation of the two-step pathway: FIB-4, while an effective population-level screening tool, is influenced by age, diabetes, and other metabolic factors that may cause it to misclassify a clinically meaningful proportion of patients. When the two tests disagree, LSM more reliably reflects true fibrosis severity, and a low FIB-4 alone should not be interpreted as definitive reassurance.

In addition, the two-step pathway was designed primarily as a referral and triage tool rather than a diagnostic instrument, and its performance in primary care settings—where the prevalence of advanced fibrosis is lower and longitudinal follow-up less structured—remains incompletely characterized. Serial NIT reassessment at 1–3-year intervals is therefore recommended even for patients initially classified as low risk, and clinical judgement should always complement algorithmic classification when individual patient features suggest a discrepancy between test results and clinical presentation.

NITs AS CLINICAL TRIAL ENDPOINTS IN MASLD

Limitations of Histological Endpoints in Clinical Trials

Liver biopsy has historically served as the cornerstone of efficacy assessment in MASLD clinical trials, providing direct histological evidence of changes in steatosis, necroinflammation, ballooning, and fibrosis stage [34]. The two regulatory-endorsed histological endpoints—resolution of steatohepatitis without worsening of fibrosis, and improvement in fibrosis by at least one stage without worsening of steatohepatitis—have formed the basis for accelerated approval pathways in both the United States and Europe [2]. However, the continued reliance on biopsy-based endpoints imposes substantial constraints on trial feasibility, efficiency, and patient acceptability that have become increasingly difficult to justify in the context of a rapidly expanding therapeutic pipeline.

The procedural limitations of liver biopsy in the trial setting are well established. Biopsy samples typically measure 1.5–2.0 cm in length and 1.2–2.0 mm in diameter, representing less

than 1/50,000th of total liver volume [35]. Given the known spatial heterogeneity of diffuse liver disease, this limited sampling introduces substantial risk of staging inaccuracy: discordance of one or more fibrosis stages between paired biopsy specimens from the same patient has been reported in up to 41% of cases [4]. Inter-observer and intra-observer variability in histological scoring further compounds this problem, particularly for necroinflammatory components such as hepatocyte ballooning and lobular inflammation, which display only fair to moderate inter-rater agreement even among experienced hepatopathologists.

Beyond sampling error, the logistical burden of serial biopsy in large, multinational phase 3 trials is considerable. Biopsy requirements increase screen failure rates, reduce patient willingness to enroll, elevate per-patient costs, and extend the timeline from trial initiation to regulatory submission [36]. In a disease affecting an estimated 30% of the global adult population—the vast majority of whom are managed in primary care settings with limited access to specialist hepatology services—the practicality of biopsy-based screening and monitoring at scale is inherently constrained. These limitations have collectively driven regulatory agencies and the research community toward a paradigm shift in which rigorously validated NITs assume a more central role in trial design, patient enrichment, and endpoint assessment.

Evolution of NITs as Trial Endpoints: from Exploratory to Co-Primary

The incorporation of NITs into MASLD clinical trials has followed a progressive trajectory over the past decade, broadly divided into three evolutionary phases that mirror the growing body of evidence validating NITs as reliable surrogates for histological outcomes.

In the earliest phase, NITs were deployed exclusively as exploratory or secondary endpoints alongside mandatory biopsy-based primary outcomes. The LEAN trial (2015), which evaluated liraglutide in biopsy-proven NASH, exemplified this approach, incorporating ALT as the sole NIT measure while histology remained the unambiguous arbiter of efficacy [37].

Similarly, the FLINT trial (2015) of obeticholic acid included MRI-PDFF as a secondary imaging endpoint; this trial provided the first multicenter demonstration that a relative reduction in MRI-PDFF of $\geq 30\%$ from baseline was associated with a significantly higher rate of histological response, establishing the 30% threshold that has since become the field's canonical NIT-based response criterion [38].

The second phase, spanning approximately 2017 to 2021, was characterized by the systematic integration of MRI-PDFF and LSM as pre-specified secondary or exploratory endpoints in larger phase 2 and phase 3 trials. The CENTAUR trial of selonsertib (2017) formally included MRI-PDFF as a secondary endpoint, while the STELLAR-3 and STELLAR-4 trials (2019) represented the first large phase 3 studies to incorporate both MRI-PDFF and LSM as pre-specified exploratory measures alongside histological primary endpoints [39]. Although selonsertib ultimately failed to demonstrate efficacy, the STELLAR trials generated important data on the correlation between NIT changes and histological outcomes, and demonstrated that LSM could be reliably measured at scale in multinational MASH trials. The AURORA trial of cenicriviroc (terminated 2020) similarly relied on biopsy as its sole primary endpoint—its early termination due to futility underscoring the efficiency costs of biopsy-dependent trial designs [40].

The third and current phase represents the full regulatory legitimization of NITs as co-primary endpoints in pivotal trials. The MAESTRO-NASH trial of resmetirom (2021–2023) marked a watershed moment: MRI-PDFF was incorporated as a co-primary endpoint alongside the two histological primary endpoints, with the FDA accepting this design under its accelerated approval framework [41]. Reductions in MRI-PDFF, LSM, ALT, and fibrosis biomarkers were observed alongside histological improvements, validating the construct that NIT responses could be meaningfully interpreted in the context of biopsy-confirmed efficacy. The ESSENCE trial of semaglutide 2.4 mg (2023–present) continued this trajectory, enrolling 1,197 patients with biopsy-proven MASH and F2–F3 fibrosis, with both histological resolution of steatohepatitis and fibrosis improvement serving as co-primary endpoints supported by a comprehensive NIT assessment framework including LSM,

MRI-PDFF, ALT, and ELF [42].

NITs in Key Pivotal Trials: FLINT to MAESTRO-NASH and ESSENCE

The specific NIT modalities and their roles have evolved considerably across the landmark trials that have shaped the current MASLD drug development landscape. Table 2 summarizes the NIT endpoints and key findings across pivotal trials. In the FLINT trial, MRI-PDFF was measured at baseline and end of treatment in a subset of 78 patients across seven clinical sites, providing the foundational multicenter evidence that the 30% relative MRI-PDFF reduction threshold predicted histological response—defined as NAS improvement of ≥ 2 points without fibrosis worsening—with an odds ratio of 6.98 compared with non-responders [38]. This seminal observation established MRI-PDFF as the benchmark steatosis biomarker for subsequent MASH trials.

In the STELLAR trials, baseline NIT data demonstrated that the sequential FIB-4 and LSM algorithm accurately identified patients with advanced fibrosis ($F \geq 3$) in the phase 3 setting, with LSM by VCTE achieving an AUROC of 0.85 for advanced fibrosis—outperforming serum-based scores—and providing the first large-scale validation of elastography-based enrichment strategies in pivotal MASH trials [39].

The MAESTRO-NASH trial represented the most comprehensive NIT assessment undertaken in a pivotal MASH study to date. Among 966 patients with biopsy-proven MASH and F1b–F3 fibrosis, resmetirom produced significant reductions in MRI-PDFF (median -32.9% with 100 mg vs. -5.6% with placebo), LSM by VCTE, ALT, LDL cholesterol, and multiple fibrosis biomarkers [41]. Importantly, the trial demonstrated that changes in NITs at 3 months were not reliably predictive of 52-week histological response, establishing that NIT-based treatment assessment should be deferred to the 12-month timepoint—a finding with direct implications for real-world monitoring protocols [43].

The phase 3 ESSENCE trial of semaglutide enrolled the largest biopsy-confirmed MASH cohort to date, and at the planned 72-week interim analysis, the trial achieved both histological

Table 2. Key clinical trials and the evolving role of NITs as endpoints in MASLD/MASH drug development

Trial (year)	Drug (mechanism)	Phase	N	Duration	Primary histologic endpoints	NIT role in trial	Key NIT findings
Era 1: Histology-only endpoints—NITs as exploratory measures (2014–2018)							
LEAN (2016)	Liraglutide (GLP-1RA)	2	52	48 wk	MASH resolution without fibrosis worsening	Exploratory only (no pre-specified NIT endpoint)	No formal NIT analysis; liver histology sole basis for efficacy assessment; highlighted need for non-invasive monitoring tools in future trials
FLINT (2015)	Obeticholic acid (FXR agonist)	2b	283	72 wk	≥2-point NAS improvement without fibrosis worsening	Post-hoc exploratory: MRI-PDFF, ALT	Post-hoc analysis showed MRI-PDFF reduced by 32% vs. 12% (placebo); suggested MRI-PDFF as candidate surrogate—first major NASH trial to report MRI-PDFF data; ALT reduction correlated with histological response
CENTAUR (2017)	Selonsertib (ASK1 inhibitor)	2	72	24 wk	≥1-stage fibrosis improvement without MASH worsening	Exploratory: MRE, MRI-PDFF, ELF	MRE and MRI-PDFF changes examined post-hoc; no significant correlation between NIT change and histological fibrosis improvement; highlighted limitations of NIT surrogacy in early-phase trials
REGENERATE (2019)	Obeticholic acid (FXR agonist)	3	931	18 mo (interim)	≥1-stage fibrosis improvement; MASH resolution (co-primary, 18-mo interim)	Pre-specified secondary: VCTE-LSM, ELF, ALT	VCTE-LSM and ELF reduced significantly with OCA vs. placebo; NIT changes did not reliably predict individual histological response; reinforced that NIT improvements can occur without histological endpoint achievement
Era 2: NITs as pre-specified secondary endpoints—emerging surrogate evidence (2019–2022)							
STELLAR-3/STELLAR-4 (2019)	Selonsertib (ASK1 inhibitor)	3	877	48 wk	≥1-stage fibrosis improvement (STELLAR-3: F3; STELLAR-4: F4/cirrhosis)	Pre-specified secondary: MRE, MRI-PDFF, VCTE, ELF	Both trials failed primary histologic endpoints; MRE and VCTE-LSM changes were modest and inconsistent; first phase 3 trial to include MRE as a pre-specified NIT endpoint; failure highlighted disconnect between NIT and histology changes in ineffective therapies
AURORA (2020, terminated)	Centiviroc (CCR2/CCR5 antagonist)	3	~1,300	~2 yr (terminated)	≥1-stage fibrosis improvement without MASH worsening (terminated early—futility)	Pre-specified secondary: VCTE, ELF, ALT	Trial terminated due to futility at interim analysis; no NIT data published; reinforced importance of early NIT-based futility assessment in long-term trials
FASCINATE-2 (2022)	Lanifibranor (pan-PPAR agonist)	2b	247	24 wk	SAF-A score improvement (steatosis+activity+fibrosis)	Pre-specified secondary: MRI-PDFF, VCTE-LSM, ELF	Significant reductions in MRI-PDFF (36% vs. ~13%), VCTE-LSM, and ELF with lanifibranor; NIT changes broadly correlated with histological activity reduction; supported MRI-PDFF as a responsive intermediate endpoint for anti-steatotic agents
Era 3: NITs as co-primary or regulatory endpoints—pivotal trial validation (2023–2025)							
MAESTRO-NASH (2024)	Resmetrirom (THRβ agonist)	3	966	52 wk	(1) MASH resolution without fibrosis worsening (2) ≥1-stage fibrosis improvement without MASH worsening (both achieved: 26–30% vs. 10%; 24–26% vs. 14%)	Co-primary endpoint: MRI-PDFF ≥30% reduction Pre-specified secondary: VCTE-LSM, ELF, ALT, lipids	MRI-PDFF ≥30% reduction achieved in 66–79% vs. 32–35% placebo—first phase 3 trial with NIT as regulatory co-primary endpoint; MRI-PDFF <30% reduction strongly associated with lack of histological response (response rate ~10%, equivalent to placebo); VCTE-LSM ≥25% reduction in 49% vs. 26% placebo; ELF and ALT improvements confirmed; established MRI-PDFF as validated surrogate in resmetrirom FDA approval

Table 2. Continued

Trial (year)	Drug (mechanism)	Phase	N	Duration	Primary histologic endpoints	NIT role in trial	Key NIT findings
ESSENCE (2025)	Semaaglutide 2.4 mg/wk (GLP-1RA)	3	1,197	72 wk (interim: 800 pts)	(1) MASH resolution without fibrosis worsening: 62.9% vs. 34.3% ($p < 0.001$) (2) ≥ 1 -stage fibrosis improvement without MASH worsening: 36.8% vs. 22.4% ($p < 0.001$) (accelerated FDA approval August 2025)	Pre-specified secondary: VCTE-LSM, CAP, ELF, ALT Response thresholds: VCTE $\geq 30\%$, MRE $\geq 20\%$, ELF ≥ 0.5 , ALT ≥ 17 U/L or $\geq 20\%$	VCTE-LSM $\geq 30\%$ reduction: 52% vs. 30.3% placebo; ELF ≥ 0.5 reduction: 55.8% vs. 25.5% placebo; ALT reduced -52% vs. -22%; CAP reduced -43 vs. -13 dB/m; NIT changes greater in histological responders than non-responders; established VCTE and ELF response thresholds as basis for AASLD semaaglutide monitoring guidance; MRI-PDFF not formally assessed in ESSENCE
Era 4: Emerging NIT-driven trial designs—NIT-based enrolment and endpoints (2024–ongoing)							
SURPASS-NASH (ongoing)	Tirzepatide (GLP-1/GIP dual agonist)	3	~1,100	52 wk	MASH resolution; ≥ 1 -stage fibrosis improvement	NIT-based enrolment: MRI-PDFF $\geq 8\%$ VCTE ≥ 8 kPa or ELF ≥ 92 Co-primary NIT endpoint: MRI-PDFF $\geq 30\%$ reduction	Phase 2 data (NEJM 2024): MRI-PDFF -55% vs. -1%; MASH resolution 62% vs. 13%; fibrosis improvement 51% vs. 30%; first pivotal MASH trial using NIT-based inclusion criteria without mandatory screening biopsy at enrolment; MRI-PDFF as co-primary regulatory endpoint
REDEFINE-1 (ongoing)	Retatrutide (GLP-1/GIP/glucagon triple agonist)	3	~1,800	48 wk	MASH resolution; ≥ 1 -stage fibrosis improvement	NIT-based enrolment: MRI-PDFF $\geq 8\%$ MRE ≥ 2.9 kPa NIT co-primary: MRI-PDFF $\geq 30\%$ reduction	Phase 2 data: weight loss -24.2% at 48 wk; MRI-PDFF -83% relative reduction; first phase 3 trial with MRE as mandatory enrolment criterion and NIT-only co-primary histologic/imaging composite endpoint; represents most advanced NIT integration to date
ALPINE (ongoing)	Survodutide (GLP-1/glucagon dual agonist)	3	~1,200	48 wk	MASH resolution; ≥ 1 -stage fibrosis improvement	NIT-based enrolment: MRI-PDFF $\geq 8\%$ VCTE ≥ 8 kPa or ELF ≥ 92 NIT secondary: MRI-PDFF, VCTE, ELF	Phase 2 data: MRI-PDFF -68% relative reduction; MASH resolution 83% vs. 18%; fibrosis improvement 63% vs. 22%; enrolment strategy mirrors SURPASS-NASH—biopsy-free screening using NIT-based criteria; NIT outcomes pre-specified as key secondary endpoints for regulatory submission

ALT, alanine aminotransferase; CAP, controlled attenuation parameter; ELF, enhanced liver fibrosis score; FXR, farnesoid X receptor; GIP, glucose-dependent insulinotropic polypeptide; GLP-1RA, glucagon-like peptide-1 receptor agonist; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; MASH, metabolic dysfunction-associated steatohepatitis; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging-derived proton density fat fraction; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD activity score; NIT, non-invasive test; OCA, obeticholic acid; PPAR, peroxisome proliferator-activated receptor; SAF, steatosis-activity-fibrosis; THRB, thyroid hormone receptor beta; VCTE, vibration-controlled transient elastography.

primary endpoints [42]. Semaglutide produced consistent 30–40% reductions in ALT, AST, and GGT from baseline relative to placebo, alongside significant reductions in LSM and ELF score—confirming that GLP-1 receptor agonist-mediated histological improvements are broadly reflected across multiple NIT domains.

Regulatory Framework: FDA and EMA Guidance on NIT-Based Endpoints

The regulatory landscape governing NIT-based endpoints in MASLD trials has evolved substantially over the past decade, with both the FDA and the European Medicines Agency (EMA) progressively formalizing the conditions under which NITs can support accelerated or conditional approval.

The FDA's 2018 guidance document for noncirrhotic NASH with liver fibrosis established the foundational framework, endorsing histological endpoints—specifically, MASH resolution without fibrosis worsening, and fibrosis improvement by ≥ 1 stage without worsening of steatohepatitis—as the accepted surrogates for accelerated approval, while explicitly encouraging sponsors to identify noninvasive imaging biomarkers capable of replacing biopsy in future trial designs [44]. The accelerated approval of resmetirom in March 2024 on the basis of these histological surrogates, supported by NIT co-primary data, operationalized this framework for the first time and demonstrated that regulators were prepared to accept a trial design in which NITs contributed substantively to the efficacy evidence package [41].

Key points of agreement between the FDA and EMA include the requirement for placebo-controlled designs targeting patients with NAS ≥ 4 and NASH Clinical Research Network (CRN) fibrosis stage 2 or 3, stable doses of comorbidity treatments, and stratification by type 2 diabetes status [45]. Both agencies endorse the use of composite clinical outcome endpoints—including all-cause mortality, cirrhosis, hepatic decompensation, and MELD score progression—as the ultimate criteria for full market approval, with histological improvement serving as the intermediate surrogate for accelerated or conditional approval pathways.

Notable divergences between the two agencies reflect differing risk tolerance and evidentiary standards. The FDA requires resolution of steatohepatitis or improvement in fibrosis, with the sponsor having flexibility to propose either endpoint based on mechanism of action, whereas the EMA mandates co-primary endpoints requiring demonstration of both steatohepatitis resolution and fibrosis improvement—a more stringent requirement that increases statistical demands and sample size [45]. With respect to NITs specifically, both agencies currently classify imaging-based surrogates such as MRE-derived LSM and MRI-PDFF as exploratory or supportive—not yet approved as stand-alone efficacy endpoints—while acknowledging that changes of $\geq 19\%$ in MRE-LSM and $\geq 30\%$ in MRI-PDFF are associated with fibrosis progression or regression and histological response, respectively [45,46]. The ELF score has received FDA de novo clearance for risk stratification but has not yet achieved status as a regulatory efficacy surrogate.

Looking ahead, the qualification of NITs as stand-alone primary endpoints—without mandatory confirmatory biopsy—represents the next frontier in MASLD trial design. This shift is already underway: several emerging agents, including tirzepatide and retatrutid have incorporated MRI-PDFF and LSM as primary or key secondary endpoints in phase 3 trials (NCT07165028) without requiring paired liver biopsies, reflecting growing confidence in NIT-based efficacy assessment at earlier stages of drug development. Achieving full regulatory acceptance of NIT-only trial designs will nonetheless require prospective demonstration that NIT-defined response categories predict long-term liver-related outcomes with sufficient sensitivity, specificity, and reproducibility across diverse populations and treatment mechanisms—a goal that the ongoing outcomes phases of the MAESTRO and ESSENCE programs, as well as emerging platform trial initiatives, are specifically designed to address.

NIT-BASED TREATMENT GUIDANCE IN MASLD: RESMETIROM AND SEMAGLUTIDE

From FDA Approval to NIT-Based Practice: Treatment Landscape Overview

The approval of resmetirom and semaglutide has fundamentally transformed the therapeutic landscape of MASLD, marking the first time that pharmacological interventions have demonstrated sufficient histological efficacy to support regulatory approval in this disease.

Resmetirom (Rezdiffra; Madrigal Pharmaceuticals), a once-daily oral liver-directed thyroid hormone receptor- β (THR β) selective agonist, received accelerated approval from the FDA in March 2024 for the treatment of adults with noncirrhotic MASH and moderate-to-advanced liver fibrosis consistent with stages F2–F3 [41]. Approval was based on the pivotal phase 3 MAE-STRO-NASH trial, in which resmetirom achieved both histological primary endpoints at 52 weeks: resolution of steatohepatitis without worsening of fibrosis (26–30% vs. 10% placebo) and improvement in fibrosis by ≥ 1 stage without worsening of steatohepatitis (24–26% vs. 14% placebo). Semaglutide (Wegovy; Novo Nordisk), a once-weekly subcutaneous glucagon-like peptide-1 receptor agonist (GLP-1RA), received accelerated FDA approval in August 2025 for the same indication—noncirrhotic MASH with moderate-to-advanced fibrosis (F2–F3)—based on the planned 72-week interim analysis of the phase 3 ESSENCE trial, in which semaglutide 2.4 mg weekly achieved both co-primary histological endpoints: resolution of steatohepatitis without worsening of fibrosis (62.9% vs. 34.3% placebo) and improvement in fibrosis without worsening of steatohepatitis (36.8% vs. 22.4% placebo) [42]. Final approval for both agents is contingent upon verification of long-term clinical benefit in their respective ongoing confirmatory outcome trials.

Both pivotal trials enrolled patients with biopsy-confirmed MASH and fibrosis stage 2 or 3 according to the NASH CRN classification, without cirrhosis. The target treatment population is therefore specifically defined as adults with at-risk MASH—that is, histologically active MASH (NAFLD activity score ≥ 4 with at

least 1 point each for steatosis, lobular inflammation, and hepatocyte ballooning) accompanied by clinically significant fibrosis (F2–F3). In clinical practice, however, the definitive histological characterization of at-risk MASH can only be achieved through liver biopsy, an invasive procedure that is neither feasible nor appropriate as a routine prerequisite for treatment initiation in most patients. Recognizing this, both the FDA prescribing information for resmetirom and semaglutide and the updated American Association for the Study of Liver Diseases (AASLD) practice guidance documents explicitly state that liver biopsy is not required for treatment eligibility, and that patient selection should instead be based on NITs consistent with MASH and F2–F3 fibrosis [8,9]. This NIT-based approach reflects the broader recognition that steatosis can be reliably estimated by imaging, and that fibrosis stage can be approximated with reasonable accuracy using validated serum-based and imaging-based NITs that are already widely deployed in clinical practice.

NIT-Based Patient Selection for MASH Treatment

The selection of candidates for resmetirom or semaglutide therapy in clinical practice is guided by a sequential NIT-based algorithm that identifies patients with MASH and clinically significant fibrosis (F2–F3) while excluding those with cirrhosis or portal hypertension, in whom neither agent is currently approved. Consistent with international guidelines, both the AASLD resmetirom and semaglutide practice guidance documents endorse a sequential NIT strategy as the starting point for identifying treatment candidates [8,9].

Imaging-based NIT criteria

Imaging-based NITs—particularly VCTE-derived LSM and MRE—are the preferred modalities for fibrosis staging in the treatment selection context, given their superior diagnostic accuracy compared with serum-based scores alone (Fig. 1B). For both resmetirom and semaglutide, the recommended LSM range by VCTE is 8–15 kPa, supported by the distribution of LSM values among trial participants with biopsy-confirmed F2–F3 in MAE-STRO-NASH and ESSENCE, where the median VCTE LSM was approximately 12 kPa and the interquartile range 10–15 kPa [8,9].

This range captured approximately 50% of trial participants with histological F2–F3. For patients with LSM between 15 and 20 kPa—present in up to 19% of ESSENCE participants—an individualized decision to treat should incorporate additional confirmatory evidence to exclude cirrhosis, including cross-sectional imaging to rule out nodular liver contour and signs of portal hypertension, or a platelet count below $150,000/\text{mm}^3$ [9]. An LSM exceeding 20 kPa by VCTE, MRE above 5.0 kPa, or clinical or radiological signs of portal hypertension are contraindications to treatment with both agents [8,9].

Where VCTE is unavailable or technically limited—particularly in patients with central adiposity ($\text{BMI} > 35 \text{ kg/m}^2$) or narrow intercostal spaces—MRE offers superior accuracy and is unaffected by severe obesity. The recommended MRE-based LSM range for treatment eligibility is 3.1–4.4 kPa for both agents, extrapolated from prior validation studies and AASLD noninvasive liver disease assessment (NILDA) guidelines, as MRE was not formally evaluated in either pivotal trial [8,9]. Values between 4.4 and 5.0 kPa warrant individualized cirrhosis exclusion as described above.

Blood-based NIT criteria

In settings where imaging-based NITs are unavailable, the ELF score is the recommended blood-based alternative. An ELF score of 9.2–10.5 is endorsed by both guidance documents as the range consistent with F2–F3 fibrosis, aligning with the interquartile range observed in both MAESTRO–NASH (median 9.7, interquartile range [IQR] 9.2–10.4) and ESSENCE (median 9.9, IQR 9.2–10.5) [41,42]. Patients with ELF values between 10.5 and 11.3 may be considered for treatment following individualized cirrhosis exclusion, while an ELF score exceeding 11.3 is a contraindication to both agents. The resmetirom guidance notes that imaging-based NITs are preferred over blood-based alternatives for treatment selection, and that the ELF cutoff range is derived from MAESTRO–NASH trial data rather than formal AASLD NILDA guideline thresholds [8].

NIT-Based Treatment Response Monitoring

Following treatment initiation, serial NIT assessment serves as the primary tool for evaluating therapeutic response in clinical

practice, given that repeat liver biopsy is neither feasible nor recommended as a routine monitoring strategy. Both the AASLD resmetirom and semaglutide guidance documents provide NIT-based frameworks for response assessment, classifying patients into three categories—beneficial response, uncertain benefit, and no response—to guide continuation or discontinuation decisions [8,9] (Fig. 1C).

Resmetirom: 12-month response assessment

For resmetirom, the AASLD guidance recommends a structured monitoring schedule with hepatic function panel testing at 3, 6, and 12 months, with imaging-based NIT reassessment at 12 months using the same modality applied at baseline [8]. The 12-month timepoint is the primary response assessment window, reflecting the 52-week duration of the MAESTRO–NASH trial [41]. A beneficial response is defined as LSM improvement of $\geq 25\%$ from baseline by VCTE or $\geq 20\%$ by MRE, combined with normalization or significant improvement in ALT in patients with elevated baseline values. These LSM thresholds are based on assay characteristics and the observation that a 20% or greater reduction in MRE-measured stiffness yielded response rates similar to histological fibrosis improvement in MAESTRO–NASH, while a 25% VCTE LSM decrease was used as the practical threshold given the absence of resmetirom-specific validation data [8]. Non-response is defined by worsening NILDA or a consistent increase in ALT, and should prompt consideration of treatment discontinuation. For patients with neither clear improvement nor worsening—a substantial proportion given that the primary histological endpoints were achieved in only 26–30% of resmetirom-treated participants—an individualized decision incorporating baseline fibrosis severity, comorbidity profile, and tolerance should guide whether to continue therapy pending longer-term outcome data.

MRI–PDFF provides additional response information where available: a reduction of $\geq 30\%$ from baseline at 12 months was strongly associated with histological response in MAESTRO–NASH, while failure to achieve this threshold was associated with response rates no better than placebo [41]. However, given limited availability in routine practice, MRI–PDFF is not a prerequisite for monitoring and should not be used as a standalone criterion for treatment continuation or discontinuation. In addition,

changes in CAP as measured by VCTE were not predictive of histological response in MAESTRO-NASH and should not be used to assess futility of resmetirom treatment [8].

Semaglutide: 72-week response assessment

For semaglutide, the primary response assessment is performed at 72 weeks, corresponding to the planned interim analysis timepoint of the ESSENCE trial [42]. Routine hepatic function panel testing is recommended only as clinically indicated during treatment—reflecting semaglutide’s favorable hepatic safety profile in ESSENCE, where no participants discontinued treatment due to liver enzyme elevations—rather than at fixed intervals as with resmetirom. NIT-based response criteria at 72 weeks include LSM reduction of $\geq 30\%$ from baseline by VCTE, $\geq 20\%$ by MRE, ALT decrease of ≥ 17 U/L or $\geq 20\%$ from baseline, ELF reduction of ≥ 0.5 , and MRI-PDFF decrease of $\geq 30\%$ [9]. A beneficial response meeting these thresholds supports continuation of semaglutide, while NIT worsening or progression toward cirrhosis constitutes non-response and should prompt treatment discontinuation. As with resmetirom, sub-threshold improvement without evidence of progression represents an area of clinical uncertainty requiring individualized evaluation.

It should be noted that the ESSENCE trial data available at the time of guidance publication were limited regarding the ability of individual NITs to predict histological response at the patient level. Aggregate data from ESSENCE demonstrated that reductions in ALT, CAP, LSM, and ELF from baseline to 72 weeks were greater in semaglutide-treated participants who achieved MASH resolution compared with non-responders, with 52% of the semaglutide group achieving $\geq 30\%$ LSM reduction versus 30.3% in the placebo group, and 55.8% achieving ≥ 0.5 ELF reduction versus 25.5% [9]. These group-level differences support the use of these thresholds as response benchmarks, while acknowledging that no single NIT reliably predicts individual histological outcomes.

CONCLUSION

NITs now constitute an indispensable component of the

MASLD management continuum, fulfilling sequential roles in population-level fibrosis triage, clinical trial endpoint validation, and individualized treatment monitoring following the first regulatory approvals of MASH-specific pharmacotherapies. The two-step pathway combining FIB-4 with LSM by VCTE has demonstrated robust prognostic validity in large multinational cohorts, and NIT-based algorithms formalized in the AASLD practice guidance documents for resmetirom and semaglutide have operationalized a treat-and-monitor paradigm that largely circumvents the need for repeat liver biopsy in routine clinical practice. Nevertheless, important limitations persist—including FIB-4/LSM discordance in approximately 30% of patients and the continued classification of imaging-based surrogates as exploratory rather than stand-alone regulatory endpoints—underscoring that prospective validation of NIT-defined response categories as reliable predictors of long-term liver-related outcomes remains the next critical frontier in MASLD drug development and clinical management.

ACKNOWLEDGEMENTS

None.

FUND

None.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

AUTHOR CONTRIBUTIONS

H.A.L.: Conceptualization, Literature search, Data curation, Writing — original draft, Writing — review & editing, Visualization, Project administration, Supervision; Y.Y.C.: Literature search, Writing — review & editing; H.J.K.: Literature search, Writing — review & editing; All authors have read and approved the final version of the manuscript.

REFERENCES

1. Younossi ZM, Kalligeros M, Henry L. Epidemiology of metabolic dysfunction-associated steatotic liver disease. *Clin Mol Hepatol* 2025;31(Suppl):S32-S50.
2. Kim GA, Jang H, Kim MY, et al. Multicenter prospective cohort study on clinical outcomes and fibrosis patterns in biopsy-proven steatotic liver disease subtypes. *Gastroenterology* 2026 January 29 [Epub]. <https://doi.org/10.1053/j.gastro.2026.01.007>.
3. Yun B, Park H, Ahn SH, Oh J, Kim BK, Yoon JH. Liver cancer risk across metabolic dysfunction-associated steatotic liver disease and/or alcohol: a nationwide study. *Am J Gastroenterol* 2025;120:410-419.
4. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 2005;128:1898-1906.
5. Allen AM, Therneau TM, Ahmed OT, et al. Clinical course of non-alcoholic fatty liver disease and the implications for clinical trial design. *J Hepatol* 2022;77:1237-1245.
6. Diaz LA, Thiele M, Louvet A, et al. Clinical trial design, biomarkers and end points in metabolic and alcohol-related liver disease. *Nat Rev Gastroenterol Hepatol* 2025;22:866-884.
7. Kim MN, Han JW, An J, et al. KASL clinical practice guidelines for noninvasive tests to assess liver fibrosis in chronic liver disease. *Clin Mol Hepatol* 2024;30(Suppl):S5-S105.
8. Chen VL, Morgan TR, Rotman Y, et al. Resmetirom therapy for metabolic dysfunction-associated steatotic liver disease: October 2024 updates to AASLD Practice Guidance. *Hepatology* 2025;81:312-320.
9. Bansal MB, Patton H, Morgan TR, Carr RM, Dranoff JA, Allen AM. Semaglutide therapy for metabolic dysfunction-associated steatohepatitis: November 2025 updates to AASLD Practice Guidance. *Hepatology* 2026;83:1326-1340.
10. Lee HH, Lee HA, Kim EJ, et al. Metabolic dysfunction-associated steatotic liver disease and risk of cardiovascular disease. *Gut* 2024;73:533-540.
11. Bose K, Mudgil P, Gupta P, et al. Deep learning for non-invasive detection of steatosis and fibrosis in MASLD: a multicenter study with a new fibroscan-labelled ultrasound dataset. *Abdom Radiol (NY)* 2026;51:2906-2918.
12. Vali Y, Lee J, Boursier J, et al. Biomarkers for staging fibrosis and non-alcoholic steatohepatitis in non-alcoholic fatty liver disease (the LITMUS project): a comparative diagnostic accuracy study. *Lancet Gastroenterol Hepatol* 2023;8:714-725.
13. McPherson S, Hardy T, Dufour JF, et al. Age as a confounding factor for the accurate non-invasive diagnosis of advanced NAFLD fibrosis. *Am J Gastroenterol* 2017;112:740-751.
14. Vali Y, Lee J, Boursier J, et al. Enhanced liver fibrosis test for the non-invasive diagnosis of fibrosis in patients with NAFLD: a systematic review and meta-analysis. *J Hepatol* 2020;73:252-262.
15. Daniels SJ, Leeming DJ, Eslam M, et al. ADAPT: an algorithm incorporating PRO-C3 accurately identifies patients with NAFLD and advanced fibrosis. *Hepatology* 2019;69:1075-1086.
16. Eslam M, Wong GL, Hashem AM, et al. A sequential algorithm combining ADAPT and liver stiffness can stage metabolic-associated fatty liver disease in hospital-based and primary care patients. *Am J Gastroenterol* 2021;116:984-993.
17. Harrison SA, Ratziu V, Magnanensi J, et al. NIS2+™, an optimisation of the blood-based biomarker NIS4® technology for the detection of at-risk NASH: a prospective derivation and validation study. *J Hepatol* 2023;79:758-767.
18. Harrison SA, Ratziu V, Boursier J, et al. A blood-based biomarker panel (NIS4) for non-invasive diagnosis of non-alcoholic steatohepatitis and liver fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol* 2020;5:970-985.
19. Canivet CM, Zheng MH, Qadri S, et al. Validation of the blood test MACK-3 for the noninvasive diagnosis of fibrotic nonalcoholic steatohepatitis: an international study with 1924 patients. *Clin Gastroenterol Hepatol* 2023;21:3097-3106.e10.
20. Tavaglione F, Jamialahmadi O, De Vincentis A, et al. Development and validation of a score for fibrotic nonalcoholic steatohepatitis. *Clin Gastroenterol Hepatol* 2023;21:1523-1532.e1.
21. Mózes 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.
22. 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.
23. Liang JX, Ampuero J, Niu H, et al. An individual patient data meta-analysis to determine cut-offs for and confounders of NAFLD-fibrosis staging with magnetic resonance elastography. *J Hepatol* 2023;79:592-604.
24. Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, quantitative assessment of liver fat by MRI-PDFF as an endpoint in NASH trials. *Hepatology* 2018;68:763-772.
25. Eddowes PJ, Sasso M, Allison M, et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with non-alcoholic fatty liver disease. *Gastroenterology* 2019;156:1717-1730.
26. Younossi ZM, Paik JM, Henry L, Stepanova M, Nader F. Pharmacoeconomic assessment of screening strategies for high-risk MASLD in primary care. *Liver Int* 2025;45:e16119.
27. Kanwal F, Shubbrook JH, Adams LA, et al. Clinical care path-

- way for the risk stratification and management of patients with nonalcoholic fatty liver disease. *Gastroenterology* 2021;161:1657-1669.
28. 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.
 29. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). *J Hepatol* 2024;81:492-542.
 30. Sohn W, Lee YS, Kim SS, et al. KASL clinical practice guidelines for the management of metabolic dysfunction-associated steatotic liver disease 2025. *Clin Mol Hepatol* 2025;31(Suppl):S1-S31.
 31. Graupera I, Thiele M, Serra-Burriel M, et al. Low accuracy of FIB-4 and NAFLD fibrosis scores for screening for liver fibrosis in the population. *Clin Gastroenterol Hepatol* 2022;20:2567-2576.e6.
 32. Yip TC, Lee HW, Lin H, et al. Prognostic performance of the two-step clinical care pathway in metabolic dysfunction-associated steatotic liver disease. *J Hepatol* 2025;83:304-314.
 33. Rabbat J, Yang B, Lee HW, et al. Histological severity and hepatic outcomes in patients with metabolic dysfunction-associated steatotic liver disease and discrepant FIB-4 and liver stiffness measurement. *Clin Mol Hepatol* 2026;32:289-304.
 34. Brunt EM, Kleiner DE, Carpenter DH, et al. NAFLD: reporting histologic findings in clinical practice. *Hepatology* 2021;73:2028-2038.
 35. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. Liver biopsy. *Hepatology* 2009;49:1017-1044.
 36. Sanyal AJ, Friedman SL, McCullough AJ, Dimick-Santos L. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American Association for the Study of Liver Diseases-U.S. Food and Drug Administration Joint Workshop. *Hepatology* 2015;61:1392-1405.
 37. Armstrong MJ, Gaunt P, Aithal GP, et al. Liraglutide safety and efficacy in patients with non-alcoholic steatohepatitis (LEAN): a multicentre, double-blind, randomised, placebo-controlled phase 2 study. *Lancet* 2016;387:679-690.
 38. Loomba R, Neuschwander-Tetri BA, Sanyal A, et al. Multicenter validation of association between decline in MRI-PDFF and histologic response in NASH. *Hepatology* 2020;72:1219-1229.
 39. 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.
 40. 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.
 41. Harrison SA, Bedossa P, Guy CD, et al. A phase 3, randomized, controlled trial of resmetirom in NASH with liver fibrosis. *N Engl J Med* 2024;390:497-509.
 42. Sanyal AJ, Newsome PN, Kliers I, et al. Phase 3 Trial of semaglutide in metabolic dysfunction-associated steatohepatitis. *N Engl J Med* 2025;392:2089-2099.
 43. Nouredin M, Charlton MR, Harrison SA, et al. Expert panel recommendations: practical clinical applications for initiating and monitoring resmetirom in patients with MASH/NASH and moderate to noncirrhotic advanced fibrosis. *Clin Gastroenterol Hepatol* 2024;22:2367-2377.
 44. Center for Drug Evaluation and Research. Noncirrhotic nonalcoholic steatohepatitis with liver fibrosis: developing drugs for treatment. Silver Spring (MD): U.S. Department of Health and Human Services Food and Drug Administration, 2018.
 45. Loomba R, Ratzu V, Harrison SA. Expert panel review to compare FDA and EMA guidance on drug development and endpoints in nonalcoholic steatohepatitis. *Gastroenterology* 2022;162:680-688.
 46. Ajmera V, Kim BK, Yang K, et al. Liver stiffness on magnetic resonance elastography and the MEFIB index and liver-related outcomes in nonalcoholic fatty liver disease: a systematic review and meta-analysis of individual participants. *Gastroenterology* 2022;163:1079-1089.e5.