Noninvasive imaging biomarkers for liver fibrosis in nonalcoholic fatty liver disease: current and future

Running title: Imaging-based noninvasive tests for liver fibrosis in NAFLD

Jung Hwan Yu¹, Han Ah Lee², and Seung Up Kim³,⁴

¹Department of Internal Medicine, Inha University Hospital and School of Medicine, Incheon, South Korea; ²Department of Internal Medicine, College of Medicine, Ewha Womans University, Seoul, South Korea; ³Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea; ⁴Yonsei Liver Center, Severance Hospital, Seoul, South Korea

Jung Hwan Yu and Han Ah Lee equally contributed to this work.

Correspondence

Seung Up Kim, MD, PhD
Department of Internal Medicine, Yonsei University College of Medicine, 250 Seongsanno, Seodaemun–gu, Seoul 120–752, Korea.
E-mail: ksukorea@yuhs.ac; Tel.: +82-2-2228-1944; Fax: +82-2-393-6884
ORCID ID: 0000-0002-9658-8050
**Conflicts of Interest**

Seung Up Kim served as an advisory committee member for Gilead Sciences, GSK, Bayer, and Eisai. He is a speaker for Gilead Sciences, GSK, Bayer, Eisai, Abbive, EchoSens, MSD, and Bristol-Myers Squibb. He also received a research grant from Abbive and Bristol-Myers Squibb. The other authors declare that they have no conflicts of interest.

**Funding**

This research was supported in part by a National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (grant no. 2018R1A5A2025286 and 2022R1I1A1A01065244), the Technology Innovation Program (or Industrial Strategic Technology Development Program-Bioindustry Strategic Technology Development Program, grant no. 20013712) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea).

**Author Contributions**

Conception and design of the study: J.H. Yu, H.A. Lee, and S.U. Kim; Drafting or revision of the manuscript: J.H. Yu, H.A. Lee, and S.U. Kim; Approval of the final version of the manuscript: J.H. Yu, H.A. Lee, and S.U. Kim
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAFLD</td>
<td>nonalcoholic fatty liver disease</td>
</tr>
<tr>
<td>NASH</td>
<td>nonalcoholic steatohepatitis</td>
</tr>
<tr>
<td>HCC</td>
<td>hepatocellular carcinoma</td>
</tr>
<tr>
<td>NIT</td>
<td>noninvasive test</td>
</tr>
<tr>
<td>VCTE</td>
<td>vibration-controlled transient elastography</td>
</tr>
<tr>
<td>IQR</td>
<td>interquartile range</td>
</tr>
<tr>
<td>SWE</td>
<td>shear wave elastography</td>
</tr>
<tr>
<td>ARFI</td>
<td>acoustic radiation force impulse imaging</td>
</tr>
<tr>
<td>MRE</td>
<td>magnetic resonance elastography</td>
</tr>
<tr>
<td>AI</td>
<td>artificial intelligence</td>
</tr>
<tr>
<td>pSWE</td>
<td>point shear wave elastography</td>
</tr>
<tr>
<td>2D-SWE</td>
<td>two-dimensional shear wave elastography</td>
</tr>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>LS</td>
<td>liver stiffness</td>
</tr>
<tr>
<td>kPa</td>
<td>kilopascals</td>
</tr>
<tr>
<td>AUROC</td>
<td>area under the receiver operating characteristic curve</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
</tr>
<tr>
<td>PPV</td>
<td>positive predictive value</td>
</tr>
<tr>
<td>HR</td>
<td>hazard ratio</td>
</tr>
<tr>
<td>FIB-4 index</td>
<td>fibrosis-4 index</td>
</tr>
<tr>
<td>CNN</td>
<td>convolutional neural networks</td>
</tr>
<tr>
<td>3D</td>
<td>three-dimensional</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance imaging</td>
</tr>
</tbody>
</table>
**Abstract**

Nonalcoholic fatty liver disease (NAFLD) is increasingly prevalent worldwide and becoming a major cause of liver disease-related morbidity and mortality. The presence of liver fibrosis in patients with NAFLD is closely related to prognosis, including the development of hepatocellular carcinoma and other complications of cirrhosis. Therefore, assessment of the presence of significant or advanced liver fibrosis is crucial. Although liver biopsy has been considered the “gold standard” method for evaluating the degree of liver fibrosis, it is not suitable for extensive use in all patients with NAFLD owing to its invasiveness and high cost. Therefore, noninvasive biochemical and imaging biomarkers have been developed to overcome the limitations of liver biopsy. Imaging biomarkers for the stratification of liver fibrosis have been evaluated in patients with NAFLD using different imaging techniques, such as transient elastography, shear wave elastography, and magnetic resonance elastography. Furthermore, artificial intelligence and deep learning methods are increasingly being applied to improve the diagnostic accuracy of imaging techniques and overcome the pitfalls of existing imaging biomarkers. In this review, we describe the usefulness and future prospects of noninvasive imaging biomarkers that have been studied and used to evaluate the degree of liver fibrosis in patients with NAFLD.

**Keywords:** noninvasive, surrogates, imaging biomarker, liver fibrosis, nonalcoholic fatty liver disease
INTRODUCTION

The prevalence of nonalcoholic fatty liver disease (NAFLD) is increasing worldwide, with approximately 25% of the global population being affected by this condition (1). Accordingly, the burden on the global healthcare system posed by the treatment of NAFLD is increasing and becoming a serious public health problem (2, 3). NAFLD comprises a spectrum of liver disorders ranging from isolated steatosis to nonalcoholic steatohepatitis (NASH), which can lead to serious conditions such as cirrhosis, hepatocellular carcinoma (HCC), and liver-related death (4, 5). In particular, the progression of liver fibrosis in patients with NAFLD is considered one of the most important factors determining prognosis, with significant and advanced liver fibrosis being an independent risk factor for both hepatic and extrahepatic complications and liver-related and overall mortality (6, 7). Therefore, accurate assessment of the degree of liver fibrosis in patients with NAFLD is the main issue to be addressed in modern medicine.

Although liver biopsy is the gold standard method for evaluating liver fibrosis in patients with NAFLD, its general clinical use is limited due to the high cost and potential complications (8). Moreover, liver biopsy has a disadvantage in that it can sample only a limited portion (1/50,000) of the entire liver. Therefore, many noninvasive tests (NITs) have been developed to overcome the limitations of liver biopsy, and their use in clinical practice is gradually increasing (9). Noninvasive imaging biomarkers can be broadly divided into ultrasound-based tests, such as vibration-controlled transient elastography (VCTE) and shear wave elastography (SWE) or acoustic radiation force impulse imaging (ARFI), and magnetic resonance imaging (MRI)-based tests, such as magnetic resonance elastography (MRE) (10) (Figure 1). As each test has its strengths and limitations, understanding the characteristics of each test is essential to selecting the optimal modality for assessing the degree of liver fibrosis in patients with NAFLD.

As research on noninvasive imaging biomarkers continues, more efficient test equipment
is expected to be developed and utilized in the future. In particular, methods that utilize artificial intelligence (AI), which have recently been in the spotlight, are expected to increase the accuracy and maximize the efficiency of existing inspection equipment (11). Recent studies on the use of AI or deep learning methods in evaluating the degree of liver fibrosis showed promising results (12, 13).

This review describes the application and advantages of noninvasive imaging biomarkers that have been studied and used to evaluate liver fibrosis in patients with NAFLD, as well as the future prospects of such biomarkers.

ELASTOGRAPHY
Elastography techniques are used to evaluate the stage of fibrosis by quantifying the shear wave velocity or tissue displacement generated by an ultrasonic or physical impulse, which represents liver stiffness (LS) (14). VCTE and MRE systems have mechanical drivers that generate shear waves and assess shear wave velocities using sonographic Doppler and magnetic resonance techniques, respectively (15). High-frequency sonographic impulses generate shear waves in point SWE (pSWE), ARFI, and two-dimensional SWE (2D-SWE). Because different elastography techniques are based on different methods and use different frequencies, their values are not identical, and caution is required when interpreting the results. Therefore, the strengths and limitations of each modality must be considered (Table 1).

ULTRASOUND-BASED ELASTOGRAPHY
Vibration-Controlled Transient Elastography

Technique
Transient elastography (FibroScan®; EchoSens, Paris, France) is an ultrasound-based elastography technique that is now a well-established noninvasive method for diagnosing and
staging liver fibrosis in patients with NAFLD (16). VCTE consists of a 3.5-MHz ultrasound transducer installed on the axis of a low-amplitude vibrator and utilizes monodimensional ultrasound to determine LS by measuring the velocity of low-frequency elastic shear waves propagating through the liver (17). For a VCTE result to be reliable, a minimum of 10 valid measurements are required, and the ratio of the median valid LS measurement to the interquartile range (IQR) should be ≤0.3 (18).

**Strengths and limitations**

A transient elastography test can be completed in a relatively short time (generally within 5 min), and many studies have validated the reliability of this test in assessing liver fibrosis in patients with NAFLD (19). Transient elastography also has excellent intraobserver and interobserver variability (20). However, transient elastography has the following limitations: the optimal cutoff point is unclear; measurements may be impossible in patients with obesity; the scan results may be unreliable in the hands of inexperienced operators; and the diagnostic accuracy is limited in the early stages of fibrosis (21).

**Clinical applications**

**Detection and staging of liver fibrosis**

Several recent studies have investigated the ideal cutoff value in VCTE to confirm significant liver fibrosis in patients with NAFLD (22-27). In those studies, the average body mass index (BMI) of patients with NAFLD was 27.1–34.8 kg/m², and the BMI of patients in Asian studies was relatively lower than that in Western studies. The LS value measured by VCTE indicating the presence of significant liver fibrosis (F2) in patients with NAFLD ranged from 7.7 to 9.8 kilopascals (kPa), and the proportion of patients with significant liver fibrosis ranged from 30.9% to 70.8% of the study population. In addition, the LS value indicating the presence of advanced
liver fibrosis or cirrhosis (F3 or higher) ranged from 7.3 to 12.5 kPa, which showed an acceptable area under the receiver operating characteristic curve (AUROC) values (0.80–0.92) (Table 2).

Prediction of liver-related outcomes

Recent studies have shown that baseline LS values measured by VCTE accurately predict the occurrence of liver decompensation, and higher baseline LS values can predict the development of liver-related events in patients with NAFLD (28, 29). In a multicenter cohort study that analyzed liver-related outcomes based on LS values measured by VCTE, baseline LS values were independently associated with the occurrence of hepatic decompensation (hazard ratio [HR]=1.03), HCC (HR=1.03), and liver-related death (HR=1.02) (29). In addition, an increase of >20% in the LS value during a mean follow-up period of 35 months was strongly associated with the risk of liver-related events and death, thus showing that LS values measured by VCTE are useful in predicting liver-related outcomes (29). However, owing to the limitations inherent in retrospective studies, the study did not follow a standardized protocol for VCTE follow-up and could not accurately identify the use of alcohol and other drugs. Therefore, future prospective and validation studies are needed to clarify the association between LS values measured by VCTE and liver-related outcomes.

Point Shear Wave Elastography/Acoustic Radiation Force Impulse Imaging

Technique

pSWE and ARFI are ultrasound-based elastography methods that enable the quantitative assessment of tissue stiffness. LS measurement with pSWE and ARFI is performed in the right lobe of the liver through the intercostal space. After selecting a region of interest (ROI), the shear wave velocity is measured within the defined region using ultrasound tracking beams
laterally adjacent to a single push beam (30). For the results of pSWE and ARFI to be reliable, the IQR/liver spasticity should be <30% (31-33).

**Strengths and limitations**

Similar to VCTE, several meta-analysis studies have confirmed that pSWE and ARFI have good diagnostic accuracy for significant liver fibrosis, with a mean AUROC of 0.84–0.87, and excellent diagnostic accuracy for cirrhosis, with a mean AUROC of 0.91–0.94 (31, 33). In addition, pSWE and ARFI have good intraobserver and interobserver agreement, with an intraclass correlation coefficient of between 0.84 and 0.87 (34, 35). In addition, unlike VCTE, the accuracy of pSWE and ARFI is generally not limited by obesity or interfering structures such as blood vessels or the biliary tract, as the ROI can be manually positioned (30). However, the disadvantages of pSWE and ARFI are that the size of the ROI is smaller than that in VCTE and the quality criteria are less evaluated.

**Clinical applications**

*Detection and staging of liver fibrosis*

Several studies have demonstrated the clinical application of pSWE and ARFI through noninvasive imaging biomarkers and the results showed that pSWE and ARFI are suitable diagnostic tools with higher diagnostic accuracy for advanced liver fibrosis (F3–4) than low-grade fibrosis (F1–2)(36, 37). However, studies on pSWE and ARFI have been mainly monocentric retrospective studies; therefore, longitudinal validation in chronic liver diseases, especially NAFLD, is required to develop standardized quality criteria.

**Two-Dimensional Shear Wave Elastography**

*Technique*
Real-time 2D-SWE is performed rather similarly to pSWE and ARFI. It combines the initiation of a radiation force in tissues using focused ultrasonic beams and the acquisition of transiently propagating resultant shear waves in real-time with a high-frequency ultrasound imaging sequence (38). In 2D-SWE, a two-dimensional parametric color map is generated by combining several shear waves over time with rapid ultrasound acquisition. Similar to pSWE and ARFI, 2D-SWE allows the operator to select the size and location of the ROI. When the operator “samples” a specific area within a color map, the shear-wave velocity is measured to obtain a quantitative measure of tissue elasticity using proprietary software (Aixplorer®; Supersonic Imaging, Aix en Provence, France) (39).

**Strengths and limitations**

The advantage of 2D-SWE is that it allows the operator to select the size and location of the ROI, thereby permitting the evaluation of the elasticity profile of a larger tissue section in a single acquisition (40). In addition, 2D-SWE has the following advantages over pSWE and ARFI: qualitative (color-coded) and quantitative measurement, easier and more manageable measurement, and stability of the measured value (41, 42). However, 2D-SWE has some limitations, including the subjective nature of the color scale, potential bias when selecting the ROI, and a lack of meta-analysis confirming its clinical applications.

**Clinical applications**

*Detection and staging of liver fibrosis*

Several recent studies have confirmed that LS measured by 2D-SWE strongly correlates with the stage of liver fibrosis on liver biopsy in patients with NAFLD (43). According to a meta-analysis conducted in Europe, 2D-SWE has good diagnostic performance for significant liver fibrosis (≥F2, AUROC=0.86) and excellent diagnostic performance for severe fibrosis (≥F3,
AUROC=0.93) and cirrhosis (F4, AUROC=0.92). The optimal cutoff values for diagnosing significant liver fibrosis and cirrhosis were reported to be 7.1 and 13.0 kPa, respectively. In addition, the AUROC for the diagnosis of significant liver fibrosis (P=0.001) and cirrhosis (P=0.022) with 2D-SWE was higher than that with VCTE (44). However, as studies on the clinical application of 2D-SWE and comparative studies with other noninvasive methods are lacking, follow-up studies are needed.

MAGNETIC RESONANCE IMAGING-BASED ELASTOGRAPHY

Technique
Liver MRE can be performed using existing magnetic resonance scanners. The setup includes an active pneumatic mechanical driver located outside the scanning room and a connected passive driver placed on the liver (45). The active driver generates continuous acoustic vibrations that are transmitted to the passive driver and subsequently to the abdomen, including the liver. These waves produce microscopic shear displacement of tissues, which is visualized using MRE sequences as propagating shear waves (46). Subsequently, a magnitude image revealing the anatomy of the upper abdomen and a phase-contrast image showing shear waves at the same level are reconstructed, and grayscale and colored stiffness maps, also known as elastograms, are produced.

Thereafter, readers draw the ROI within the confidence map of the liver, avoiding the liver edge, artifacts, fissures, fossa, and regions of wave interference (45). The mean LS value is calculated using ROIs on four slices. The LS value measured by MRE is expressed in kPa, representing both the elasticity and viscosity of the tissue.

Strengths and Limitations
MRE can examine the entire liver, and technical failure occurs in <5% of the examinations (47-
MRE measurements are highly reproducible, with robust intraobserver and interobserver agreements (50-53). The LS value measured by MRE is not significantly affected by hepatic steatosis, and MRE can measure LS in patients with obesity (54-57). In addition, hepatic inflammation does not affect the accuracy of MRE in patients with NAFLD (55).

The most common cause of technical failure in MRI is iron overload (55). Poor transmission of shear waves into the liver because of massive ascites increased subcutaneous fat thickness, and poor contact between the passive driver and the abdominal wall also led to a measurement failure. Inconsistent breath-holding and motion during the sequence are common causes of technical failure in patients with massive ascites (45). The heterogeneity of fibrosis progression in different liver lesions may lead to inaccurate LS measurements, particularly in small ROIs (58). MRE cannot differentiate LS caused by congestion from that caused by increased vascular pressure; thus, the LS value measured by MRE should be carefully interpreted (59). Differences in MRI specifications and vendors among institutions and studies are another concern in the interpretation of LS values measured by MRE. Finally, considering its cost and limited availability, MRE cannot be generally used in clinical practice at present.

Clinical Applications

Detection and staging of liver fibrosis

Multiple studies have demonstrated that MRE has excellent accuracy in diagnosing and stratifying liver fibrosis in patients with NAFLD, predicting significant or advanced liver fibrosis and cirrhosis with consistent AUROC values of >0.90 (60-63) (Table 3). A recent meta-analysis showed the excellent accuracy of MRE, with an AUROC of 0.96 for advanced liver fibrosis and 0.92 for cirrhosis and LS cutoff values of 3.62–4.8 and 4.15–6.7 kPa, respectively (58). A meta-analysis of nine studies that included 232 patients with NAFLD suggested reliable LS cutoff values of 2.88, 3.54, 3.77, and 4.09 kPa for detecting fibrosis stages 1, 2, 3, and 4,
respectively (55).

In a recent meta-analysis with individual data of 230 patients with biopsy-proven NAFLD, MRE outperformed VCTE in detecting all stages of fibrosis (AUROC for fibrosis stage ≥1, 0.87 vs. 0.82 [P=0.04]; stage ≥2, 0.92 vs. 0.87 [P=0.03]; stage ≥3, 0.93 vs. 0.84 [P=0.001]; and stage ≥4, 0.94 vs. 0.84 [P=0.005]) (64). Comparative studies between MRE and pSWE are limited; however, one study demonstrated that MRE was more accurate than pSWE in diagnosing any fibrosis stage in patients with NAFLD, especially in those with obesity (56). A recent study demonstrated that MRE was more accurate than 2D-SWE in diagnosing stage ≥1 and ≥2 fibrosis but not stages ≥3 or 4 fibrosis (27). Other MRI techniques, including diffusion-weighted imaging or contrast-enhanced MRI, were also reported to be less accurate than MRE in assessing liver fibrosis (65, 66). Consequently, the LS value measured by MRE can be considered the most accurate noninvasive imaging biomarker for detecting all stages of fibrosis (Table 4).

Recently, noninvasive LS-based models combining two different biomarkers have shown promising results in identifying patients with significant liver fibrosis, with increased positive predictive value (PPV), thereby reducing screening failure rates in clinical trials and reducing unnecessary liver biopsies (67-69). In previous studies, MEFIB (MRE plus fibrosis-4 [FIB-4]) had a significantly higher diagnostic accuracy than MRE alone and the FIB-4 index alone (67). Notably, a recent study compared MEFIB, MAST (MRI–aspartate aminotransferase), and FAST (FibroScan–aspartate aminotransferase) in detecting stage ≥2 fibrosis among patients with NAFLD and demonstrated the superiority of MEFIB (PPV, 95%; negative predictive value, 90%) over MAST and FAST (both P<0.001) (69).

Prediction of liver-related outcomes

Multiple retrospective studies have suggested that MRE can play a role in predicting the long-
term prognosis of patients with NAFLD (70-72). A recent meta-analysis of six cohorts, including 1,707 patients with a median follow-up of 3 years, investigated the association between the LS value measured by MRE and liver-related outcomes (67). The HR for liver-related outcomes in patients with an LS value of 5–8 kPa was 11.0 (P<0.001) and that in patients with an LS value of ≥8 kPa was 15.9 (P<0.001), compared with those with an LS value of <5 kPa. Furthermore, the MEFIB index was developed using the identified best cutoff values for LS and the FIB-4 index (defined as positive when the LS value measured by MRE was ≥3.3 kPa and the FIB-4 index was ≥1.6). A positive MEFIB index had a robust association with liver-related outcomes (HR=20.6; P<0.001), and a negative MEFIB had a high negative predictive value for liver-related outcomes (99.1% at 5 years).

However, few retrospective studies have described the association of MRE with the clinical outcomes of patients with NAFLD. Therefore, future multicenter prospective studies are required to clarify the association between LS measured by MRE and liver-related clinical outcomes.

**Emerging Magnetic Resonance Imaging-Based Techniques**

Advances in MRE techniques, including automated liver elasticity calculations and improvements in shear-wave delivery, are promising to provide a faster and more reliable evaluation of the liver. Three-dimensional (3D)-MRE is a newly developed imaging technique that assesses shear-wave propagation in multiple planes to avoid mathematical assumptions (63). For the 3D-MRE examination, a separate motion-sensitized, multislice, spin-echo echo-planar imaging sequence is performed to assess shear-wave displacements along the x-, y-, and z-directions.

Although 3D-MRE is more accurate than 2D-MRE in predicting advanced liver fibrosis in patients with NAFLD, further validation is required to prove the benefits of this technique.
Multiparametric MRI measures shear stiffness, loss modulus, and MRI-derived fat fraction in a single scan. 3D-MRE incorporates a damping ratio at a lower frequency, which may further help in the detection of NASH and NASH-related fibrosis (73).

**ARTIFICIAL INTELLIGENCE**

Recently, AI and deep learning methods have been incorporated into MRE and shown encouraging results. AI can make quantitative assessments objective, reproducible, and less ambiguous. Traditional (supervised) machine learning and deep learning algorithms use approaches that are dependent on predefined information or ROIs determined by experts (11). Deep learning does not rely on predefined features and does not always require a focus on ROIs. Convolutional neural networks (CNNs) are the most commonly applied deep learning methods in imaging analysis. In a retrospective study, LS measurements using an automated CNN-based method strongly agreed with manual ROI-based analysis across MRE systems (intraclass correlation coefficient, 0.98–0.99) and showed excellent discriminative performance for histology-determined stages of liver fibrosis (AUROC=0.89–0.93) in patients with NAFLD (74). Considering the high incidence of NAFLD, CNN-based analysis may reduce reliance on expert image analysts.

Radiomic texture analysis is an evolving translational tool used to extract imaging information, which is prone to subjective and variable interpretation. A recent study applied texture analysis–derived parameters combined with machine learning to MRI-based techniques for the quantification of liver fibrosis (12). Texture analysis and machine learning techniques were tested on T1- and T2-weighted MRI and MRE images of 62 participants with histologic evidence of chronic liver disease. The diagnostic accuracy for advanced liver fibrosis in T1-weighted MRI and MRE images was excellent (AUROC=0.82 vs. 0.92, P=0.41); however, T2-weighted MRI had a lower accuracy (AUROC=0.57).
Integrating AI into conventional noninvasive tools can provide an optimal balance between sensitivity and specificity in assessing liver fibrosis. Thus far, few studies have investigated the application of AI in the assessment of imaging biomarkers in NAFLD; however, studies evaluating liver fibrosis in patients with NAFLD are expected to gradually increase in the future.

ROLE OF NONINVASIVE TESTS IN DISEASE MONITORING

Repeated measurements using NITs can stratify the risk of liver-related events in patients with NAFLD. Currently, limited data are available on the impact of dynamic changes in LS values measured using NITs on the long-term outcomes of patients with NAFLD.

VCTE is useful for monitoring the severity of liver fibrosis not only in patients with NAFLD but also in patients with NASH-related cirrhosis, and LS can be a useful biomarker for predicting varices, HCC, and liver-related death (75). According to a multinational study conducted in Europe in 790 patients with NAFLD-related compensated cirrhosis, the LS value measured by VCTE can effectively identify varices requiring treatment and reduce unnecessary endoscopies (76). In addition, some studies have indicated that VCTE can be used to monitor fibrosis changes after treatment, although this should be confirmed by further studies using paired liver biopsies (77, 78).

In a prospective cohort study, 102 patients with biopsy-proven NAFLD underwent contemporaneous MRE and liver biopsy at baseline, followed by repeat paired liver biopsy and MRE assessment (79). A 15% increase in the LS value measured by MRE was associated with histologic fibrosis progression and progression from early to advanced liver fibrosis. A retrospective study of 128 patients with NAFLD who underwent at least two serial MRE examinations showed a significantly higher risk of the development of cirrhosis and decompensation or death in patients with a ≥19% increase in LS value from baseline than in
those without (80).

Further studies are warranted to assess the implication of changes in LS measured using NITs over time on the risk of future liver-related events and mortality. Furthermore, although evidence is lacking and the optimal time interval remains to be determined, repeating NITs every 3 years in patients with early-stage NAFLD and every year in patients with advanced-stage disease seems reasonable.
CONCLUSION

Currently, the main utility of noninvasive imaging biomarkers in NAFLD is discriminating patients with significant or advanced liver fibrosis from those with mild or no fibrosis for prognosis prediction and clinical decision-making. VCTE is the most widely validated test; pSWE and 2D-SWE have comparable performance to VCTE; and MRE is currently considered the most accurate noninvasive tool for the detection and staging of liver fibrosis. However, the clinical use of these tests is usually determined by the availability of the technology and the local expertise at each institution.

A major limitation of NITs is their suboptimal accuracy in diagnosing fibrosis in the early stages and in adequately discriminating between adjacent fibrosis stages. Differentiating other processes that cause increased LS values, such as inflammation, biliary obstruction, cholestasis, passive congestion, and increased portal venous pressure, from liver fibrosis is another challenge. Research on noninvasive imaging biomarkers in NAFLD, especially concerning their use in screening and risk prediction, will continue as the prevalence of the disease increases and as newer treatment methods emerge. Finally, noninvasive imaging biomarkers, liver biopsies, and clinical parameters must be used in combination for the accurate assessment of the fibrosis stage and risk stratification in patients with NAFLD.
References


9. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease


15. Yin M, Venkatesh SK. Ultrasound or MR elastography of liver: which one shall I use? Abdom Radiol (NY) 2018;43:1546-1551.


34. Myers RP, Crotty P, Pomier-Layrargues G, Ma M, Urbanski SJ, Elkashab M. Prevalence, risk factors and causes of discordance in fibrosis staging by transient elastography


59. Venkatesh SK, Yin M, Ehman RL. Magnetic resonance elastography of liver:


**Figure 1.** Currently used noninvasive imaging biomarkers in NAFLD

NAFLD, nonalcoholic fatty liver disease; SWE, shear wave elastography
Figure 2. Algorithm for risk discrimination in patients with NAFLD using noninvasive imaging biomarkers

NAFLD, nonalcoholic fatty liver disease; LS, liver stiffness; VCTE, vibration-controlled transient elastography; MRE, magnetic resonance elastography; HCC, hepatocellular carcinoma
Table 1. Strengths and limitations of noninvasive imaging tests for liver fibrosis in NAFLD

<table>
<thead>
<tr>
<th>Methods</th>
<th>Validation</th>
<th>Reliability</th>
<th>Reliability criteria</th>
<th>Indeterminate cases</th>
<th>Failure rate</th>
<th>Factors related to failure</th>
<th>Invalid result rate</th>
<th>Confounders</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE</td>
<td>Fibrosis stage, liver-related outcomes</td>
<td>0.99</td>
<td>Yes</td>
<td>30-40%</td>
<td>3-14%</td>
<td>Obesity (less with XL probe), ascites</td>
<td>1-9%</td>
<td>Acute hepatitis, cholestasis, food ingestion, obesity, congestion</td>
</tr>
<tr>
<td>pSWE</td>
<td>Fibrosis stage</td>
<td>0.98</td>
<td>Yes</td>
<td>NA</td>
<td>0-1%</td>
<td>Obesity</td>
<td>16-24%</td>
<td>Acute hepatitis, food ingestion, obesity,</td>
</tr>
<tr>
<td>2D-SWE</td>
<td>Fibrosis stage</td>
<td>0.98-1.0</td>
<td>No</td>
<td>NA</td>
<td>1-13%</td>
<td>Obesity</td>
<td>0%</td>
<td>Acute hepatitis, food ingestion</td>
</tr>
<tr>
<td>MRE</td>
<td>Fibrosis stage, liver-related outcomes</td>
<td>0.99</td>
<td>Yes</td>
<td>NA</td>
<td>&lt;5%</td>
<td>Massive ascites, poor contact between the passive driver and the abdominal wall, inconsistent breath holding and motion, claustrophobia, inability to fit in the MRI machine</td>
<td>Negligible</td>
<td>Iron overload, acute hepatitis, massive ascites</td>
</tr>
</tbody>
</table>

NAFLD, nonalcoholic fatty liver disease; TE, transient elastography; pSWE, point shear wave elastography; 2D-SWE, two-dimensional shear wave elastography; MRE, magnetic resonance elastography; NA, not available.
<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Country</th>
<th>Mean BMI, kg/m²</th>
<th>Patients with advanced fibrosis (F≥2), n, %</th>
<th>AUROC for F≥2 (95% CI)</th>
<th>Cut off for F≥2</th>
<th>Sensitivity/ specificity</th>
<th>AUROC for F≥3 (95% CI)</th>
<th>Cut off for F≥3</th>
<th>Sensitivity/ specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cassinotto et al. 2016</td>
<td>291</td>
<td>France</td>
<td>32.1</td>
<td>206, 70.8%</td>
<td>0.82</td>
<td>9.8</td>
<td>0.60/0.90</td>
<td>0.86</td>
<td>12.5</td>
<td>0.57/0.90</td>
</tr>
<tr>
<td>Lee et al. 2017 (25)</td>
<td>94</td>
<td>Korea</td>
<td>27.1</td>
<td>46, 47.9%</td>
<td>0.76</td>
<td>7.4</td>
<td>0.62/0.92</td>
<td>0.87</td>
<td>8</td>
<td>0.82/0.85</td>
</tr>
<tr>
<td>Park et al. 2017 (24)</td>
<td>94</td>
<td>US</td>
<td>30.4</td>
<td>29, 30.9%</td>
<td>0.86</td>
<td>6.9</td>
<td>0.79/0.85</td>
<td>0.80</td>
<td>7.3</td>
<td>0.78/0.78</td>
</tr>
<tr>
<td>Fulan et al. 2019 (26)</td>
<td>62</td>
<td>US</td>
<td>34.8</td>
<td>44, 70.1%</td>
<td>0.77</td>
<td>8.8</td>
<td>0.51/0.94</td>
<td>0.86</td>
<td>10.5</td>
<td>0.50/0.92</td>
</tr>
<tr>
<td>Eddowes et al. 2019 (22)</td>
<td>450</td>
<td>UK</td>
<td>33.8</td>
<td>225, 60%</td>
<td>0.77</td>
<td>8.2</td>
<td>0.71/0.70</td>
<td>0.80</td>
<td>9.7</td>
<td>0.71/0.75</td>
</tr>
<tr>
<td>Imajo et al. 2022 (27)</td>
<td>201</td>
<td>Japan</td>
<td>27.1</td>
<td>-</td>
<td>0.89</td>
<td>8.4</td>
<td>0.86/0.74</td>
<td>0.92</td>
<td>9.7</td>
<td>0.84/0.85</td>
</tr>
</tbody>
</table>

NAFLD, nonalcoholic fatty liver disease; AUROC, area under the receiver operating characteristics curve; CI, confidence interval; BMI, body mass index.
<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Country</th>
<th>Mean BMI, kg/m²</th>
<th>Patients with advanced fibrosis (F≥2), %</th>
<th>AUROC for F≥2 (95% CI)</th>
<th>Cut off for F≥2, kPa</th>
<th>Sensitivity/ specificity</th>
<th>AUROC for F≥3 (95% CI)</th>
<th>Cut off for F≥3, kPa</th>
<th>Sensitivity/ specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imajo et al. 2022 (27)</td>
<td>201</td>
<td>Japan</td>
<td>27.1</td>
<td>-</td>
<td>0.927</td>
<td>3.19</td>
<td>0.90/0.81</td>
<td>0.929</td>
<td>3.90</td>
<td>0.83/0.92</td>
</tr>
<tr>
<td>Cui et al. 2016 (56)</td>
<td>125</td>
<td>United States</td>
<td>31.8</td>
<td>26.4</td>
<td>0.885</td>
<td>3.62</td>
<td>0.67/0.96</td>
<td>0.934</td>
<td>3.62</td>
<td>0.91/0.93</td>
</tr>
<tr>
<td>Loomba et al. 2014 (61)</td>
<td>117</td>
<td>United States</td>
<td>32.4</td>
<td>29.9</td>
<td>0.856</td>
<td>3.58</td>
<td>0.66/0.92</td>
<td>0.894</td>
<td>4.67</td>
<td>0.8/0.94</td>
</tr>
<tr>
<td>Costa-Silva et al. 2018 (62)</td>
<td>90</td>
<td>Brazil</td>
<td>32.2</td>
<td>24.4</td>
<td>0.932</td>
<td>4.14</td>
<td>0.92/0.97</td>
<td>0.928</td>
<td>4.39</td>
<td>0.91/0.97</td>
</tr>
<tr>
<td>Loomba et al. 2016 (63)</td>
<td>100</td>
<td>United States</td>
<td>32.1</td>
<td>27.0</td>
<td>0.878</td>
<td>3.65</td>
<td>-</td>
<td>0.921</td>
<td>3.80</td>
<td>0.87/0.94</td>
</tr>
</tbody>
</table>

NAFLD, nonalcoholic fatty liver disease; AUROC, area under the receiver operating characteristics curve; CI, confidence interval; BMI, body mass index.
Table 4. Diagnostic accuracy of noninvasive imaging biomarkers for each stage of fibrosis in NAFLD

<table>
<thead>
<tr>
<th>Stage ≥2 fibrosis</th>
<th>Method</th>
<th>TE</th>
<th>pSWE</th>
<th>2D-SWE</th>
<th>MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUROC</td>
<td>0.77</td>
<td>0.87</td>
<td>0.86</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>71.0</td>
<td>79.0</td>
<td>94.0</td>
<td>84.9</td>
<td></td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>70.0</td>
<td>85.0</td>
<td>52.0</td>
<td>85.4</td>
<td></td>
</tr>
<tr>
<td>PPV (%)</td>
<td>78.0</td>
<td>91.0</td>
<td>65.1</td>
<td>79.8</td>
<td></td>
</tr>
<tr>
<td>NPV (%)</td>
<td>61.0</td>
<td>66.0</td>
<td>86.7</td>
<td>89.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stage ≥3 fibrosis</th>
<th>Method</th>
<th>TE</th>
<th>pSWE</th>
<th>2D-SWE</th>
<th>MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUROC</td>
<td>0.80</td>
<td>0.91</td>
<td>0.93</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>71.0</td>
<td>92.0</td>
<td>93.0</td>
<td>82.5</td>
<td></td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>75.0</td>
<td>86.0</td>
<td>81.0</td>
<td>83.2</td>
<td></td>
</tr>
<tr>
<td>PPV (%)</td>
<td>63.0</td>
<td>82.0</td>
<td>77.0</td>
<td>61.8</td>
<td></td>
</tr>
<tr>
<td>NPV (%)</td>
<td>81.0</td>
<td>89.0</td>
<td>97.4</td>
<td>93.5</td>
<td></td>
</tr>
</tbody>
</table>

NAFLD, nonalcoholic fatty liver disease; TE, transient elastography; pSWE, point shear-wave elastography; 2D-SWE, two-dimensional shear-wave elastography; MRE, magnetic resonance elastography; AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value