Medical Policy

Blue Cross Blue Shield Blue Care Network of Michigan

Nonprofit corporations and independent licensees of the Blue Cross and Blue Shield Association

Joint Medical Policies are a source for BCBSM and BCN medical policy information only. These documents are not to be used to determine benefits or reimbursement. Please reference the appropriate certificate or contract for benefit information. This policy may be updated and is therefore subject to change.

***Current Policy Effective Date: 5/1/24**
(See policy history boxes for previous effective dates)

Title: Noninvasive Techniques for the Evaluation and Monitoring of Patients with Chronic Liver Disease

Description/Background

Noninvasive techniques to monitor liver fibrosis are being investigated as alternatives to liver biopsy in patients with chronic liver disease. There are two options for noninvasive monitoring: (1) multianalyte serum assays with algorithmic analysis of either direct or indirect biomarkers; and (2) specialized radiographic methods, including magnetic resonance elastography, multiparametric magnetic resonance imaging (MRI), transient elastography, acoustic radiation force impulse imaging, and real-time transient elastography.

BIOPSY FOR CHRONIC LIVER DISEASE

The diagnosis of non–neoplastic liver disease is often made from needle biopsy samples. In addition to establishing a disease etiology, liver biopsy can determine the degree of inflammation present and stage the degree of fibrosis. The degree of inflammation and fibrosis may be assessed by different scoring schemes. Most of these scoring schemes grade inflammation from 0 (no or minimal inflammation) to 4 (severe) and fibrosis from 0 (no fibrosis) to 4 (cirrhosis). There are several limitations to liver biopsy, including its invasive nature, small tissue sample size, and subjective grading system. Regarding small tissue sample size, liver fibrosis can be patchy and thus missed on a biopsy sample, which includes only 0.002% of the liver tissue. A noninvasive alternative to liver biopsy would be particularly helpful, both to initially assess patients and then as a monitoring tool to assess response to therapy. The implications of using liver biopsy as a reference standard are discussed in the Rationale.

Hepatitis C Virus

Infection with hepatitis C virus (HCV) can lead to permanent liver damage. Before noninvasive tests were available, liver biopsy is typically recommended before the initiation of antiviral therapy. Repeat biopsies may be performed to monitor fibrosis progression. Liver biopsies are analyzed according to a histologic scoring system; the most commonly used one for hepatitis C is the Metavir scoring system, which scores the presence and degree of inflammatory activity and fibrosis. The fibrosis is graded from F0 to F4, with a Metavir score of F0 signifying no fibrosis and F4 signifying cirrhosis (which is defined as the presence throughout the liver of

fibrous septa that subdivide the liver parenchyma into nodules and represents the final and irreversible form of disease). The stage of fibrosis is the most important single predictor of morbidity and mortality in patients with hepatitis C. Biopsies for hepatitis C are also evaluated according to the degree of inflammation present, referred to as the grade or activity level. For example, the Metavir system includes scores for necroinflammatory activity ranging from A0 to A3 (A0 = no activity, A1 = minimal activity, A2 = moderate activity, A3 = severe activity).

Hepatitis B Virus

Most people who become infected with hepatitis B virus (HBV) recover fully, but a small portion will develop chronic HBV, which can lead to permanent liver damage. As with HCV, identification of liver fibrosis is needed to determine timing and management of treatment, and liver biopsy is the criterion standard for staging fibrosis. The grading of fibrosis in HBV also uses the Metavir system.

Alcoholic Liver Disease

Alcoholic liver disease (ALD) is the leading cause of liver disease in most Western countries. Histologic features of ALD usually include steatosis, alcoholic steatohepatitis (ASH), hepatocyte necrosis, Mallory bodies (tangled proteins seen in degenerating hepatocytes), a large polymorphonuclear inflammatory infiltrate, and, with continued alcohol abuse, fibrosis and possibly cirrhosis. The grading of fibrosis is similar to the scoring system used in hepatitis C. The commonly used Laënnec scoring system uses grades 0 to 4, with 4 being cirrhosis.

Nonalcoholic Fatty Liver Disease

Nonalcoholic fatty liver disease (NAFLD) is defined as a condition that pathologically resembles ALD but occurs in patients who are not heavy users of alcohol. It may be associated with a variety of conditions, including obesity, diabetes, and dyslipidemia. The characteristic feature of NAFLD is steatosis. At the benign end of the disease spectrum, there is usually no appreciable inflammation, hepatocyte death, or fibrosis. In contrast, nonalcoholic steatohepatitis (NASH), which shows overlapping histologic features with ALD, is an intermediate form of liver damage, and liver biopsy may show steatosis, Mallory bodies, focal inflammation, and degenerating hepatocytes. NASH can progress to fibrosis and cirrhosis. A variety of histologic scoring systems have been used to evaluate NAFLD. The NAFLD Activity Score (NAS) system for NASH includes scores for steatosis (0-3), lobular inflammation (0-3), and ballooning (0-2). Cases with scores of 5 or greater are considered NASH, while cases with scores of 3 and 4 are considered borderline (probable or possible) NASH. The grading of fibrosis is similar to the scoring system used in hepatitis C. The commonly used Laënnec scoring system uses grades 0 to 4, with 4 being cirrhosis.

NONINVASIVE ALTERNATIVES TO LIVER BIOPSY

Multianalyte Assays

A variety of noninvasive laboratory tests are being evaluated as an alternative to liver biopsy. Biochemical tests can be broadly categorized into indirect and direct markers of liver fibrosis. Indirect markers include liver function tests such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), the ALT/AST ratio (also referred to as the AAR), platelet count, and prothrombin index. In recent years, there has been growing understanding of the underlying pathophysiology of fibrosis, leading to direct measurement of the factors involved. For example, the central event in the pathophysiology of fibrosis is activation of the hepatic stellate cell. Normally, stellate cells are quiescent but are activated in the setting of liver injury, producing a variety of extracellular matrix (ECM) proteins. In normal livers, the rate of ECM production

equals its degradation, but in the setting of fibrosis, production exceeds degradation. Metalloproteinases are involved in intracellular degradation of ECM, and a profibrogenic state exists when there is either a down regulation of metalloproteinases or an increase in tissue inhibitors of metalloproteinases (TIMP). Both metalloproteinases and TIMP can be measured in the serum, which directly reflects the fibrotic activity. Other direct measures of ECM deposition include hyaluronic acid or α2-macroglobulin.

While many studies have been done on these individual markers, or on groups of markers in different populations of patients with liver disease, there has been interest in analyzing multiple markers using mathematical algorithms to generate a score that categorizes patients according to the biopsy score. It is proposed that these algorithms can be used as an alternative to liver biopsy in patients with liver disease. The following proprietary, algorithm-based tests are commercially available in the United States.

FIBROSURE

There are 3 different FibroSURE tests available depending on the indication for use: HCV FibroSURE, ASH FibroSURE, and NASH FibroSURE.

HCV FibroSURE™

The HCV FibroSURE uses a combination of 6 serum biochemical indirect markers of liver function plus age and sex in a patented algorithm to generate a measure of fibrosis and necroinflammatory activity in the liver that correspond to the Metavir scoring system for stage (ie, fibrosis) and grade (ie, necroinflammatory activity). The biochemical markers include the readily available measurements of α2-macroglobulin, haptoglobin, bilirubin, γ-glutamyl transpeptidase (GGT), ALT, and apolipoprotein A1. Developed in France, the test has been clinically available in Europe under the name FibroTest since 2003 and is exclusively offered by LabCorp in the United States as HCV FibroSURE.

ASH FibroSURE™

ASH FibroSURE (ASH Test) uses a combination of 10 serum biochemical markers of liver function together with age, sex, height, and weight in a proprietary algorithm and is proposed to provide surrogate markers for liver fibrosis, hepatic steatosis, and alcoholic steatohepatitis (ASH). The biochemical markers include α2-macroglobulin, haptoglobin, apolipoprotein A1, bilirubin, GGT, ALT, AST, total cholesterol, triglycerides, and fasting glucose. The test has been available in Europe under the name ASH Test and is exclusively offered by LabCorp in the United States as ASH FibroSURE.

NASH FibroSURE™

NASH FibroSURE (NASH Test) uses a proprietary algorithm of the same 10 biochemical markers of liver function in combination with age, sex, height, and weight and is proposed to provide surrogate markers for liver fibrosis, hepatic steatosis, and NASH. The biochemical markers include α2-macroglobulin, haptoglobin, apolipoprotein A1, bilirubin, GGT, ALT, AST, total cholesterol, triglycerides, and fasting glucose. The test has been available in Europe under the name NASH Test and is exclusively offered by LabCorp in the United States as NASH FibroSURE.

FIBROSpect II

FIBROSpect II uses a combination of 3 markers that directly measure fibrogenesis of the liver, analyzed with a patented algorithm. The markers include hyaluronic acid, TIMP-1, and α2-

macroglobulin. FIBROSpect II is offered exclusively by Prometheus Laboratories. The measures are combined using a logistic regression algorithm to generate a FIBROSpect II index score, ranging from 1 to 100 (or sometimes reported between 0 and 1), with higher scores indicating more severe disease.

Enhanced Liver Fibrosis Test

The Enhanced Liver Fibrosis (ELF) test uses a proprietary algorithm to produce a score based on 3 serum biomarkers involved in matrix biology: hyaluronic acid, Procollagen III amino terminal peptide and tissue inhibitor of metalloproteinase 1. The manufacturer recommends the following cutoffs for interpretation for risk of development of cirrhosis or liver-related events in patients with NASH: <9.80 (lower risk) and ≥11.30 (higher risk).

NONINVASIVE IMAGING TECHNOLOGIES

Noninvasive imaging technologies to detect liver fibrosis or cirrhosis among patients with chronic liver disease have been developed as an alternative to liver biopsy. The noninvasive imaging technologies include transient elastography (eg, FibroScan®), magnetic resonance elastography (MRE), multiparametric magnetic resonance imaging (MRI), acoustic radiation force impulse (ARFI) (eg, Acuson S2000™), and real-time tissue elastography (eg, HI VISION Preirus). Noninvasive imaging tests have been used in combination with multianalyte serum tests such as FibroTest or FibroSURE with FibroScan.

Transient Elastography

Transient elastography (FibroScan®) uses a mechanical vibrator to produce mild amplitude and low frequency (50 Hz) waves, inducing an elastic shear wave that propagates throughout the liver. Ultrasonography tracks the wave, measuring its speed, which correlates with liver stiffness. Increases in liver fibrosis also increase liver stiffness and resistance of liver blood flow. Transient elastography does not perform as well in patients with ascites, higher body mass index, or narrow intercostal margins. Although FibroScan® may be used to measure fibrosis, unlike liver biopsy, it does not provide information on necroinflammatory activity and steatosis, nor is it accurate during acute hepatitis or hepatitis exacerbations.

Acoustic Radiation Force Impulse Imaging

Acoustic radiation force impulse imaging (ARFI) uses an ultrasound probe to produce an acoustic “push” pulse, which generates shear waves that propagate in tissue to assess liver stiffness. ARFI elastography evaluates the wave propagation speed to assess liver stiffness. The faster the shear wave speed, the harder the object. ARFI technologies include Virtual Touch™ Quantification and Siemens Acuson S2000™ system. ARFI elastography can be performed at the same time as a liver sonographic evaluation, even in patients with a significant amount of ascites.

Magnetic Resonance Elastography

Magnetic resonance elastography (MRE) uses a driver to generate 60-Hz mechanical waves on the patient’s chest well. The magnetic resonance equipment creates elastograms by processing the acquired images of propagating shear waves in the liver using an inversion algorithm. These elastograms represent the shear stiffness as a pixel value in kilopascals. MRE has several advantages over ultrasound elastography, including: (1) the ability to analyze larger liver volumes; (2) the ability to analyze liver volumes of obese patients or patients with ascites; and (3) the ability to precisely analyze viscoelasticity using a 3-dimensional displacement vector.

Real-Time Tissue Elastography

Real-time tissue elastography is a type of strain elastography which uses a combined autocorrelation method to measure tissue strain caused by manual compression or a person's heartbeat. The relative tissue strain is displayed on conventional color B mode ultrasound images in real time. Hitachi manufacturers the real-time tissue elastography devices, including one called HI VISION Preirus. The challenge is to identify a region of interest while avoiding areas likely to introduce artifacts, such as large blood vessels, the area near the ribs, and the surface of the liver. Areas of low strain increase as fibrosis progresses and strain distribution becomes more complex. Various subjective and quantitative methods have been developed to evaluate the results. Real-time tissue elastography can be performed in patients with ascites or inflammation. This technology does not perform as well in severely obese individuals.

Multiparametric Magnetic Resonance Imaging

Multiparametric MRI combines proton density fat-fraction, T2, and T1 mapping. Proton density fat-fraction provides an assessment of hepatic fat content and can be used to determine the grade of liver steatosis. T1 relaxation times are used to assess increases in extracellular fluid, which correlates with the extent of fibrosis and inflammation of the liver. Hepatic iron quantification is measured through T2 relaxation times as T1 relaxation times are decreased by excess iron in the liver tissue. LiverMultiScan® uses a clinical algorithm that accounts for an iron-corrected T1 value, based on the T2 relaxation time, and proton density fat-fraction to assess the presence of fat, inflammation, and fibrosis.

Regulatory Status

In 2008, Acuson S2000™ Virtual Touch (Siemens AG), which provides acoustic radiation force impulse imaging, was cleared for marketing by the U.S. Food and Drug Administration (FDA) through the 510(k) process (K072786).

In 2009, AIXPLORER® Ultrasound System (SuperSonic Imagine), which provides shear wave elastography, was cleared for marketing by the FDA through the 510(k) process (K091970).

In 2010, Hitachi HI VISION Preirus Diagnostic Ultrasound Scantier (Hitachi Medical Systems America), which provides real-time tissue elastography, was cleared for marketing by the FDA through the 510(k) process (K093466).

In 2013, FibroScan® (EchoSens), which uses transient elastography, was cleared for marketing by the FDA through the 510(k) process (K123806).

In June 2015, LiverMultiScan (Perspectum), which is a magnetic resonance diagnostic device software application, was cleared for marketing by the FDA through the 510(k) process (K143020).

In February 2017, ElastQ Imaging shear wave elastography (Royal Phillips) was cleared for marketing by the FDA through the 510(k) process (K163120).

In August 2021, ADVIA Centaur ELFTM test (Siemens Healthcare) was cleared for marketing by the FDA through the 513(f)(2) De Novo review pathway (DEN190056). In 2018, the device had been granted a Breakthrough Device designation.

FDA product codes: IYO, LNH, QQB.

Medical Policy Statement

The safety and effectiveness of ultrasonic transient elastography (FibroScan®) for the evaluation and/or monitoring of individuals with chronic liver disease have been established. It may be considered a useful diagnostic option when indicated.

Magnetic resonance elastography for the diagnosis and management of advanced hepatic fibrosis or cirrhosis has been established. It may be considered a useful option when indicated.

Multiparametric MRI (LiverMultiScan) is considered a useful option for diagnosis and management of advanced hepatic fibrosis/cirrhosis.

The use of FibroSURE™ multianalyte assays (HCV FibroSURE, ASH FibroSURE, NASH FibroSURE) in chronic liver disease has been established. It may be considered a useful diagnostic option when indicated.

The use of other noninvasive imaging, including but not limited to acoustic radiation force impulse imaging (ARFI), or real-time tissue elastography, is considered experimental/ investigational for the evaluation and/or monitoring of patients with chronic liver disease. While these services may be safe, their clinical utility for this clinical indication has not been determined.

The peer reviewed medical literature has not demonstrated the clinical utility of other multianalyte assays with algorithmic analyses (eg, FIBROSpect II, Enhanced Liver Fibrosis Test) for the evaluation or monitoring of patients with chronic liver disease. Therefore, these services are experimental/investigational.

Inclusionary and Exclusionary Guidelines

Inclusions

Noninvasive Imaging Techniques:

- Ultrasound transient elastography (FibroSCAN®), using an FDA-approved probe (eg, S+ M+ or XL+ Probe), may be considered established for the evaluation and/or monitoring of chronic liver disease
- Magnetic resonance elastography (MRE) may be considered established for the diagnosis and/or management of advanced hepatic fibrosis or cirrhosis for:
    - Individuals with nonalcoholic fatty liver disease who have high risk for cirrhosis due to advanced age, obesity, diabetes, or alanine aminotransferase (ALT) level more than twice the upper limit of normal OR
    - Individuals with other established chronic liver diseases when ultrasound elastography cannot be performed or is nondiagnostic
- Multiparametric MRI (LiverMultiScan) is considered a useful option for the diagnosis and management of advanced hepatic fibrosis/cirrhosis when diagnostic testing such as an ultrasound is inconclusive or non-diagnostic.

Multianalyte Assays:

- A FibroSURE™ multianalyte assay (either HCV FibroSURE™, ASH FibroSURE™ or NASH FibroSURE™) may be considered established for the evaluation and/or monitoring of chronic liver disease

Exclusions

Noninvasive Imaging Techniques:

- Ultrasound transient elastography in individuals with ascites
- Acoustic radiation force impulse imaging (ARFI)
- Real-time tissue elastography
- Use of ultrasound elastography to differentiate benign from malignant liver lesions

Multianalyte Assays:

- Multianalyte assays with algorithmic analyses for the evaluation or monitoring of patients with chronic liver disease not listed above (eg, Fibrospect, ELF, etc. -- this is not a complete list)

CPT/HCPCS Level II Codes *(Note: The inclusion of a code in this list is not a guarantee of coverage. Please refer to the medical policy statement to determine the status of a given procedure.)*

Established codes:

76391	81596	76981	76982	76983	87467
91200	0002M	0003M	0648T	0649T	

Other codes (investigational, not medically necessary, etc.):

76498	81517	81599	84999

Note: Code(s) may not be covered by all contracts or certificates. Please consult customer or provider inquiry resources at BCBSM or BCN to verify coverage.

Rationale

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

NONINVASIVE TESTING FOR CHRONIC LIVER DISEASE

Liver biopsy is an imperfect reference standard. There is a high rate of sampling error in biopsy, which can lead to underdiagnosis of liver disease.(1,2) These errors will bias estimates of performance characteristics of the noninvasive tests to which it is compared and must be considered in apprising the body of evidence. Mehta et al (2009) estimated that, under the best

scenario where sensitivity and specificity of liver biopsy are 90% and the prevalence of significant disease (increased liver fibrosis, scored as Metavir ≥ F2) is 40%; a perfect alternative marker would have calculated the area under the receiver operating characteristic (AUROC) curve of 0.90.(3) Therefore, the effectiveness of alternative technologies may be underestimated. In fact, when the accuracy of biopsy is presumed to be 80%, a comparative technology with an AUROC curve of 0.76 may actually have an AUROC curve of 0.93 to 0.99 for diagnosing true disease.

Due to the large number of primary studies published on this topic, this evidence review focuses on systematic reviews when available. The validation of multiple noninvasive tests will be assessed individually in the following sections. In this section, systematic reviews that compare several noninvasive tests will be discussed. Although options exist for performing systematic reviews with imperfect reference standards(4), most available reviews did not use any correction for the imperfect reference.

A systematic review by Crossan et al (2015) was performed for the National Institute for Health Research.(5) The first objective of the review was to determine the diagnostic accuracy of different noninvasive liver tests compared with liver biopsy in the diagnosis and monitoring of liver fibrosis and cirrhosis in patients with hepatitis C virus (HCV), hepatitis B virus (HBV), nonalcoholic fatty liver disease (NAFLD), and alcoholic liver disease (ALD). Reviewers selected 302 publications and presentations from 1998 to April 2012. Patients with HCV were the most common population included in the studies while patients with ALD were the least common. FibroScan and FibroTest were the most commonly assessed tests across liver diseases. Aminotransferase to platelet ratio index (APRI) was also widely assessed in HBV and HCV but not in NAFLD or ALD. The estimates of diagnostic accuracy for each test by disease are discussed in further detail in the following sections. Briefly, for diagnosing significant fibrosis (stage ≥F2) in HCV, the summary sensitivities and specificities were: FibroScan, 79% and 83%; FibroTest, 68% and 72%; APRI (low cutoff), 82% and 57%; acoustic radiation force impulse imaging (ARFI), 85% and 89%; HepaScore, 73% and 73%, FIBROSpect II, 78% and 71%; and FibroMeter, 79% and 73%, respectively. For diagnosing advanced fibrosis in HBV, the summary sensitivities and specificities were: FibroScan, 71% and 84%; FibroTest, 66% and 80%, respectively. There are no established or validated cutoffs for fibrosis stages across the diseases for most tests. For FibroTest, established cutoffs exist but were used inconsistently across studies. Test failures or reference standard(s) were frequently not captured in analyses. Most populations included in the studies were from tertiary care settings that have more advanced disease than the general population, which would overestimate the prevalence of the disease and diagnostic accuracy. These issues likely cause overestimates of sensitivities and specificities. The quality of the studies was generally rated as poor, with only 1.6% receiving a high-quality rating.

Houot et al (2016) reported on a systematic review funded by BioPredictive, the manufacturer of FibroTest.(6) This review included 71 studies published between January 2002 to February 2014 with over 12,000 participants with HCV and HBV comparing the diagnostic accuracy of FibroTest, FibroScan, APRI, and fibrosis-4 (FIB-4) index. Included studies directly compared the tests and calculated median differences in the AUROC curve using Bayesian methods. There was no evaluation of the methodologic quality of the included studies. The Bayesian difference in AUROC curve for significant fibrosis (stage ≥F2) between FibroTest and FibroScan was based on 15 studies and estimated to be 0.06 (95% credible interval [CrI], 0.02 to 0.09) favoring FibroTest. The difference in AUROC curve for cirrhosis for FibroTest versus FibroScan was based on 13 studies and estimated to be 0.00 (95% CrI, 0.04 to 0.04). The

difference for advanced fibrosis between FibroTest and APRI was based on 21 studies and estimated to be 0.05 (95% CrI, 0.03 to 0.07); for cirrhosis, it was based on 14 studies and estimated to be 0.05 (95% CrI, 0.00 to 0.11), both favoring FibroTest.

FIBROSURE SERUM PANEL

Clinical Context and Test Purpose

The purpose of noninvasive testing in individuals with chronic liver disease is to detect liver fibrosis so that patients can avoid the potential adverse effects of an invasive liver biopsy and receive appropriate treatment. The degree of liver fibrosis is an important factor in determining the appropriate approach for managing patients with liver disease (eg, hepatitis, ALD, NAFLD).

The question addressed in this portion of the evidence review is: Does use of the FibroSURE multianalyte serum assay for detecting liver fibrosis improve the health outcomes in individuals with chronic liver disease?

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with chronic liver disease.

Interventions

The test being considered is FibroSURE serum panel.

Comparators

The following tests and practices are currently being used to diagnose chronic liver disease: liver biopsy, noninvasive radiologic methods, and other multianalyte serum assays.

Outcomes

The general outcomes of interest are test validity, morbid events, and treatment-related morbidity. Follow-up over months to years is of interest for the relevant outcomes.

Study Selection Criteria

For the evaluation of clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores)
- Included a suitable reference standard (describe the reference standard)
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

Hepatitis C Virus

Clinical Validity

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Following the initial research into FibroSURE (patients with liver fibrosis who had undergone biopsy)(7), the next step in the development of this test was the further evaluation of the

algorithm in a cross-section of patients, including patients with hepatitis C virus (HCV) participating in large clinical trials before and after the initiation of antiviral therapy. A study by Poynard et al (2003) focused on patients with HCV who were participating in a randomized study of pegylated interferon and ribavirin.(8) From the 1530 participants, 352 patients with stored serum samples and liver biopsies at study entry and at 24-week follow-up were selected. The HCV FibroSURE score was calculated and then compared with the Metavir liver biopsy score. At a cutoff of 0.30, the HCV FibroSURE score had 90% sensitivity and 88% positive predictive value (PPV) for the diagnosis of Metavir F2 to F4 fibrosis; the specificity was 36%, and the negative predictive value (NPV) was 40%.

Poynard et al (2004) also evaluated discordant results in 537 patients who underwent liver biopsy and the HCV FibroSURE and ActiTest on the same day; discordance was attributed to either the limitations in the biopsy or serum markers.(9) In this study, cutoff values were used for individual Metavir scores (ie, F0 to F4) and for combinations of Metavir scores (ie, F0 to F1, F1 to F2). The definition of a significant discordance between FibroTest and ActiTest and biopsy scores was at least 2 stages or grades in the Metavir system. Discordance was observed in 29% of patients. Risk factors for failure of HCV FibroSURE scoring system were as follows: the presence of hemolysis, inflammation, possible Gilbert syndrome, acute hepatitis, drugs inducing cholestasis, or an increase in transaminases. Discordance was attributable to markers in 2.4% of patients, to the biopsy in 18%, and unattributed in 8.2% of patients. As noted in two reviews, the bulk of the research on HCV FibroSURE was conducted by researchers with an interest in the commercialization of the algorithm.(10,11)

In the Crossan et al (2015) systematic review, FibroTest was the most widely validated commercial serum test.(5) Seventeen studies were included in the pooled estimate of the diagnostic accuracy of FibroTest for significant fibrosis (stage ≥F2) in HCV. With varying cutoffs for positivity between 0.32 and 0.53, the summary sensitivity in HCV was 68% (95% confidence interval [CI], 58% to 77%) and specificity was 72% (95% CI, 70% to 77%). Eight studies were included for cirrhosis (stage F4) in HCV. The cutoffs for positivity ranged from 0.56 to 0.74 and the summary sensitivity and specificity were 60% (95% CI, 43% to 76%) and 86% (95% CI, 81% to 91%), respectively. Uninterpretable results were rare for tests based on serum markers.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials (RCTs). The primary benefit of the FibroSURE (FibroTest in Europe) for HCV is the ability to avoid liver biopsy in patients without significant fibrosis. There are currently no such published studies to demonstrate the effect on patient outcomes.

The FibroTest has been used as an alternative to biopsy for the purposes of establishing trial eligibility in terms of fibrosis or cirrhosis; several trials with FibroTest (ION-1,-3; VALENCE; ASTRAL-2, -3, -4) have established the efficacy of HCV treatments.(12-17) For example, in the

ASTRAL-2 and -3 trials, cirrhosis could be defined by a liver biopsy; a FibroScan or a FibroTest score of more than 0.75; or an APRI of more than 2.

These tests also need to be adequately compared with other noninvasive tests of fibrosis to determine their comparative efficacy. In particular, the proprietary, algorithmic tests should demonstrate superiority to other readily available, nonproprietary scoring systems to demonstrate that the tests improve health outcomes.

The FibroSURE test also has potential effect on patient outcomes as a means to follow response to therapy. In this case, evidence needs to demonstrate that use of the test for response to therapy impacts decision making and that these changes in management decisions lead to improved outcomes. It is not clear whether HCV FibroSURE could be used as an interval test in patients receiving therapy to determine whether an additional liver biopsy is necessary.

Alcoholic Liver Disease and Alcoholic Steatohepatitis

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

The diagnostic value of FibroSURE (FibroTest in Europe) has also been evaluated for the prediction of liver fibrosis in patients with ALD and NAFLD.(18,19) Thabut et al (2006) reported the development of a panel of biomarkers (ASH FibroSURE [ASH Test]) for the diagnosis of alcoholic steatohepatitis (ASH) in patients with chronic ALD.(20) Biomarkers were initially assessed with a training group consisting of 70 patients, and a panel was constructed using a combination of the 6 biochemical components of the FibroTest-ActiTest plus aspartate aminotransferase (AST). The algorithm was subsequently studied in 2 validation groups (1 prospective study for severe ALD, 1 retrospective study for non-severe ALD) that included 155 patients and 299 controls. The severity of ASH (none, mild, moderate, severe) was blindly assessed from biopsy samples. In the validation groups, there were 28 (18%) cases of discordance between the diagnosis of ASH predicted by the ASH Test and biopsy; 10 (36%) were considered to be false negatives of the ASH Test, and 11 were suspected failures of biopsy. Seven cases were indeterminate by biopsy. The AUROC curves were 0.88 and 0.89 in the validation groups. The median ASH Test value was 0.005 in controls, 0.05 in patients without or with mild ASH, 0.64 in the moderate ASH grade, and 0.84 in severe ASH grade 3. Using a cutoff value of 0.50, the ASH Test had sensitivity of 80% and specificity of 84%, with PPVs and NPVs of 72% and 89%, respectively.

Several authors have an interest in the commercialization of this test, and no independent studies on the diagnostic accuracy of ASH FibroSURE (ASH Test) were identified. In addition, it is not clear if the algorithm used in this study is the same as that used in the currently commercially available test, which includes 10 biochemicals.

FibroTest has been studied in patients with ALD. In the Crossan (2015) systematic review, 1 study described the diagnostic accuracy of FibroTest for significant fibrosis (stage ≥F2) or cirrhosis in ALD.(5) With a high cutoff for positivity (0.7) the sensitivity and specificity for advanced fibrosis were 55% (95% CI, 47% to 63%) and 93% (95% CI, 85% to 97%) and for cirrhosis were 91% (95% CI, 82% to 96%) and 87% (95% CI, 81% to 91%), respectively. With

a low cutoff for positivity (0.3) the sensitivity and specificity for advanced fibrosis were 84% (95% CI, 77% to 89%) and 65% (95% CI, 55% to 75%), respectively. The sensitivity and specificity for cirrhosis were 100% (95% CI, 95% to 100%) and 50% (95% CI, 42% to 58%), respectively.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

No studies were identified that assessed clinical outcomes following use of ASH FibroSURE (ASH Test) in ALD and ASH.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Poynard et al (2006) reported the development of a panel of biomarkers (NASH FibroSURE[NASH Test]) for the prediction of nonalcoholic steatohepatitis (NASH) in patients with NAFLD.(21) Biomarkers were initially assessed with a training group of 160 patients, and a panel was constructed using a combination of 13 of 14 parameters of the currently available test. The algorithm was subsequently studied in a validation group of 97 patients and 383 controls. Patients in the validation group were from a prospective multicenter study with hepatic steatosis at biopsy and suspicion of NAFLD. Histologic diagnoses used Kleiner et al's scoring system, with 3 classes for NASH (NASH, borderline NASH, no NASH). The main endpoint was steatohepatitis, defined as a histologic NASH score of 5 or greater. The AUROC curve for the validation group was 0.79 for the diagnosis of NASH, 0.69 for the diagnosis of borderline NASH, and 0.83 for the diagnosis of no NASH. Results showed sensitivity of 33% and specificity of 94% for NASH with a PPV and NPV of 66% and 81%, respectively. For borderline NASH or no NASH, sensitivity was 88%, specificity 50%, PPV 74%, and NPV72%. Clinically significant discordance (2 class difference) was observed in 8 (8%) patients. None of the 383 controls were considered to have NASH by NASH FibroSURE (NASH Test). Authors propose that this test would be suitable for mass screening for NAFLD in patients with obesity and diabetes.

An independent study by Lassailly et al (2011) attempted to prospectively validate the NASH Test (along with the FibroTest, SteatoTest, and ActiTest) in a cohort of 288 patients treated with bariatric surgery.(20) Included were patients with severe or morbid obesity (body mass

index, >35 kg/m^2), at least 1 co-morbidity for at least 5 years, and resistance to medical treatment. Excluded were patients with current excessive drinking, long-term consumption of hepatotoxic drugs, and positive screening for chronic liver diseases including hepatitis. Histology and biochemical measurements were centralized and blinded to other characteristics. The NASH Test provided a 3-category score for no NASH (0.25), possible NASH (0.50), and NASH (0.75). The prevalence of NASH was 6.9%, while the prevalence of NASH or possible NASH was 27%. The concordance rate between histologic NAS and the NASH Test was 43.1% with a weak kappa reliability test (0.14). In 183 patients who were categorized as possible NASH by the NASH Test, 124 (68%) were classified as no NASH by biopsy. In 15 patients categorized as NASH by the NASH Test, 7 (47%) were no NASH and 4 (27%) were possible NASH by biopsy. The NPV of the NASH Test for possible NASH or NASH was 47.5%. Authors suggested that the power of this study to validate agreement between the NASH Test and biopsy was low, due to the low prevalence of NASH. However, the results showed poor concordance between the NASH Test and biopsy, particularly for intermediate values.

In the Crossan (2015) systematic review, 4 studies were included in the pooled estimate of the diagnostic accuracy of FibroTest for advanced fibrosis (stage ≥ 3) in NAFLD.(5) The summary sensitivities and specificities were 40% (95% CI, 24% to 58%) and 96% (95% CI, 91% to 98%). Only 1 study included reported accuracy for cirrhosis, with sensitivity and specificity of 74% (95% CI, 54%, to 87%) and 92% (95% CI, 88% to 95%), respectively.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No studies were identified that assessed clinical outcomes following use of NASH FibroSURE (NASH Test) in NAFLD and NASH.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Hepatitis B Virus

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
While most multianalyte assay studies that have identified fibrosis have been in patients with HCV, studies are also being conducted in patients with chronic HBV.(22,23) In a study, Park et al (2013) compared liver biopsy with the FibroTest results obtained on the same day from 330 patients who had chronic HBV.(24) Discordance was found in 30 (9.1%) patients for whom the

FibroTest underestimated fibrosis in 25 patients and overestimated it in 5 patients. Those with Metavir liver fibrosis stages F3 or F4 (15.4%) had a significantly higher discordance rate than with stages F1 or F2 (3.0%; p<.001). The only independent factor for discordance on multivariate analysis was a Metavir stages F3 or F4 on liver biopsy (p<.001).

Salkic et al (2014) conducted a meta-analysis of studies on the diagnostic accuracy of FibroTest in chronic HBV.(25) Included in the meta-analysis were 16 studies (n=2494) on liver fibrosis diagnosis and 13 studies (n=1754) on cirrhosis diagnosis. There was strong evidence of heterogeneity in the 16 fibrosis studies and evidence of heterogeneity in the cirrhosis studies. For significant liver fibrosis (Metavir F2 to F4) diagnosis using all of the fibrosis studies, the AUROC curve was 0.84 (95% CI, 0.78 to 0.88). At the recommended FibroTest threshold of 0.48 for a significant liver fibrosis diagnosis, the sensitivity was 60.9%, specificity was 79.9%, and the diagnostic odds ratio (OR) was 6.2. For liver cirrhosis (Metavir F4) diagnosis using all of the cirrhosis studies, the AUROC curve was 0.87 (95% CI, 0.85 to 0.9). At the recommended FibroTest threshold of 0.74 for cirrhosis diagnosis, the sensitivity was 61.5%, specificity was 90.8%, and the diagnostic OR was 15.7. While the results demonstrated FibroTest may be useful in excluding a diagnosis of cirrhosis in patients with chronic HBV, the ability to detect significant fibrosis and cirrhosis and exclude significant fibrosis is suboptimal.

Xu et al (2014) reported on a systematic review and meta-analysis of studies on biomarkers to detect fibrosis in HBV.(26) Included in the analysis on FibroTest were 11 studies (N=1640 patients). In these 11 studies, AUROC curves ranged from 0.69 to 0.90. Heterogeneity in the studies was statistically significant.

In the Crossan (2015) systematic review, 6 studies were included in the pooled estimate of the diagnostic accuracy of FibroTest for significant fibrosis (stage ≥F2) in HBV.(5) The cutoffs for positivity ranged from 0.40 to 0.48, and the summary sensitivities and specificities were 66% (95% CI, 57% to 75%) and 80% (95% CI, 72% to 86%), respectively. The accuracy for diagnosing cirrhosis in HBV was based on 4 studies with cutoffs for positivity ranging from 0.58 to 0.74; sensitivities and specificities were 74% (95% CI, 25% to 96%) and 90% (95% CI, 83% to 94%), respectively.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

There are no studies of the effect on patient outcomes for patients with HBV. Of note, some researchers have noted that different markers (eg, HBV FibroSURE) may be needed for this assessment in patients with hepatitis B.(27)

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: FibroSURE Serum Panel
For individuals who have chronic liver disease who receive FibroSURE serum panels, the evidence includes systematic reviews of more than 30 observational studies (>5000 patients). FibroSURE has been studied in populations with viral hepatitis, NAFLD, and ALD. There are established cutoffs, although they were not consistently used in validation studies. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. However, for the purposes of deciding whether a patient has severe fibrosis or cirrhosis, FibroSURE results provide data sufficiently useful to determine therapy. Specifically, FibroSURE has been used as an alternative to biopsy to establish eligibility regarding the presence of fibrosis or cirrhosis in several RCTs that showed the efficacy of HCV treatments, which in turn demonstrated that the test can identify patients who would benefit from therapy.

MULTIANALYTE SERUM ASSAYS OTHER THAN FIBROSURE

Clinical Context and Test Purpose
The purpose of noninvasive testing in individuals with chronic liver disease is to detect liver fibrosis so that patients can avoid the potential adverse events of an invasive liver biopsy and receive appropriate treatment. The degree of liver fibrosis is an important factor in determining the appropriate approach for managing patients with liver disease (eg, hepatitis, ALD, NAFLD).

The question addressed in this portion of the evidence review is: Does the use of the multianalyte serum assays (other than FibroSURE) for detecting liver fibrosis improve the net health outcome in patients with chronic liver disease?

The following PICO was used to select literature to inform this review.

Populations
The relevant population of interest is individuals with chronic liver disease.

Interventions
The tests being considered are multianalyte serum assays (other than FibroSURE).

Comparators
The following tests and practices are currently being used to diagnose chronic liver disease: liver biopsy, noninvasive radiologic methods, and other multianalyte serum assays.

Outcomes
The general outcomes of interest are test validity, morbid events, and treatment-related morbidity. Follow-up over months to years is of interest to the relevant outcomes.

Study Selection Criteria
For the evaluation of the clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:
- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores).
- Included a suitable reference standard (describe the reference standard).
- Patient/sample clinical characteristics were described.
- Patient/sample selection criteria were described.

FIBROSpect II

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Patel et al (2004) investigated the use of these serum markers in an initial training set of 294 patients with HCV and further validated the resulting algorithm in a validation set of 402 patients.(28) The algorithm was designed to distinguish between no or mild fibrosis (F0 to F1) and moderate-to-severe fibrosis (F2 to F4). With the prevalence of F2 to F4 disease of 52% and a cutoff value of 0.36, the PPVs and NPVs were 74.3% and 75.8%, respectively.

The published studies for this combination of markers continue to focus on test characteristics such as sensitivity, specificity, and accuracy.(29-31) In Crossan et al (2015), the summary diagnostic accuracy for detecting significant fibrosis (stage ≥F2) in 5 studies of HCV with FIBROSpect II with cutoffs ranging from 42 to 72 was 78% (95% CI, 49% to 93%) and the summary specificity was 71% (95% CI, 59% to 80%).(5)

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

The issues of effect on patient outcomes are similar to those discussed for the FibroSURE (FibroTest in Europe). No studies were identified in the published literature in which the results of the FIBROSpect test were actively used in the management of the patient.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.
Because the clinical validity of FIBROSpect has not been established, a chain of evidence supporting the clinical utility of this test for this population cannot be constructed.

Other Multianalyte Scoring Systems

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Other scoring systems have been developed, including FIB-4, NAFLD fibrosis score (NFS), APRI, AST/ALT ratio, combined body mass index, AST/ALT ratio and diabetes status (BARD), and Enhanced Liver Fibrosis (ELF). The ELF test combines measurements of biomarkers into a proprietary algorithm to produce a score. The other scoring systems use a simple

nonproprietary formula that can be calculated at the bedside to produce a score for the prediction of fibrosis. Tables 1 and 2 summarize the characteristics and results of systematic reviews that have assessed the diagnostic accuracy of various noninvasive scoring systems. There are no established cutoffs for ruling in or ruling out advanced fibrosis for most tests. In the systematic reviews, 2 cutoffs were analyzed for each test (as selected by the authors); a lower threshold to rule out advanced fibrosis and a higher threshold to rule in advanced fibrosis. Patients that fall between the 2 thresholds are classified as "indeterminate" risk for whom a liver biopsy may be considered. Castellana et al (2021) conducted an meta-analytic head-to-head comparison between FIB-4 and NFS and found no significant differences regarding relative diagnostic OR, positive likelihood ratio, and negative likelihood ratio.(32) FIB-4 was associated with fewer indeterminate findings compared to NFS. Mozes et al (2021) found that FibroScan, a transient elastography test, outperformed all of the serum-based tests.(33) Sharma et al (2021) qualitatively evaluated the diagnostic performance of ELF in patients with chronic liver disease.(34)

Table 1. Characteristics of Systematic Reviews Assessing NonInvasive Scoring Systems

Study	Dates	Studies	N (range)	Population	Index Tests	Reference Standard
Castellana et al (2021)[32,]	2012-2020	18	12,604 (102 to 3202)	NAFLD	FIB-4 NFS	Histology
Mozes et al (2021)[33,]	Up to 2020	37	5735 (13 to 1063)	NAFLD	FibroScan FIB-4 NFS APRI AST/ALT	Histology
Sharma et al (2021)[34,]	Up to 2020	36	NR (38 to 3202)	Chronic liver disease (NAFLD, ALD, hepatitis, mixed etiologies)	ELF	Histology

ALD: alcoholic liver disease; ALT: alanine aminotransferase; APRI: AST-to-platelet ratio; AST: aspartate aminotransferase; ELF: Enhanced Liver Fibrosis; FIB-4: fibrosis-4 index; NAFLD: nonalcoholic fatty liver disease; NFS: NAFLD fibrosis score; NR: not reported.

Table 2. Results of Systematic Reviews Assessing the Diagnostic Accuracy of NonInvasive Scoring Systems

Index Test (Threshold)	Studies/Sample Size	Index Test Threshold (low, high)	AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)
Castellana et al (2021)[32,]			Advanced Fibrosis (ie, Stages F3 to F4)
FIB-4	14 (9968)	1.3, 2.67	NR 65% (51% to 77%) 93% (89% to 96%)
NFS	14 (9113)	-1.455, 0.676	NR 61% (45% to 76%) 93% (89% to 96%)
Mozes et al (2021)[33,]			Advanced Fibrosis (ie, Stages F3 to F4)
FibroScan	NR (5489)	7.4, 12.1	0.85 (0.84 to 0.86) 84% (81% to 87%) 87% (85% to 88%)

FIB-4	NR (5393)	0.88, 2.31	0.76 (0.74 to 0.77) 80% (76% to 83%) 79% (77% to 81%)
NFS	NR (3248)	-2.55, 0.28	0.73 (0.71 to 0.75) 74% (70% to 79%) 78% (76% to 81%)
APRI	NR (5477)		0.70 (0.69 to 0.72)[a] NE NE
AST/ALT	NR (5434)		0.64 (0.62 to 0.65)[a] NE NE
Sharma et al (2021)[34,]			Advanced Fibrosis
ELF - HCV	11 (NR)	Varied among studies	AUROC range, 0.773 (0.697 to 0.848) to 0.98 (0.93 to 1.00)
ELF - HBV	4 (NR)	Varied among studies	AUROC range, 0.69 (0.63 to 0.75) to 0.86 (0.81 to 0.92)
ELF - NAFLD	7 (NR)	Varied among studies	AUROC range, 0.78 (0.70 to 0.89) to 0.97 (no CI reported)
ELF - ALD	3 (NR)	Varied among studies	AUROC range, 0.92 (0.89 to 0.96) to 0.944 (0.836 to 1.000)
ELF - mixed etiology	7 (NR)	Varied among studies	AUROC range, 0.63 (no CI reported) to 0.91 (0.88 to 0.95)

ALD: alcoholic liver disease; ALT: alanine aminotransferase; APRI: aminotransferase-to-platelet ratio index; AST: aspartate aminotransferase; AUROC: area under the receiver operating characteristic; CI: confidence interval; ELF: enhanced liver fibrosis; FIB-4: fibrosis-4 index; HBV: hepatitis B virus; HCV: hepatitis C virus; NAFLD: nonalcoholic fatty liver disease; NE: not evaluated; NFS: NAFLD fibrosis score; NR: not reported.
a Diagnostic performance not further evaluated after modest performance on AUROC.

The APRI requires only the serum level of AST and the number of platelets as part of its calculation.(35) Using an optimized cutoff value derived from a training set and validation set of patients with HCV, authors have reported that the NPV for fibrosis was 86% and that the PPV was 88%. In Crossan et al (2015), APRI was frequently evaluated and has been tested in HCV, HBV, NAFLD, and ALD.(5) The summary diagnostic accuracies are in Table 3.

Table 3. Diagnostic Accuracy for Aminotransferase to Platelet Ratio Index

Disease	Metavir Stage	Cutoff	Studies	Sensitivity, % (95% CI)	Specificity, % (95% CI)
HCV	≥F2 (significant)	Low: 0.4 to 0.7	47	82 (77 to 86)	57 (49 to 65)
HCV	≥F2 (significant)	High: 1.5	36	39 (32 to 47)	92 (89 to 95)
HCV	F4 (cirrhosis)	Low: 0.75 to 1	24	77 (73 to 81)	78 (74 to 81)
HCV	F4 (cirrhosis)	High: 2	19	48 (41 to 56)	94 (91 to 95)
HBV	≥F2 (significant)	Low: 0.4 to 0.6	8	80 (68 to 88)	65 (52 to 77)
HBV	≥F2 (significant)	High: 1.5	6	37 (22 to 55)	93 (85 to 97)
HBV	F4 (cirrhosis)	Low: 1	4	58 (49 to 66)	76 (70 to 81)
HBV	F4 (cirrhosis)	High: 2	3	24 (8 to 52)	91 (83 to 96)

Disease	Metavir Stage	Cutoff	Studies	Sensitivity, % (95% CI)	Specificity, % (95% CI)
NAFLD	≥F3 (significant)	0.5 to 1.0	4	40 (7 to 86)	82 (78 to 60)
NAFLD	F4 (cirrhosis)	0.54 and NA	2	78 (71 to 99)	71 (30 to 93)
ALD	≥F2 (significant)	Low: 0.5	2	72 (60 to 82)	46 (33 to 60)
ALD	≥F2 (significant)	High: 1.5	2	54 (42 to 66)	78 (64 to 88)
ALD	F4 (cirrhosis)	High: 2.0	1	40 (22 to 61)	62 (41 to 79)

Adapted from Crossan et al (2015).(5)
ALD: alcoholic liver disease; APRI: aspartate aminotransferase-platelet ratio index; CI: confidence interval; HBV: hepatitis B virus; HCV: hepatitis C virus; NA: not available; NAFLD: nonalcoholic fatty liver disease.

Giannini et al (2006) reported that the use of the AST/ALT ratio and platelet counts in a diagnostic algorithm would have avoided liver biopsy in 69% of patients with chronic hepatitis C and would have correctly identified the absence/presence of significant fibrosis in 80.5% of these cases.(36) In Crossan et al (2015), the cutoffs for the positivity of AST/ALT ratio for diagnosis of significant fibrosis (stage ≥F2) varied from 0.6 to 1 in 7 studies.(5) Summary sensitivity and specificity were 44% (95% CI, 27% to 63%) and 71% (95% CI, 62% to 78%), respectively. Thirteen studies used a cutoff of 1 to estimate diagnostic accuracy of cirrhosis with AST/ALT ratio, and summary sensitivity and specificity were 49% (95% CI, 39% to 59%) and 87% (95% CI, 75% to 94%), respectively.

A number of studies have compared HCV FibroSURE (FibroTest) and other noninvasive tests of fibrosis with biopsy using receiver operating characteristic (ROC) analysis. For example, Burlier et al (2006) reported on the validation of FibroSURE(FibroTest) and found that, based on ROC analysis, FibroSURE (FibroTest) was superior to APRI for identifying significant fibrosis, with AUROC curves of 0.81 and 0.71, respectively.(37) A prospective multicenter study by Sarkis et al (2012) compared 9 of the best-evaluated blood tests in 436 patients with HCV and found similar performance for HCV FibroSURE (FibroTest), FibroMeter, and HepaScore (ROC curve, 0.84, 0.86, 0.84, respectively).(38) These 3 tests were significantly superior to the 6 other tests, with 70% to 73% of patients considered well-classified according to a dichotomized score (F0/F1 vs ≥F2). The number of “theoretically avoided liver biopsies” for the diagnosis of significant fibrosis was calculated to be 35.6% for HCV FibroSURE (FibroTest). To improve diagnostic accuracy, algorithms that combine HCV FibroSURE (FibroTest) with other tests (eg, APRI) are also being evaluated.(38-40) One of these, the sequential algorithm for fibrosis evaluation, combines the APRI and FibroSURE (FibroTest). Crossan et al (2015) reported that the algorithm has been assessed in 4 studies of HCV for diagnosing both significant fibrosis (stage ≥F2) and cirrhosis.(5) Summary sensitivity and specificity for significant fibrosis were estimated to be 100% (95% CI, 100% to 100%) and 81% (95% CI, 80% to 83%), respectively. The summary sensitivity and specificity for cirrhosis were 74% (95% CI, 42% to 92%) and 93% (95% CI, 91% to 94%), respectively.

Rosenberg et al (2004) developed a scoring system based on an algorithm combining hyaluronic acid, amino terminal propertied of type III collagen, and tissue inhibitors of metalloproteinase 1.(41) This test is manufactured by Siemens Healthcare as the Enhanced Liver Fibrosis (ELF) Test.(42) The algorithm was developed in a test set of 400 patients with a wide variety of chronic liver diseases and then validated in another 521 patients. The algorithm was designed to discriminate between no or mild fibrosis and moderate-to-severe fibrosis. The NPV for fibrosis was 92%.

Younus et al (2021) evaluated the diagnostic value of ELF to assess liver fibrosis in patients with NAFLD.(43) This was a retrospective, cross-sectional study including 829 patients; 462 had transient elastography data and 463 had liver biopsy data. A significant increase in ELF scores was correlated in patients with advanced fibrosis by biopsy or transient elastography. The AUROC for ELF for identifying fibrosis was 0.81 (95% CI, 0.77 to 0.85) with biopsy as the reference standard and 0.79 (95% CI, 0.75 to 0.82) with transient elastography as the reference standard. Predictive combinations of ELF and FIB-4 scores were additionally evaluated. For ELF score ≥7.2 with a FIB-4 score ≥0.74, the sensitivity and NPV were 92.5% (95% CI, 87.4% to 97.5%) and 95.1% (95% CI, 91.8% to 98.4%), respectively, for ruling out fibrosis. For ELF score ≥9.8 with a FIB-4 score ≥2.9, the specificity and PPV were 99.7% (95% CI, 99.1% to 100%) and 95.0% (95% CI, 85.5% to 100%), respectively, for ruling in fibrosis.

The FIB-4 index was developed in a cohort of patients with HCV and is similar to APRI in that it uses a simple nonproprietary formula to produce a score for the prediction of fibrosis, incorporating patient age, AST level, ALT level, and platelet count. In the original cohort studied by Sterling et al (2006) (44), a low cutoff score of <1.45 had an NPV of 90% for advanced fibrosis whereas a high cutoff score >3.25 had a 97% specificity and PPV of 65% for advanced fibrosis. Overall, 70% of patients were stratified <1.45 or >3.25 and represented potential cases that could have avoided liver biopsy with a corresponding diagnostic accuracy of 86%. In a comparative study by Eagle-Picher et al (2007) in patients with HCV utilizing the same cutoff values, an NPV of 94.7% with a sensitivity of 74.3% and a specificity of 80.1% and a PPV of 82.1% with a specificity of 98.2% and sensitivity of 37.6% were reported.(45) When the diagnostic performance of FIB-4 was compared against FibroTest (Fibro Sure in the U.S.), the exclusion of severe fibrosis and the detection of severe fibrosis were found to agree between tests in 92.1% and 76.0% of cases, respectively.

Yan et al (2020) evaluated the diagnostic value of total bile acid-to-cholesterol ratio (TBA/TC) as a serum marker for cirrhosis and fibrosis in chronic HBV-infected patients without cholestasis.(46) This was a cross-sectional study including 667 patients. In a multivariate analysis, TBA/TC was independently correlated with cirrhosis in the study population (OR, 1.102, 95% CI, 1.085 to 1.166). ROC curve analyses yielded similar areas under the curve (AUCs) for TBA/TC, APRI, and FIB-4 at 0.87, 0.84, and 0.80, respectively. For diagnosing cirrhosis, the specificity and PPV of TBA/TC (83.33%, 91.10%) were higher than those of APRI (73.61%, 87.20%). The AUC of TBA/TC that distinguished significant liver cirrhosis was 2.70. In another multivariate analysis, TBA/TC was also independently correlated with significant fibrosis (OR, 1.040, 95% CI, 1.001 to 1.078). The AUC of TBA/TC that distinguished significant liver fibrosis was 0.70. Among 32 patients who also had a liver biopsy performed, TBA/TC was significantly higher in both fibrosis and cirrhosis as well as significantly correlated with fibrosis stage (p<.001 for all).

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs. The primary benefit of the multivariate serum assays is the ability to avoid liver biopsy.

A systematic review and meta-analysis conducted by Ciancia et al (2022) evaluated the use of noninvasive biomarkers for prediction of all-cause and cardiovascular mortality in patients with NAFLD.47, Of 24 studies included in the review, noninvasive scoring systems were assessed in 16 studies, 4 of which had adequate data for meta-analysis based on review criteria that required 2 or more studies reporting the same outcome measure using equivalent cut-off values and statistical methods in a similar study population. All of the studies included in the meta-analysis studies were retrospective (N=9,725; n range=320 to 4,680), and NAFLD diagnosis was based on liver biopsy or clinical diagnosis. Mean duration of follow-up ranged from 9 to 20 years in 3 of the studies and was not reported in the fourth study, but the total study duration was 17 years. A total of 1,697 deaths were reported in the 4 studies. Results of the meta-analyses appear in Table 4. Although high scores were associated with an increased risk of mortality relative to low scores across all scoring systems, the evidence is limited by the small number of included studies and high heterogeneity and imprecision for some estimates.

Table 4. Pooled Diagnostic Accuracy of Noninvasive Scoring Systems for Prediction of All-Cause and Cardiovascular Mortality in Patients with NAFLD

Scoring System	Number of Studies	Comparison (Score Cut-off)	Pooled HR (95% CI)
All-cause mortality			
NFS	4	High (>0.676) vs. Low (< -1.455)	3.07 (1.62 to 5.83; I^2=76%)
NFS	4	Intermediate (-1.455 to 0.676) vs. Low (< -1.455)	1.91 (1.18 to 3.09; I^2=82%
FIB-4	3	High (>2.67) vs. Low (<1.30)	3.06 (1.54 to 6.07; I^2=73%)
FIB-4	3	Intermediate (1.30 to 2.67) vs. Low (<1.30)	1.60 (1.33 to 1.91; I^2=0%)
APRI	3	High (>1.5) vs. Low (<0.5)	1.90 (1.32 to 2.73; I^2=0%)
APRI	3	Intermediate (0.5 to 1.5) vs. Low (<0.5)	0.98 (0.76 to 1.26; I^2=0%)
BARD	2	High (4) vs. Low (0 to 1)	2.87 (1.27 to 6.46; I^2=45%)
BARD	2	Intermediate (2 to 3) vs. Low (0 to 1)	1.64 (1.21 to 2.23; I^2=0%)
Cardiovascular mortality			
NFS	2	High (>0.676) vs. Low (< -1.455)	3.09 (1.78 to 5.34; I^2=0%)

NFS	2	Intermediate (-1.455 to 0.676) vs. Low (< -1.455)	2.12 (1.41 to 3.17; I^2=0%)

Adapted from Ciancia et al 2022[47,]
ALT: alanine aminotransferase; APRI: aminotransferase-to-platelet ratio index; AST: aspartate aminotransferase; BARD: body mass index, AST/ALT ratio and diabetes status; CI: confidence interval; FIB-4: fibrosis-4 index; HR: hazard ratio; NAFLD: nonalcoholic fatty liver disease; NFS: NAFLD fibrosis score

Sanyal et al (2019) reported on findings of two, phase 2b, placebo-controlled trials of nimotuzumab in NASH in patients with bridging fibrosis (F3; n=217) or compensated cirrhosis (F4; n=258) that assessed patients with liver biopsy and serum biomarker tests, including ELF, APRI, Fibro Sure/FibroTest, and the FIB-4 index.(48) Laboratory screening was conducted at baseline and at every 3 months during the course of the trials. The trials were terminated after 96 weeks due to nimotuzumab inefficacy, at which point data from treatment groups were combined for analysis. In patients with bridging fibrosis, increased risk of progression to cirrhosis was observed with higher baseline levels of all serum fibrosis tests (p<.001). Change in the ELF score over time was also associated with progression to cirrhosis (p<.001). For a cutoff score of 9.76, progression to cirrhosis had a reported hazard ratio (HR) of 4.12 (95% CI: 2.14 to 7.93; p<.001). For patients with compensated cirrhosis, higher levels of baseline biomarker tests were also associated with liver-related clinical events in 19% of patients, such as ascites, hepatic encephalopathy, newly diagnosed varices, esophageal variceal bleed, increase in Child-Pugh and/or Model for end-stage liver disease (MELD) score, or death (p<.001 to .006). While the manufacturer of the test differentiates moderate from severe fibrosis with a cutoff ELF score of 9.8, current National Institute for Health and Care Excellence guidelines for NAFLD recommend reserving a diagnosis of advanced fibrosis to NAFLD patients with an ELF score of 10.51 or greater, limiting the clinical significance of these findings.(49) Furthermore, serum fibrosis test results were not directly used in patient management in the nimotuzumab trials.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Multianalyte Serum Assays Other Than FibroSURE
For individuals who have chronic liver disease who receive multianalyte serum assays for liver function assessment other than FibroSURE, the evidence includes a number of observational studies and systematic reviews of those studies. Studies have frequently included varying cutoffs, some of which were standardized and others not validated. Cutoff thresholds have often been modified over time, may be specific to certain patient populations, and in some cases, guideline recommendations differ from cutoffs designated by manufacturers and those utilized in studies. A comparison of transient elastography to various serum-based tests found that the former were superior in detecting fibrosis, and a meta-analysis of 4 studies found higher multianalyte scores associated with an increased risk of mortality relative to lower scores, but the evidence is limited by the small number of included studies and high heterogeneity and imprecision for some estimates. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. There is no direct evidence that other multianalyte serum assays improve health outcomes; further, it is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence on clinical validity. FIBROSpect II has been studied in populations with HCV. Cutoffs for positivity varied across studies and were not well validated. The methodologic quality of the

validation studies was generally poor. There is no direct evidence that FIBROSpect II improves health outcomes.

NONINVASIVE IMAGING: TRANSIENT ELASTOGRAPHY

Clinical Context and Test Purpose

The purpose of noninvasive testing in individuals with chronic liver disease is to detect liver fibrosis so that patients can avoid the potential adverse events of an invasive liver biopsy and receive appropriate treatment. The degree of liver fibrosis is an important factor in determining the appropriate approach for managing patients with liver disease (eg, hepatitis, ALD, NAFLD).

The question addressed in this portion of the evidence review is: Does the use of transient elastography for detecting liver fibrosis improve the net health outcome in patients with chronic liver disease?

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with chronic liver disease.

Interventions

The test being considered is transient elastography.

Comparators

The following tests and practices are currently being used to diagnose chronic liver disease: liver biopsy, other noninvasive radiologic methods, and multianalyte serum assays.

Outcomes

The general outcomes of interest are test validity, morbid events, and treatment-related morbidity. Follow-up over months to years is of interest to the relevant outcomes.

Study Selection Criteria

For the evaluation of the clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores).
- Included a suitable reference standard (describe the reference standard).
- Patient/sample clinical characteristics were described.
- Patient/sample selection criteria were described.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

There is extensive literature on the use of transient elastography (eg, FibroScan) to gauge liver fibrosis and cirrhosis. Summaries of systematic reviews are shown in Tables 5 and 6. Brener et al (2015) performed a health technology assessment summarizing many of the systematic reviews below.(50) The assessment focused on reviews of the diagnostic accuracy and effect

on patient outcomes of transient elastography for liver fibrosis in patients with HCV, HBV, NAFLD, ALD, or cholestatic diseases. Fourteen systematic reviews of transient elastography with biopsy reference standard shown below were included in the Brener assessment, summarizing more than 150 primary studies.(51-64) There was variation in the underlying cause of liver disease and the cutoff values of transient elastography stiffness used to define Metavir stages in the systematic reviews. There did not appear to be a substantial difference in diagnostic accuracy for 1 disease over any other. The reviews demonstrated that transient elastography has good diagnostic accuracy compared to biopsy for the assessment of liver fibrosis and steatosis.

Crossan et al (2015) found that FibroScan was the noninvasive liver test most assessed in validation studies across liver diseases (37 studies in HCV, 13 in HBV, 8 in NAFLD, 6 in ALD).(5) Cutoffs for positivity for fibrosis staging varied between diseases and were frequently not prespecified or validated: HCV, 5.2 to 10.1 kilopascal (kPa) in the 37 studies for Metavir stages ≥F2; HBV, 6.3 to 8.9 kPa in 13 studies for stages ≥F2; NAFLD, 7.5 to 10.4 kPa in 8 studies for stages ≥F3; ALD, 11.0 to 12.5 kPa in 4 studies for stages ≥F3. Summary sensitivities and specificities by disease are shown in Table 6. The overall sensitivity and specificity for cirrhosis including all diseases (65 studies; cutoffs range, 9.2-26.5 kPa) were 89% (95% CI, 86% to 91%) and 89% (95 % CI, 87% to 91%), respectively. The rate of uninterpretable results, when reported, with FibroScan (due to <10 valid measurements; success rate, <60%; interquartile range, >30%) was 8.5% in HCV, and 9.6% in NAFLD.

Table 5. Transient Elastography Systematic Review Characteristics

Study	Dates	Studies	N	Population
Bota et al (2013)[51,]	To May 2012	13	1163	Chronic hepatitis
Cai et al (2021)[65,]	To Mar 2019	62	NR	ALD, NAFLD
Chon et al (2012)[52,]	2002 to Mar 2011	18	2772	HBV
Crossan et al (2015)[5,]	1998 to Apr 2012	66	NR	HCV, HBV, NAFLD, ALD
Friedrich-Rust et al (2008)[53,]	2002 to Apr 2007	50	11,275	All causes of liver disease
Geng et al (2016)[67,]	To Jan 2015	57	10,569	Multiple causes of liver disease
Jiang et al (2018)[68,]	To Dec 2017	11	1735	NAFLD
Kwok et al (2014)[54,]	To Jun 2013	22	1047	NAFLD
Li et al (2016)[69,]	Jan 2003 to Nov 2014	27	4386	HBV
Njei et al (2016)[70,]	To Jan 2016	6	756	HCV/HIV coinfection
Pavlov et al (2015)[71,]	To Aug 2014	14	834	ALD
Poynard et al (2011)[56,]	Feb 2001 to Dec 2010	18	2714	HBV
Shaheen et al (2007)[57,]	Jan 1997 to Oct 2006	12	1981	HCV
Shi et al (2014)[58,]	To May 2013	9	1771	All causes of steatosis
Steadman et al (2013)[59,]	2001 to Jun 2011	64	6028	HCV, HBV, NAFLD, CLD, liver transplant

Study	Dates	Studies	N	Population
Stebbing et al (2010)[60,]	NR, prior to Feb 2009	22	4625	All causes of liver disease
Talwalkar et al (2007)[61,]	To Jan 2027	9	2083	All causes of liver disease
Tsochatzis et al (2011)[62,]	To May 2009	40	7661	All causes of liver disease
Tsochatzis et al (2014)[63,]	1998 to Apr 2012	302	NR	HCV, HBV, ALD, NAFLD
Xu et al (2015)[72,]	To Dec 2013	19	3113	HBV
Xue-Ying (2020)[64,]	Jan 2008 to Dec 2018	81	32,694	HBV

ALD: alcoholic liver disease; CLD: chronic liver disease; HBV: hepatitis B virus; HCV: hepatitis C virus; HIV: human immunodeficiency virus; NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Table 6. Transient Elastography Systematic Reviews Diagnostic Accuracy Results

		Significant Fibrosis (ie, Metavir Stages F2 to F4)		Cirrhosis (ie, Metavir Stage F4)	
Study	**Population**	**Studies/ Sample Size**	**AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)**	**Studies/ Sample Size**	**AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)**
Bota et al 2013)[51,]	Multiple diseases	10/1016	0.87 (0.83 to 0.89) 78% (72% to 83%) 84% (75% to 90%)	13/1163	0.93 (0.91 to 0.95) 89% (80% to 94%) 87% (82% to 91%)
	HCV			4/NR	NR 92% (78% to 97%) 86% (82% to 90%)
Cai et al (2021)[65,]	ALD/NAFLD	40/2569	0.86 (0.83 to 0.89) 77% (73% to 81%) 82% (78% to 86%)	34/914	0.95 (0.92 to 0.96) 91% (87% to 94%) 86% (83% to 89%)
Chon et al (2012)[52,]	Chronic HBV	12/2000	0.86 (0.86 to 0.86) 74.3% (NR) 78.3% (NR)	16/2614	0.93 (0.93 to 0.93) 84.6% (NR) 81.5% (NR)
Crossan et al (2015)[5,]	HCV	37/NR	NR 79% (74% to 84%) 83% (77% to 88%)	36/NR	NR 89% (84% to 92%) 91% (89% to 93%)
	HBV	13/NR	NR 71% (62% to 78%) 84% (74% to 91%)	19/NR	NR 86% (79% to 91%) 85% (78% to 89%)
	NAFLD			4/NR	NR 96% (83% to 99%) 89% (85% to 92%)
	ALD	1/NR	NR 81% (70% to 88%) 92% (76% to 98%)	4/NR	NR 87% (64% to 96%) 82% (67% to 91%)
Friedrich-Rust (2008)[53,]	Multiple diseases	25/3685	0.84 (0.82 to 0.86) NR NR	25/4557	0.94 (0.93 to 0.95) NR NR
	HCV	NR	0.84 (0.80 to 0.86) NR NR		

Geng et al (2016)[67],	Multiple diseases				0.93 (NR) 81% (79% to 83%) 88% (87% to 89%)
Jiang et al (2018)[68],	NAFLD	10/NR	0.85 (0.82 to 0.88) 77% (70% to 84%) 80% (74% to 84%)	11/NR	0.96 (0.93 to 0.97) 90% (73% to 97%) 91% (87% to 94%)
Kwok et al (2014)[54],	NAFLD	7/800	0.83 (0.79 to 0.87) 0.79 (0.72 to 0.84) 0.75 (0.71 to 0.79)	57/10,569	0.96 (0.94 to 0.99) 92% (82% to 97%) 92% (86% to 98%)
Li et al (2016)[69],	HBV	19/NR	0.88 (0.85 to 0.91) 81% (76% to 85%) 82% (71% to 87%)	24/NR	0.93 (0.91 to 0.95) 86% (82% to 90%) 88% (84% to 90%)
Njei et al (2016)[70],	HCV/HIV	6/756	NR 97% (82% to 91%) 64% (45% to 79%)	6/756	NR 90% (74% to 91%) 87% (80% to 92%)
Pavlov et al (2015)[71],	ALD	7/338	NR 94% (86% to 97%) 89% (76% to 95%)	7/330	NR 95% (87% to 98%) 71% (56% to 82%)
Poynard et al (2011)[56],	HBV	4/NR	0.84 (0.78 to 0.89) NR NR	NR	0.93 (0.87 to 0.99) NR NR
Shaheen et al (2007)[57],	HCV	4/NR	0.84 (0.78 to 0.89) NR NR	NR	0.93 (0.87 to 0.99) NR NR
Shi et al (2014)[58],	No summary statistics reported. Concluded that transient elastography controlled attenuation parameter has good sensitivity and specificity for diagnosing steatosis, but it has limited utility.				
Steadman et al (2013)[59],	Multiple diseases	45/NR	0.88 (0.84 to 0.90) 80% (76% to 83%) 81% (77% to 85%)	49/NR	0.94 (0.91 to 0.96) 86% (82% to 89%) 89% (87% to 91%)
	HBV	5/710	0.81 (0.78 to 0.84) 77% (68% to 84%) 72% (55% to 85%)	8/1092	0.86 (0.82 to 0.89) 67% (57% to 75%) 87% (83% to 91%)
	HCV	13/2732	0.89 (0.86 to 0.91) 76% (61% to 86%) 86% (77% to 92%)	12/2887	0.94 (0.92 to 0.96) 85% (77% to 91%) 91% (87% to 93%)
	NAFLD	5/630	0.78 (0.74 to 0.82) 77% (70% to 83%) 75% (70% to 79%)	4/469	0.96 (0.94 to 0.97) 92% (77% to 98%) 95% (88% to 98%)
Stebbing et al (2010)[60],	Multiple diseases	17/3066	NR 72% (71% to 72%) 82% (82% to 83%)	17/4052	NR 84% (84% to 85%) 95% (94% to 95%)
Talwalkar et al (2007)[61],	Multiple diseases	7/>1100	0.87 (0.83 to 0.91) 70% (67% to 73%) 84% (80% to 88%)	9/2083	0.96 (0.94 to 0.98) 87% (84% to 90%) 91% (89% to 92%)
Tsochatzis et al (2011)[62],	Multiple diseases	31/5919	NR 79% (74% to 82%) 78% (72% to 83%)	30/6530	NR 83% (79% to 86%) 89% (87% to 91%)
	HCV	14/NR	NR 78% (71% to 84%) 80% (71% to 86%)	11/NR	NR 83% (77% to 88%) 90% (87% to 93%)

	HBV	4/NR	NR 84% (67% to 93%) 78% (68% to 85%)	6/NR	NR 80% (61% to 91%) 86% (82% to 94%)
Tsochatzis et al (2014)[63,]	HCV	37/NR	0.87 (0.83 to 0.90) 79% (74% to 84%) 83% (77% to 88%)	36/NR	0.96 (0.94 to 0.97) 89% (84% to 92%) 91% (89% to 93%)
	HBV	13/NR	0.83 (0.76 to 0.90) 71% (62% to 78%) 84% (74% to 91%)	13/NR	0.92 (0.89 to 0.96) 86% (79% to 91%) 85% (78% to 89%)
	NAFLD			4/NR	0.96 (0.94 to 0.99) 96% (83% to 99%) 89% (85% to 92%)
	ALD			6/NR	0.90 (0.87 to 0.94) 86% (76% to 92%) 83% (74% to 89%)
Xu et al (2015)[72,]	HBV	14/2318	0.82 (0.78 to 0.86) NR NR	18/2996	0.91 (0.89 to 0.93) NR NR
Xue-Ying (2020)[64,]	HBV	29/5035	0.83 (0.80 to 0.86) 72% (68% to 76%) 82% (77% to 86%)	NR/NR	NR NR NR

ALD: alcoholic liver disease; AUROC: area under the receiver operating characteristic curve; CI: confidence interval; HBV: hepatitis B virus; HCV: hepatitis C virus; HIV: human immunodeficiency virus; NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

There are currently no published studies that directly demonstrate the effect of transient elastography (eg, FibroScan) on patient outcomes.

FibroScan is used extensively in practice to make management decisions. In addition, FibroScan was used as an alternative to biopsy to diagnose fibrosis or cirrhosis to establish trial eligibility in several trials (ION-1,-3; VALENCE; ASTRAL-2, -3, -4) that confirmed efficacy of HCV treatments.(12-17) For example, in the VALENCE trial, cirrhosis could be defined by liver biopsy Or a confirmatory FibroTest or FibroScan result at 12.5 kPa or greater. In VALENCE, FibroScan was used to determine cirrhosis in 74% of the participants. In a retrospective, multicenter analysis of 7256 chronic HCV patients by Abdel Alem et al (2019), both transient elastography and FIB-4 were found to be predictors of treatment failure to sofosbuvir-based treatment regimens with an NPV of 95%.(73)

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Transient Elastography (FibroScan)
For individuals who have chronic liver disease who receive transient elastography (eg, FibroScan), the evidence includes many systematic reviews of more than 50 observational studies (>10,000 patients). Transient elastography has been studied in populations with viral hepatitis, NAFLD, and ALD. There are varying cutoffs for positivity. Failures of the test are not uncommon, particularly for those with high body mass index, but these failures often went undetected in analyses of the validation studies. Given these limitations and the imperfect reference standard, it can be difficult to interpret performance characteristics. However, for the purposes of deciding whether a patient has severe fibrosis or cirrhosis, the FibroScan results provide data sufficiently useful to determine therapy. In fact, FibroScan has been used as an alternative to biopsy to establish eligibility regarding the presence of fibrosis or cirrhosis in the participants of several RCTs. These trials showed the efficacy of HCV treatments, which in turn demonstrated that the test could identify patients who would benefit from therapy.

NONINVASIVE IMAGING: MULTIPARAMETRIC MAGNETIC RESONANCE IMAGING

Clinical Context and Test Purpose
The purpose of noninvasive testing in individuals with chronic liver disease is to detect liver fibrosis so that patients can avoid the potential adverse events of an invasive liver biopsy and receive appropriate treatment. The degree of liver fibrosis is an important factor in determining the appropriate approach for managing patients with liver disease (eg, hepatitis, ALD, NAFLD).

The question addressed in this portion of the evidence review is: Does the use of multiparametric MRI for detecting liver fibrosis improve the net health outcome in patients with chronic liver disease?

The following PICO was used to select literature to inform this review.

Populations
The relevant population of interest is individuals with chronic liver disease.

Interventions
The test being considered is multiparametric MRI (eg, LiverMultiScan).

Comparators
The following tests and practices are currently being used to diagnose chronic liver disease: liver biopsy, other noninvasive radiologic methods, and multianalyte serum assays.

Outcomes
The general outcomes of interest are test validity, morbid events, and treatment-related morbidity. Follow-up over months to years is of interest to the relevant outcomes.

Study Selection Criteria
For the evaluation of the clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores).
- Included a suitable reference standard (describe the reference standard).

- Patient/sample clinical characteristics were described.
- Patient/sample selection criteria were described.

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Tables 7 and 8 summarize studies that have evaluated the diagnostic accuracy of multiparametric MRI, which incorporates assessment of proton density fat-fraction, T2*, and T1 mapping to characterize liver fat, iron, fibrosis, and inflammation. Generally, technical failures were less common with MRI than transient elastography.(74,75,76)

Table 7. Characteristics of Studies Assessing the Diagnostic Accuracy of Multiparametric Magnetic Resonance Imaging

Study	Population	Design	Index Test(s)	Reference Standard	Timing of Reference and Index Tests
Beyer et al (2021)[74,]	N=580 patients with suspected NAFLD/NASH	Retrospective evaluation of patients from 2 clinical trials	MRI PDFF (LMS-IDEAL)* CAP (FibroScan)	Liver biopsy	Not reported
Imajo et al (2021)[75,]	N=145 patients with suspected NASH	Prospective, observational	MRI liver fat* MRI cT_1 measurements* MRI cT_1 + PDFF* MRE VCTE-LSM (FibroScan) CAP (FibroScan) 2D-SWE	Liver biopsy	All performed at first clinical visit
McDonald et al (2018)[76,]	N=149 patients with known or suspected liver disease	Prospective, validation cohort	MRI cT_1* ELF test TE (FibroScan)	Liver biopsy	Liver biopsy performed within 2 weeks of noninvasive assessments

*Measurements obtained with LiverMultiscan protocol.
2D-SWE: 2-dimensional shear-wave elastography; CAP: controlled attenuation parameter; ELF: Enhanced Liver Fibrosis; LMS-IDEAL: LiverMultiScan-Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation; MRE: magnetic resonance elastography; MRI: magnetic resonance imaging; NAFLD: non-alcoholic fatty liver disease; NASH: non-alcoholic steatohepatitis; PDFF: proton density fat-fraction; TE: transient elastography; VCTE-LSM: vibration-controlled transient elastography-liver stiffness measure.

Table 8. Results of Studies Assessing the Diagnostic Accuracy of Multiparametric Magnetic Resonance Imaging

		Significant Fibrosis				Steatosis			Advanced NASH (NAS ≥4 and ≥F2)	
Study	Population	Test	AUROC (95% CI) Sensitivity Specificity		Test	AUROC (95% CI) Sensitivity Specificity			Test	AUROC (95% CI) Sensitivity Specificity
						Grade ≥1	Grade ≥2	Grade ≥3		
Beyer et al (2021)[74,]	Suspected NAFLD/ NASH				MRI PDFF (LMS-IDEAL)*	1.0 (0.99 to 1.00)	0.77 (0.73 to 0.82)	0.81 (0.76 to 0.87)		

						99% 100%	72% 72%	68% 81%		
					CAP (FibroScan)	0.95 (0.91 to 0.99) 89% 100%	0.60 (0.55 to 0.65) 78% 41%	0.63 (0.57 to 0.70) 61% 59%		
			Stage ≥2							
Imajo et al (2021)[75,]	Suspected NASH	MRE	0.92 (0.87 to 0.97) NR NR		MRI liver fat*	0.92 (0.87 to 0.98) NR NR	0.86 (0.80 to 0.93) NR NR		MRI cT_1*	0.74 (0.66 to 0.82) NR NR
		VCTE-LSM	0.88 (0.81 to 0.95) NR NR		CAP (FibroScan)	0.75 (0.58 to 0.92) NR NR	0.68 (0.59 to 0.78) NR NR		MRI liver fat*	0.71 (0.63 to 0.80) NR NR
		2D-SWE	0.87 (0.76 to 0.99) NR NR						MRE	0.66 (0.57 to 0.75) NR NR
		MRI cT_1*	0.62 (0.49 to 0.74) NR NR						VCTE-LSM	0.64 (0.54 to 0.74) NR NR
			Stage ≥3	Stage ≥5						
McDonald et al (2018)[76,]	Known or suspected liver disease (unselected)	MRI cT_1*	0.72 (0.63 to 0.80) 88% 51%	0.72 (0.64 to 0.81) 71% 64%						
		ELF test	0.70 (0.61 to 0.78) 49% 77%	0.68 (0.57 to 0.79) 19% 91%						
		TE	0.84 (0.76 to 0.91) NR NR	0.86 (0.79 to 0.93) NR NR						

*Measurements obtained with LiverMultiscan protocol.
2D-SWE: 2-dimensional shear-wave elastography; AUROC: area under the receiver operating characteristic curve; CAP: controlled attenuation parameter; CI: confidence interval; ELF: Enhanced Liver Fibrosis; LMS-IDEAL: LiverMultiScan-Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation; MRE: magnetic resonance elastography; MRI: magnetic resonance imaging; NAFLD: non-alcoholic fatty liver disease; NASH: non-alcoholic steatohepatitis; NR: not reported; PDFF: proton density fat-fraction; TE: transient elastography; VCTE-LSM: vibration-controlled transient elastography-liver stiffness measure.

Jayaswal et al (2020) compared the prognostic value of MRI cT1 measurements, transient elastography, and multianalyte serum assays in a cohort of 197 patients with compensated chronic liver disease.(77) Patients who were referred for a clinically indicated liver biopsy, or with a known diagnosis of liver cirrhosis, were eligible. At baseline, patients underwent multiparametric MRI scans, transient elastography, and blood tests. Additionally, all patients received a liver biopsy and had their fibrosis rated on the Ishak scale; results of the biopsies informed clinical care. The most common underlying disease states were NAFLD (n=85, 43%),

viral hepatitis (n=50, 25%), and ALD (n=22, 11%). The primary endpoint was a composite of ascites, variceal bleeding, hepatic encephalopathy, hepatocellular carcinoma, liver transplantation and mortality. Binary cutoff values were predefined. Patients were followed for a median of 43 months. Over this period, 14 new clinical events were recorded, including 11 deaths. The prognostic value of the noninvasive testing is summarized in Table 8. Technical failures were also reported (eg, poor quality scan); reliable measurements were obtained in 182 of 197 (92%) patients for multiparametric MRI and in 121 of 160 (76%) patients for transient elastography (transient elastography was additionally not attempted in 37 patients). The study was limited by having variable follow-up periods and the effect of patients being censored at different time points was not taken into account, so sensitivities, specificities, PPVs and NPVs should be interpreted cautiously. The CI for the survival analysis were wide likely due to the relatively small number of new clinical events observed.

Table 9. Survival Analysis and Performance in Identifying Development of a New Clinical Event[a]

Test, Binary Cutoff	Cox Regression Analysis, HR (95% CI)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Liver cT_1 >825 ms	9.91 (1.287 to 76.24)	92.3	47.3	11.9	98.8
Transient elastography >8 kPa	7.79 (0.974 to 62.3)	88.9	51.8	12.9	98.3
FIB-4 >1.45	4.11 (0.91 to 18.56)	84.6	47.7	10.9	97.6
APRI >1	2.645 (0.886 to 7.9)	46.2	79.2	14.3	95.1
AST/ALT >1	6.093 (1.673 to 22.19)	76.9	65.6	14.3	97.4
Ishak >F4 (liver biopsy)	12.64 (2.8 to 57.08)	84.6	73.9	20.4	98.4

[a]Composite of ascites, variceal bleeding, hepatic encephalopathy, HCC, liver transplantation, and mortality
ALT: alanine aminotransferase; APRI: AST-to-platelet ratio; AST: aspartate aminotransferase; CI: confidence interval; FIB-4: fibrosis-4 index; HR: hazard ratio; kPa: kilopascal.

Pavlides et al (2016) evaluated whether data obtained from multiparametric MRI was predictive of all-cause mortality and liver-related clinical events.(78) Patients who were referred for a clinically indicated liver biopsy, or with a diagnosis of liver cirrhosis on MRI scan, were eligible. Liver-related clinical events were defined as liver-related death, hepatocellular carcinoma, and new hepatic decompensation (ie, clinically evident ascites, variceal bleeding, and hepatic encephalopathy). Patients received multiparametric MRI and liver cT1 values were mapped into a Liver Inflammation and Fibrosis (LIF) score. One hundred twenty three patients were recruited to the study; 6 were excluded due to claustrophobia or incomplete MRI data. Of the 117 patients who had complete MRI data, follow-up data were available for 112; the study reported outcomes on these 112 patients. The most common underlying disease states were NAFLD (35%), viral hepatitis (30%), and ALD (10%). Over a median follow-up time of 27

months, 10 patients had a liver-related clinical event and 6 patients died. No patients who had a LIF <2 (no or mild liver disease) developed a clinical event. Ten of 56 (18%) patients with a LIF ≥2 (moderate or severe liver disease) experienced a clinical event. A study limitation is the use of LIF scores, which are no longer used in clinical practice. The authors further described the study as a small proof of principle study.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs. The primary benefit of multiparametric MRI for chronic liver disease is the ability to avoid liver biopsy in patients without significant fibrosis.

LiverMultiScan (LMS) is a multiparametric MRI protocol consisting of proton density fat fraction (PDFF), T1, and T2 mapping sequences. A 2018 prospective validation study of 161 patients who had liver biopsies, transient elastography, Enhanced Liver Fibrosis (ELF) test, and contemporaneous LMS found sensitivity of 83% and negative predictive value of 96% for LMS, when evaluating for iron accumulation.(95)

Multiparametric MRI has been used as an alternative to biopsy for measuring fibrosis or cirrhosis in clinical trials. Phase 2 clinical trials have used multiparametric MRI to measure therapeutic efficacy of an investigational treatments for NASH (79) and NAFLD.(80) Guidelines outline the specific population in which LiverMultiScan is recommended for risk stratification (i.e. only in patients who have indeterminate/intermediate test results using current non-invasive testing methods).

The utility of multiparametric MRI to provide clinically useful information on the presence and extent of liver fibrosis and inflammation has been evaluated in smaller prospective studies. Specifically, it has been evaluated in the setting of biochemical remission in liver diseases where noninvasive testing for continued disease activity could further aid in direct management of patients as a prognostic marker of future liver-related complications. Quantitative multiparametric MRI has been used to measure disease burden after treatment in patients with chronic HCV (81) and autoimmune hepatitis. (82,83,84)

Section Summary: Multiparametric Magnetic Resonance Imaging
For individuals who have chronic liver disease who receive multiparametric MRI, the evidence includes several prospective and retrospective observational studies. Multiparametric MRI (eg, LiverMultiScan) has been studied in mixed populations, including NAFLD, viral hepatitis, and ALD. Quantitative MRI provides various measures assessing both liver fat content, fibrosis and inflammation. Various cutoffs have been utilized for positivity. Generally, multiparametric MRI performed similarly to transient elastography, and fewer technical failures of multiparametric MRI were reported. The prognostic ability of quantitative MRI to predict liver-related clinical events has been evaluated in 2 studies; both reported positive correlations, with wide confidence intervals. Additionally, multiparametric MRI has been used to measure the presence of fibrosis or cirrhosis in the patients who have achieved biochemical remission after treatment in small prospective studies.

OTHER NONINVASIVE IMAGING

Clinical Context and Test Purpose

The purpose of noninvasive testing in individuals with chronic liver disease is to detect liver fibrosis so that patients can avoid the potential adverse events of an invasive liver biopsy and receive appropriate treatment. The degree of liver fibrosis is an important factor in determining the appropriate approach for managing patients with liver disease (eg, hepatitis, ALD, NAFLD).

The question addressed in this portion of the evidence review is: Does the use of other noninvasive imaging for detecting liver fibrosis improve the net health outcome in patients with chronic liver disease?

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with chronic liver disease.

Interventions

The tests being considered are other noninvasive imaging, including magnetic resonance elastography (MRE), ARFI (eg, Acuson S2000), and real-time tissue elastography (RTE; eg, HI VISION Preirus).

Comparators

The following tests and practices are currently being used to diagnose chronic liver disease: liver biopsy, other noninvasive radiologic methods, and multianalyte serum assays.

Outcomes

The general outcomes of interest are test validity, morbid events, and treatment-related morbidity. Follow-up over months to years is of interest to the relevant outcomes.

Study Selection Criteria

For the evaluation of the clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores).
- Included a suitable reference standard (describe the reference standard).
- Patient/sample clinical characteristics were described.
- Patient/sample selection criteria were described.

Acoustic Radiation Force Impulse Imaging

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Tables 10 and 11 summarize the characteristics and results of systematic reviews that have assessed the diagnostic accuracy of ARFI imaging.

Table 10. Characteristics of Systematic Reviews Assessing Acoustic Radiation Force Impulse Imaging

Study	Dates	Studies	N	Population
Bota et al (2013)[51,]	To May 2012	6	518	Chronic hepatitis
Crossan et al (2015)[5,]	1998 to Apr 2012	4	NR	HCV
Guo et al (2015)[74,]	To Jun 2013	15	2128	Multiple diseases
Hu et al (2017)[75,]	To Jul 2014	7	723	NAFLD
Lin et al (2020)[76,]	To Apr 2019	29	NR	Non-viral liver disease
Jiang et al (2018)[68,]	To Dec 2017	9	982	NAFLD
Liu et al (2015)[77,]	To Apr 2016	23	2691	Chronic HBV or HCV
Nierhoff et al (2013)[78,]	2007 to Feb 2012	36	3951	Multiple diseases

HBV: hepatitis B virus; HCV: hepatitis C virus; NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Table 11. Results of Systematic Reviews Assessing the Diagnostic Accuracy of Acoustic Radiation Force Impulse Imaging

		Significant Fibrosis (ie, Metavir Stages F2 to F4)		Cirrhosis (ie, Metavir Stage F4)	
Study	**Population**	**Studies/ Sample Size**	**AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)**	**Studies/ Sample Size**	**AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)**
Bota et al (2013)[51,]	Chronic hepatitis	6/518	0.88 (0.83 to 0.93) NR NR		0.92 (0.87 to 0.98) NR NR
Crossan et al (2015)[5,]	HCV	4/NR	NR 85% (69% to 94%) 89% (72% to 97%)		
Guo et al (2015)[74,]	Multiple diseases	13/NR	NR 76% (73% to 78%) 80% (77% to 83%)	14/NR	NR 88% (84% to 91%) 80% (81% to 84%)
Hu et al (2017)[75,]	HBV, HCV	15/NR	88% (85% to 91%) 75% (69% to 78%) 85% (81% to 89%)		
Jiang et al (2018)[68,]	NAFLD	6/NR	0.86 (0.83 to 0.89) 70% (59% to 79%) 84% (79% to 88%)	7/NR	0.95 (0.93 to 0.97) 89% (60% to 98%) 91% (82% to 95%)
Liu et al (2015)[77,]	NAFLD	7/723	NR 80% (76% to 84%) 85% (81% to 89%)		
Lin et al (2020)[76,]	Non-viral liver disease	23/NR	0.87 (0.83 to 0.89) 79% (73% to 83%) 81% (75% to 86%)	14/NR	0.94 (0.92 to 0.96) 89% (79% to 95%) 89% (85% to 92%)
Nierhoff et al (2013)[78,]	Multiple diseases	26/NR	0.83 (0.80 to 0.86) NR NR	27/NR	0.91 (0.89 to 0.93) NR NR

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; HBV: hepatitis B virus; HCV: hepatitis C virus; NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

There are currently no published studies that directly demonstrate the effect of ARFI imaging on patient outcomes.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.
Because the clinical validity of ARFI has not been established, a chain of evidence supporting the clinical utility of this test for this population cannot be constructed.

Magnetic Resonance Elastography

Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Tables 12 and 13 summarize the characteristics and results of systematic reviews that have assessed the diagnostic accuracy of MRE. MRE has been studied primarily in hepatitis and NAFLD.

Table 12. Characteristics of Systematic Reviews Assessing Magnetic Resonance Elastography

Study	Dates	Studies	N	Population
Crossan et al (2015)[5,]	1998 to Apr 2012	3	NR	Chronic liver disease
Guo et al (2015)[74,]	To Jun 2013	11	982	Multiple diseases
Singh et al (2015)[79,]	2003 to Sep 2013	12	697	Chronic liver disease
Singh et al (2016)[80,]	To Oct 2014	9	232	NAFLD
Xiao et al (2017)[81,]	To 2016	5	628	NAFLD

NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Table 13. Results of Systematic Reviews Assessing the Diagnostic Accuracy of Magnetic Resonance Elastography

Study	Population	Studies/ Sample Size	AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)	Studies/ Sample Size	AUROC (95% CI) Sensitivity (95% CI) Specificity (95% CI)
Crossan et al (2015)[5,]	Chronic liver disease	3/NR	NR 94% (13% to 100%) 92% (72% to 98%)		

Guo et al (2015)[74,]	Multiple diseases	9/NR	NR 87% (84% to 90%) 94% (91% to 97%)		NR 93% (88% to 96%) 91% (88% to 93%)
Singh et al (2015)[79,]	Chronic hepatitis	12/697	0.84 (0.76 to 0.92) 73% (NR) 79% (NR)	12/697	0.92 (0.90 to 0.94) 91% (NR) 81% (NR)
Singh et al (2016)[80,]	NAFLD	9/232	0.87 (0.82 to 0.93) 79% (76% to 90%) 81% (72% to 91%)	9/232	0.91 (0.76 to 0.95) 88% (82% to 100%) 87% (77% to 97%)
Xiao et al (2017)[81,]	NAFLD	3/384	0.88 (0.83 to 0.92) 73.2% (65.7% to 87.3%) 90.7% (85.0% to 95.7%)	3/384	0.92 (0.80 to 1.00) 86.6% (80.0% to 90.9%) 93.4% (91.4% to 94.5%)

AUROC: area under the receiver operating characteristic curve; CI: confidence interval; NAFLD: nonalcoholic fatty liver disease; NR: not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

Chain of Evidence

For individuals who have chronic liver disease who receive magnetic resonance elastography, MRE has high diagnostic accuracy, particularly for the detection of fibrosis in NAFLD, independent of body mass index and degree of inflammation. MRE is also highly reproducible. The evidence is sufficient to determine that the technology results in an improvement in the net health outcome.

Real-Time Tissue Elastography (HI VISION 15 Preirus)

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Kobayashi et al (2015) published results of a meta-analysis of RTE for staging liver fibrosis.(82) The authors selected 15 studies (N=1626) published through December 2013, including patients with multiple liver diseases and healthy adults. A bivariate random-effects model was used to estimate summary sensitivity and specificity. The summary AUROC, sensitivity, and specificity were 0.69, 79% (95% CI, 75% to 83%) and 76% (95% CI, 68% to 82%), respectively, for detection of significant fibrosis (stage ≥F2) and 0.72, 74% (95% CI, 63% to 82%), and 84% (95% CI, 79% to 88%) for detection of cirrhosis. Reviewers found evidence of heterogeneity due to differences in study populations, scoring methods, and cutoffs for positivity. They also found evidence of publication bias based on funnel plot asymmetry.

Hong et al (2014) reported results of a meta-analysis RTE for staging fibrosis in multiple diseases.(83) Thirteen studies (N=1347) published between April 2000 and April 2014 that used a liver biopsy or transient elastography as the reference standard were included. Different quantitative methods were used to measure liver stiffness in the included studies: Liver Fibrosis Index (LFI), Elasticity Index (EI), elastic ratio 1 (ER1), and elastic ratio 2 (ER2). For predicting significant fibrosis (stage ≥F2), the pooled sensitivities for LFI and ER1 were 78% (95% CI, 70% to 84%) and 86% (95% CI, 80% to 90%), respectively. The specificities were 63% (95% CI, 46% to 78%) and 89% (95% CI, 83% to 94% and the AUROCs were 0.79 (95% CI, 0.75 to 0.82) and 0.94 (95% CI, 0.92 to 0.96), respectively. For predicting cirrhosis (stage F4), the pooled sensitivities of LFI, ER1, and ER2 were 79% (95% CI, 61% to 91%), 96% (95% CI, 87% to 99%), and 79% (95% CI, 61% to 91%), respectively. The specificities were 88% (95% CI, 81% to 93%) for LFI, 89% (95% CI, 83% to 93%) for ER1, and 88% (95% CI, 81% to 93%) for ER2, and the AUROCs were 0.85 (95% CI, 0.81 to 0.87), 0.93 (95% CI, 0.94 to 0.98), and 0.92 (95% CI, not reported), respectively. Pooled estimates for EI were not performed due to insufficient data.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy or avoid testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

There are currently no published studies that directly demonstrate the effect of RTE on patient outcomes.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Because the clinical validity of RTE has not been established, a chain of evidence supporting the clinical utility of this test for this population cannot be constructed.

Section Summary: Noninvasive Radiological Methods Other Than Transient Elastography

The use of ARFI imaging has been evaluated in viral hepatitis and NAFLD. Moreover, many have noted that ARFI imaging has potential advantages over FibroScan. ARFI can be implemented on a standard ultrasound machine, may be more applicable for assessing complications such as ascites, and may be more applicable in obese patients. ARFI imaging appears to have similar diagnostic accuracy to FibroScan, but there are fewer data available on performance characteristics. Validation studies have used varying cutoffs for positivity. MRE has a high success rate and is highly reproducible. The diagnostic accuracy also appears to be high. In particular, MRE has high diagnostic accuracy for the detection of fibrosis in NAFLD, independent of body mass index and degree of inflammation. However, further validation is needed to determine standard cutoffs and confirm performance characteristics because CI for estimates are wide. MRE is also not widely available. RTE has been evaluated in multiple diseases with varying scoring methods and cutoffs. Although data are limited, the accuracy of RTE appears to be similar to FibroScan for the evaluation of significant liver

fibrosis, but less accurate for the evaluation of cirrhosis. However, there was evidence of publication bias in the systematic review and the diagnostic accuracy may be overestimated.

For individuals who have chronic liver disease who receive noninvasive radiologic methods other than transient elastography for liver fibrosis measurement, the evidence includes systematic reviews of observational studies. Other radiologic methods (eg, MRE, RTE, ARFI, LMS) may have similar performance for detecting significant fibrosis or cirrhosis. Studies have frequently included varying cutoffs not prespecified or validated. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. There is no direct evidence that other noninvasive radiologic methods improve health outcomes; further, it is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence on clinical validity.

SUMMARY OF EVIDENCE

Multianalyte Serum Assays

For individuals who have chronic liver disease who receive FibroSURE serum panels, the evidence includes systematic reviews of more than 30 observational studies (>5000 patients). Relevant outcomes are test validity, morbid events, and treatment-related morbidity. FibroSURE has been studied in populations with viral hepatitis, NAFLD, and ALD. There are established cutoffs, although they were not consistently used in validation studies. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. However, for the purposes of deciding whether a patient has severe fibrosis or cirrhosis, FibroSURE results provide data sufficiently useful to determine therapy. Specifically, FibroSURE has been used as an alternative to biopsy to establish eligibility regarding the presence of fibrosis or cirrhosis in several RCTs that showed the efficacy of HCV treatments, which in turn demonstrated that the test can identify patients who would benefit from therapy. The evidence is sufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have chronic liver disease who receive multianalyte serum assays for liver function assessment other than FibroSURE, the evidence includes a number of observational studies and systematic reviews of those studies. Relevant outcomes are test validity, morbid events, and treatment-related morbidity. Studies have frequently included varying cutoffs, some of which were standardized and others not validated. Cutoff thresholds have often been modified over time, may be specific to certain patient populations, and in some cases, guideline recommendations differ from cutoffs designated by manufacturers and those utilized in studies. A comparison of transient elastography to various serum-based tests found that the former was superior in detecting fibrosis, and a meta-analysis of 4 studies found higher multianalyte scores associated with an increased risk of mortality relative to lower scores, but the evidence is limited by the small number of included studies and high heterogeneity and imprecision for some estimates. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. There is no direct evidence that other multianalyte serum assays improve health outcomes; further, it is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence on clinical validity. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Noninvasive Imaging

For individuals who have chronic liver disease who receive transient elastography, the evidence includes many systematic reviews of more than 50 observational studies (>10,000

patients). Relevant outcomes are test validity, morbid events, and treatment-related morbidity. Transient elastography (FibroScan) has been studied in populations with viral hepatitis, nonalcoholic fatty liver disease (NAFLD), and alcoholic liver disease (ALD). There are varying cutoffs for positivity. Failures of the test are not uncommon, particularly for those with high body mass index, but these failures often went undetected in analyses of the validation studies. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. However, for the purposes of deciding whether a patient has severe fibrosis or cirrhosis, the FibroScan results provide data sufficiently useful to determine therapy. In fact, FibroScan has been used as an alternative to biopsy to establish eligibility regarding the presence of fibrosis or cirrhosis in the participants of several RCTs. These trials showed the efficacy of hepatitis C virus (HCV) treatments, which in turn demonstrated that the test can identify patients who would benefit from therapy. The evidence is sufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have chronic liver disease who receive magnetic resonance elastography, MRE has high diagnostic accuracy, particularly for the detection of fibrosis in NAFLD, independent of body mass index and degree of inflammation. MRE is also highly reproducible. The evidence is sufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have chronic liver disease who receive multiparametric MRI, the evidence includes several prospective and retrospective observational studies. Multiparametric MRI (eg, LiverMultiScan) has been studied in mixed populations, including NAFLD, NASH, viral hepatitis, and alcoholic liver disease (ALD). Quantitative MRI provides various measures to assess liver fat content, fibrosis and inflammation. Various cutoffs have been utilized for positivity. Given these limitations and the imperfect reference standard, it can be difficult to interpret performance characteristics. Otherwise, multiparametric MRI performed similarly to transient elastography, and fewer technical failures of multiparametric MRI were reported. The prognostic ability of quantitative MRI to predict liver-related clinical events has been evaluated in 2 studies. Both studies reported positive correlations, but the confidence intervals were wide. Multiparametric MRI has been used to measure the presence of fibrosis, cirrhosis, or hematochromatosis in the patients who have achieved biochemical remission after treatment in small prospective studies. The evidence is sufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have chronic liver disease who received noninvasive radiologic methods other than transient elastography, magnetic resonance elastography, and multiparametric MRI (LiverMultiScan) for liver fibrosis measurement, the evidence includes systematic reviews of observational studies. The relevant outcomes are test validity, morbid events, and treatment-related morbidity. Other radiologic methods (eg, acoustic radiation force impulse imaging, real-time transient elastography) may have similar performance for detection of significant fibrosis or cirrhosis. Studies have frequently included varying cutoffs not prespecified or validated. Given these limitations and the imperfect reference standard, it is difficult to interpret performance characteristics. There is no direct evidence that other noninvasive radiologic methods improve health outcomes; further, it is not possible to construct a chain of evidence for clinical utility due to the lack of sufficient evidence on clinical validity. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Supplemental Information

The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

CLINICAL INPUT FROM PHYSICIAN SPECIALTY SOCIETIES AND ACADEMIC MEDICAL CENTERS

While the various physician specialty societies and academic medical centers may collaborate with and make recommendations during this process through the provision of appropriate reviewers, input received does not represent an endorsement or position statement by the physician specialty societies or academic medical centers, unless otherwise noted.

2015 Input

In response to requests, the Blue Cross Blue Shield Association received input from 3 physician specialty societies and three academic medical centers while their policy was under review in 2015. Most reviewers considered noninvasive techniques for the evaluation and monitoring of chronic liver disease to be investigational, both individually and in combination.

PRACTICE GUIDELINES AND POSITION STATEMENTS

Guidelines or position statements will be considered for inclusion in ‘Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

Nonalcoholic Fatty Liver Disease

American Gastroenterological Association et al

In 2018, the practice guidelines on the diagnosis and management of nonalcoholic fatty liver disease (NAFLD), developed by the American Gastroenterological Association (AGA), the American Association for the Study of Liver Diseases and the American College of Gastroenterology stated that “NFS [NAFLD fibrosis score] or FIB-4 [Fibrosis-4] index are clinically useful tools for identifying NAFLD patients with higher likelihood of having bridging fibrosis (stage 3) or cirrhosis (stage 4).”(84) This guideline also cited vibration-controlled transient elastography (VCTE) and magnetic resonance elastography (MRE) as “clinically useful tools for identifying advanced fibrosis in patients with NAFLD.”

A 2022 consensus-based clinical care pathway was published by the AGA on risk stratification and management of NAFLD, including some recommendations regarding the use of non-invasive testing for individuals with chronic liver disease. (93) Among individuals with increased risk of NAFLD or nonalcoholic steatohepatitis (NASH)-related fibrosis (i.e., individuals with type-2 diabetes, ≥2 metabolic risk factors, or an incidental finding of hepatic steatosis or elevated aminotransferases), assessment with a nonproprietary fibrosis scoring system such as FIB-4 is recommended, although aspartate transaminase to platelet ratio index can be used in lieu of FIB-4 scoring. Depending on the fibrosis score, imaging-based testing for liver stiffness may be warranted with transient elastography (FibroScan), although bidimensional shear wave elastography or point shear wave elastography are also imaging options included in the clinical care pathway.

American Association of Clinical Endocrinology and American Association for the Study of Liver Diseases

A 2022 joint clinical practice guideline issued by the American Association of Clinical Endocrinology and American Association for the Study of Liver Diseases included the following recommendations on the use of noninvasive techniques for diagnosis of NAFLD with clinically significant fibrosis (stage F2 to F4) (94)

- Clinicians should use liver fibrosis prediction calculations to assess the risk of NAFLD with liver fibrosis. The preferred noninvasive initial test is the FIB-4 (Grade B, Level 2 evidence)
- High-risk individuals with indeterminate or high FIB-4 score for further workup with an transient elastography or enhanced liver fibrosis test, as available (Grade B, Level 2 evidence)
- Clinicians should prefer the use of transient elastography as best validated to identify advanced disease and predict liver-related outcomes. Alternative imaging approaches may be considered, including shear wave elastography (less well validated) and/or magnetic resonance elastography (most accurate but with a high cost and limited availability; best if ordered by liver specialist for selected cases) (Grade B, Level 2 evidence).

National Institute for Health and Care Excellence

In 2016, the National Institute for Health and Care Excellence (NICE) published guidelines on the assessment and management of NAFLD.(49) The guidance did not reference elastography. The guidance recommended the enhanced liver fibrosis test to test for advanced liver fibrosis, utilizing a cutoff enhanced liver fibrosis score of 10.51.

American Gastroenterological Association Institute

In 2017, the American Gastroenterological Association Institute published guidelines on the role of elastography in chronic liver disease. The guidelines indicate that, in adults with NAFLD, VCTE has superior diagnostic sensitivity and specificity for diagnosing cirrhosis when compared to the aspartate aminotransferase-platelet ratio index (APRI) or FIB-4 tests (very low quality of evidence).(85) Moreover, the guidelines stated that, in adults with NAFLD, magnetic resonance-guided elastography has little or no increased diagnostic accuracy for identifying cirrhosis compared with VCTE in patients who have cirrhosis, and has higher diagnostic accuracy than VCTE in patients who do not have cirrhosis (very low quality of evidence).

Hepatitis B and C Viruses

National Institute for Health and Care Excellence

In 2017, the NICE published updated guidance on the management and treatment of patients with hepatitis B.(86) The guidance recommends offering transient elastography as the initial test in adults diagnosed with chronic hepatitis B, to inform the antiviral treatment decision (Table 14).

Table 14. Antiviral Treatment Recommendations by Transient Elasticity Score

Transient Elasticity Score	Antiviral Treatment
>11 kPa	Offer antiviral treatment
6 to 10 kPa	Offer liver biopsy to confirm fibrosis level prior to offering antiviral treatment
<6 kPa plus abnormal ALT	Offer liver biopsy to confirm fibrosis level prior to offering antiviral treatment
<6 kPa plus normal ALT	Do not offer antiviral treatment

ALT: alanine aminotransferase; kPa: kilopascal.

American Association for the Study of Liver Diseases and Infectious Diseases Society of America

In 2020, the American Association for the Study of Liver Disease and Infectious Diseases Society of America guidelines for testing, managing, and treating hepatitis C virus (HCV) recommended that for counseling and pretreatment assessment purposes, the following should be completed:

> "Evaluation for advanced fibrosis, using liver biopsy, imaging and/or noninvasive markers, is recommended in all persons with HCV infection to facilitate an appropriate decision regarding HCV treatment strategy and determine the need for initiating additional measures for the management of cirrhosis (eg, hepatocellular carcinoma screening). Rating: Class I, Level A [evidence and/or general agreement; data derived from multiple randomized trials, or meta-analyses]"(87)

The guidelines note that there are several noninvasive tests to stage the degree of fibrosis in patients with hepatitis C. Tests included indirect serum biomarkers, direct serum biomarkers, and vibration-controlled liver elastography. The guidelines assert that no single method is recognized to have high accuracy alone and careful interpretation of these tests is required.

American Gastroenterological Association Institute

In 2017, guidelines published by the American Gastroenterological Association Institute on the role of elastography in chronic liver disease indicated that, in adults with chronic hepatitis B virus and chronic HCV, VCTE has superior diagnostic performance for diagnosing cirrhosis when compared to the APRI and FIB-4 tests (moderate quality of evidence for HCV, low quality of evidence for hepatitis B virus).(85) In addition, the guidelines stated that, in adults with HCV, magnetic resonance-guided elastography has little or no increased diagnostic accuracy for identifying cirrhosis compared with VCTE in patients who have cirrhosis, and has lower diagnostic accuracy than VCTE in patients who do not have cirrhosis (very low quality of evidence).

Chronic Liver Disease

American College of Radiology

In 2020, the American College of Radiology appropriateness criteria rated ultrasound shear wave elastography as an 8 (usually appropriate) for the diagnosis of liver fibrosis in patients with chronic liver disease.(88) The criteria noted that high quality data can be difficult to obtain in obese patients, and assessments of liver stiffness can be confounded by parenchyma, edema, inflammation, and cholestasis.

U.S. PREVENTIVE SERVICES TASK FORCE RECOMMENDATIONS

A 2020 U.S. Preventive Services Task Force Recommendation Statement for HCV screening notes that a diagnostic evaluation for fibrosis stage or cirrhosis with a noninvasive test reduces the risk for harm compared to a liver biopsy.(89) This statement does not give preference to a specific noninvasive test.

ONGOING AND UNPUBLISHED CLINICAL TRIALS

Some currently unpublished trials that might influence this review are listed in Table 15.

Table 15. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT03789825	Screening for Liver Fibrosis. A Population-based Study in European Countries. The "LiverScreen" Project.	20000	Dec 2023 (enrolling by invitation)
NCT03308916[a]	Screening At-risk Populations for Hepatic Fibrosis With Non-invasive Markers (SIPHON)	6500	Oct 2035 (recruiting)
NCT02037867	The Stratification of Liver Disease in the Community Using Fibrosis Biomarkers	2000	May 2033 (recruiting)
NCT04435054	Screening for NAFLD-related Advanced Fibrosis in High Risk popuLation: Optimization of the Diabetology Pathway Referral Using Combinations of Non-invAsive Biological and elastogRaphy paramEters	1000	Oct 2023 (recruiting)
NCT04365855	The Olmsted NAFLD Epidemiology Study (TONES)	800	Jun 2028 (recruiting)
NCT04550481	Role of Lisinopril in Preventing the Progression of Non-Alcoholic Fatty Liver Disease, RELIEF-NAFLD Study	45	Mar 2023 (recruiting)

NCT: national clinical trial.

[a] Denotes industry-sponsored or cosponsored trial.

Government Regulations

National:

There is no Medicare national coverage determination.

Local:

There is no Medicare local coverage determination found on this topic.

The CMS 2022 Laboratory Fee Schedule has fees associated with codes 81596, 0002M, 0003M. An assigned fee is not a guarantee of coverage.

(The above Medicare information is current as of the review date for this policy. However, the coverage issues and policies maintained by the Centers for Medicare & Medicare Services [CMS, formerly HCFA] are updated and/or revised periodically. Therefore, the most current CMS information may not be contained in this document. For the most current information, the reader should contact an official Medicare source.)

Related Policies

Breast Elastography

References

1. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol. Oct 2002;97(10):2614-2618. PMID 12385448
2. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology. Mar 2009;49(3):1017-1044.
3. Mehta SH, Lau B, Afdhal NH, et al. Exceeding the limits of liver histology markers. J Hepatol. Jan 2009;50(1):36-41.
4. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard". J Gen Intern Med. Jun 2012;27 Suppl 1:S67-75.
5. Crossan C, Tsochatzis EA, Longworth L, et al. Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in patients with chronic liver disease: systematic review and economic evaluation. Health Technol Assess. Jan 2015;19(9):1-409, v-vi.
6. Houot M, Ngo Y, Munteanu M, et al. Systematic review with meta-analysis: direct comparisons of biomarkers for the diagnosis of fibrosis in chronic hepatitis C and B. Aliment Pharmacol Ther. Jan 2016;43(1):16-29.
7. Imbert-Bismut F, Ratziu V, Pieroni L, et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. Lancet. Apr 7 2001;357(9262):1069-1075. PMID 11297957
8. Poynard T, McHutchison J, Manns M, et al. Biochemical surrogate markers of liver fibrosis and activity in a randomized trial of peginterferon alfa-2b and ribavirin. Hepatology. Aug 2003;38(2):481-492. PMID 12883493
9. Poynard T, Munteanu M, Imbert-Bismut F, et al. Prospective analysis of discordant results between biochemical markers and biopsy in patients with chronic hepatitis C. Clin Chem. Aug 2004;50(8):1344-1355. PMID 15192028
10. Afdhal NH, Nunes D. Evaluation of liver fibrosis: a concise review. Am J Gastroenterol. Jun 2004;99(6):1160-1174. PMID 15180741
11. Lichtinghagen R, Bahr MJ. Noninvasive diagnosis of fibrosis in chronic liver disease. Expert Rev Mol Diagn. Sep 2004;4(5):715-726. PMID 15347264
12. Afdhal N, Reddy KR, Nelson DR, et al. Ledipasvir and sofosbuvir for previously treated HCV genotype 1 infection. N Engl J Med. Apr 17 2014;370(16):1483-1493.
13. Afdhal N, Zeuzem S, Kwo P, et al. Ledipasvir and sofosbuvir for untreated HCV genotype 1 infection. N Engl J Med. May 15 2014;370(20):1889-1898.
14. Curry MP, O'Leary JG, Bzowej N, et al. Sofosbuvir and velpatasvir for HCV in patients with decompensated cirrhosis. N Engl J Med. Dec 31 2015;373(27):2618-2628.

15. Foster GR, Afdhal N, Roberts SK, et al. Sofosbuvir and velpatasvir for HCV genotype 2 and 3 infection. N Engl J Med. Dec 31 2015;373(27):2608-2617.
16. Kowdley KV, Gordon SC, Reddy KR, et al. Ledipasvir and sofosbuvir for 8 or 12 weeks for chronic HCV without cirrhosis. N Engl J Med. May 15 2014;370(20):1879-1888.
17. Zeuzem S, Dusheiko GM, Salupere R, et al. Sofosbuvir and ribavirin in HCV genotypes 2 and 3. N Engl J Med. May 22 2014;370(21):1993-2001.
18. Naveau S, Raynard B, Ratziu V, et al. Biomarkers for the prediction of liver fibrosis in patients with chronic alcoholic liver disease. Clin Gastroenterol Hepatol. Feb 2005;3(2):167-174.
19. Ratziu V, Massard J, Charlotte F, et al. Diagnostic value of biochemical markers (FibroTest-FibroSURE) for the prediction of liver fibrosis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol. 2006;6:6.
20. Lassailly G, Caiazzo R, Hollebecque A, et al. Validation of noninvasive biomarkers (FibroTest, SteatoTest, and NashTest) for prediction of liver injury in patients with morbid obesity. Eur J Gastroenterol Hepatol. Jun 2011;23(6):499-506. PMID 21499110
21. Poynard T, Ratziu V, Charlotte F, et al. Diagnostic value of biochemical markers (NashTest) for the prediction of non alcoholo steato hepatitis in patients with non-alcoholic fatty liver disease. BMC Gastroenterol. Nov 10 2006; 6:34. PMID 17096854
22. Mohamadnejad M, Montazeri G, Fazlollahi A, et al. Noninvasive markers of liver fibrosis and inflammation in chronic hepatitis B-virus related liver disease. Am J Gastroenterol. Nov 2006;101(11):2537-2545.
23. Zeng MD, Lu LG, Mao YM, et al. Prediction of significant fibrosis in HBeAg-positive patients with chronic hepatitis B by a noninvasive model. Hepatology. Dec 2005;42(6):1437-1445.
24. Park MS, Kim BK, Cheong JY, et al. Discordance between liver biopsy and FibroTest in assessing liver fibrosis in chronic hepatitis B. PLoS One. 2013;8(2):e55759.
25. Salkic NN, Jovanovic P, Hauser G, et al. FibroTest/Fibrosure for significant liver fibrosis and cirrhosis in chronic hepatitis B: a meta-analysis. Am J Gastroenterol. Jun 2014;109(6):796-809
26. Xu XY, Kong H, Song RX, et al. The effectiveness of noninvasive biomarkers to predict hepatitis B-related significant fibrosis and cirrhosis: a systematic review and meta-analysis of diagnostic test accuracy. PLoS One. 2014;9(6):e100182.
27. Wai CT, Cheng CL, Wee A, et al. Non-invasive models for predicting histology in patients with chronic hepatitis B. Liver Int. Aug 2006;26(6):666-672.
28. Patel K, Gordon SC, Jacobson I, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. J Hepatol. Dec 2004;41(6):935-942. PMID 15582126
29. Mehta P, Ploutz-Snyder R, Nandi J, et al. Diagnostic accuracy of serum hyaluronic acid, FIBROSpect II, and YKL-40 for discriminating fibrosis stages in chronic hepatitis C. Am J Gastroenterol. Apr 2008;103(4):928-936. PMID 18371145
30. Patel K, Nelson DR, Rockey DC, et al. Correlation of FIBROSpect II with histologic and morphometric evaluation of liver fibrosis in chronic hepatitis C. Clin Gastroenterol Hepatol. Feb 2008;6(2):242-247. PMID 18187364
31. Snyder N, Nguyen A, Gajula L, et al. The APRI may be enhanced by the use of the FIBROSpect II in the estimation of fibrosis in chronic hepatitis C. Clin Chim Acta. Jun 2007;381(2):119-123. PMID 17442291
32. Castellana M, Donghia R, Guerra V, et al. Fibrosis-4 Index vs Nonalcoholic Fatty Liver Disease Fibrosis Score in Identifying Advanced Fibrosis in Subjects With Nonalcoholic Fatty Liver Disease: A Meta-Analysis. Am J Gastroenterol. Sep 01 2021; 116(9): 1833-1841. PMID 34160377

33. Mozes FE, Lee JA, Selvaraj EA, et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. Gut. May 17 2021. PMID 34001645
34. Sharma C, Cococcia S, Ellis N, et al. Systematic review: Accuracy of the enhanced liver fibrosis test for diagnosing advanced liver fibrosis and cirrhosis. J Gastroenterol Hepatol. Jul 2021; 36(7): 1788-1802. PMID 33668077
35. Wai CT, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. Aug 2003;38(2):518-526. PMID 12883497
36. Giannini EG, Zaman A, Ceppa P, et al. A simple approach to noninvasively identifying significant fibrosis in chronic hepatitis C patients in clinical practice. J Clin Gastroenterol. Jul 2006;40(6):521-527. PMID 16825935
37. Bourliere M, Penaranda G, Renou C, et al. Validation and comparison of indexes for fibrosis and cirrhosis prediction in chronic hepatitis C patients: proposal for a pragmatic approach classification without liver biopsies. J Viral Hepat. Oct 2006;13(10):659-670. PMID 16970597
38. Zarski JP, Sturm N, Guechot J, et al. Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: the ANRS HCEP-23 study. J Hepatol. Jan 2012;56(1):55-62. PMID 21781944
39. Sebastiani G, Halfon P, Castera L, et al. SAFE biopsy: a validated method for large-scale staging of liver fibrosis in chronic hepatitis C. Hepatology. Jun 2009;49(6):1821-1827. PMID 19291784.
40. Boursier J, de Ledinghen V, Zarski JP, et al. Comparison of eight diagnostic algorithms for liver fibrosis in hepatitis C: new algorithms are more precise and entirely noninvasive. Hepatology. Jan 2012;55(1):58-67. PMID 21898504
41. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. Gastroenterology. Dec 2004;127(6):1704-1713. PMID 15578508.
42. Siemens Healthineers. Liver Fibrosis Assays: Enhanced Liver Fibrosis (ELF) Test. 2019. https://www.siemenshealthineers.com/laboratory-diagnostics/assays-by-diseases-conditions/liver-disease/elf-test. Accessed 5/20/22.
43. Younossi ZM, Felix S, Jeffers T, et al. Performance of the Enhanced Liver Fibrosis Test to Estimate Advanced Fibrosis Among Patients With Nonalcoholic Fatty Liver Disease. JAMA Netw Open. Sep 01 2021; 4(9): e2123923. PMID 34529067
44. Sterling RK, Lissen E, Clumeck N et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology, 2006 May 27;43(6). PMID 16729309.
45. Vallet-Pichard A, Mallet V, Nalpas B et al. “FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. Comparison with liver biopsy and fibrotest.” Hepatology, 2007 Jun 15;46(1). PMID 17567829.
46. Yan LT, Wang LL, Yao J, et al. Total bile acid-to-cholesterol ratio as a novel noninvasive marker for significant liver fibrosis and cirrhosis in patients with non-cholestatic chronic hepatitis B virus infection. Medicine (Baltimore). Feb 2020; 99(8): e19248. PMID 32080129
47. Cianci N, Subhani M, Hill T, et al. Prognostic non-invasive biomarkers for all-cause mortality in non- alcoholic fatty liver disease: A systematic review and meta-analysis. World J Hepatol. May 27 2022; 14(5): 1025-1037. PMID 35721296
48. Sanyal AJ, Harrison SA, Ratziu V et al. The Natural History of Advanced Fibrosis Due to Nonalcoholic Steatohepatitis: Data From the Simtuzumab Trials. Hepatology, 2019 Apr 18. PMID 30993748.

49. National Institute for Health and Care Excellence (NICE). Non-alcoholic fatty liver disease (NAFLD): assessment and management [NG49]. 2016; https://www.nice.org.uk/guidance/ng49.
50. Brener S. Transient elastography for assessment of liver fibrosis and steatosis: an evidence-based analysis. Ont Health Technol Assess Ser. 2015;15(18):1-45.
51. Bota S, Herkner H, Sporea I, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int. Sep 2013;33(8):1138-1147. PMID 23859217
52. Chon YE, Choi EH, Song KJ, et al. Performance of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B: a meta-analysis. PLoS One. 2012;7(9):e44930. PMID 23049764
53. Friedrich-Rust M, Ong MF, Martens S, et al. Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. Gastroenterology. Apr 2008;134(4):960-974.
54. Kwok R, Tse YK, Wong GL, et al. Systematic review with meta-analysis: non-invasive assessment of non-alcoholic fatty liver disease--the role of transient elastography and plasma cytokeratin-18 fragments. Aliment Pharmacol Ther. Feb 2014;39(3):254-269.
55. Poynard T, Morra R, Ingiliz P, et al. Assessment of liver fibrosis: noninvasive means. Saudi J Gastroenterol. Oct 2008;14(4):163-173.
56. Poynard T, Ngo Y, Munteanu M, et al. Noninvasive markers of hepatic fibrosis in chronic hepatitis B. Curr Hepat Rep. Jun 2011;10(2):87-97.
57. Shaheen AA, Wan AF, Myers RP. FibroTest and FibroScan for the prediction of hepatitis C-related fibrosis: a systematic review of diagnostic test accuracy. Am J Gastroenterol. Nov 2007;102(11):2589-2600.
58. Shi KQ, Tang JZ, Zhu XL, et al. Controlled attenuation parameter for the detection of steatosis severity in chronic liver disease: a meta-analysis of diagnostic accuracy. J Gastroenterol Hepatol. Jun 2014;29(6):1149-1158.
59. Steadman R, Myers RP, Leggett L, et al. A health technology assessment of transient elastography in adult liver disease. Can J Gastroenterol. Mar 2013;27(3):149-158.
60. Stebbing J, Farouk L, Panos G, et al. A meta-analysis of transient elastography for the detection of hepatic fibrosis. J Clin Gastroenterol. Mar 2010;44(3):214-219
61. Talwalkar JA, Kurtz DM, Schoenleber SJ, et al. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. Clin Gastroenterol Hepatol. Oct 2007;5(10):1214-1220.
62. Tsochatzis EA, Gurusamy KS, Ntaoula S, et al. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. J Hepatol. Apr 2011;54(4):650-659.
63. Tsochatzis EA, Crossan C, Longworth L, et al. Cost-effectiveness of noninvasive liver fibrosis tests for treatment decisions in patients with chronic hepatitis C. Hepatology. Sep 2014;60(3):832-843.
64. Xu XY, Wang WS, Zhang QM, et al. Performance of common imaging techniques vs serum biomarkers in assessing fibrosis in patients with chronic hepatitis B: A systematic review and meta-analysis. World J Clin Cases. Aug 06 2019; 7(15): 2022-2037. PMID 31423434
65. Cai C, Song X, Chen X, et al. Transient Elastography in Alcoholic Liver Disease and Nonalcoholic Fatty Liver Disease: A Systemic Review and Meta-Analysis. Can J Gastroenterol Hepatol. 2021; 2021: 8859338. PMID 33542909
66. Friedrich-Rust M, Nierhoff J, Lupsor M, et al. Performance of Acoustic Radiation Force Impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. J Viral Hepat. Feb 2012;19(2):e212-219. PMID 22239521.

67. Geng XX, Huang RG, Lin JM, et al. Transient elastography in clinical detection of liver cirrhosis: A systematic review and meta-analysis. Saudi J Gastroenterol. Jul-Aug 2016;22(4):294-303.
68. Jiang W, Huang S, Teng H, et al. Diagnostic accuracy of point shear wave elastography and transient elastography for staging hepatic fibrosis in patients with non-alcoholic fatty liver disease: a meta-analysis. BMJ Open. Aug 23, 2018;8(8):e021787. PMID 30139901
69. Li Y, Huang YS, Wang ZZ, et al. Systematic review with meta-analysis: the diagnostic accuracy of transient elastography for the staging of liver fibrosis in patients with chronic hepatitis B. Aliment Pharmacol Ther. Feb 2016;43(4):458-469
70. Njei B, McCarty TR, Luk J, et al. Use of transient elastography in patients with HIV-HCV coinfection: A systematic review and meta-analysis. J Gastroenterol Hepatol. Oct 2016;31(10):1684-1693. PMID 26952020
71. Pavlov CS, Casazza G, Nikolova D, et al. Transient elastography for diagnosis of stages of hepatic fibrosis and cirrhosis in people with alcoholic liver disease. Cochrane Database Syst Rev. 2015;1:CD010542.
72. Xu X, Su Y, Song R, et al. Performance of transient elastography assessing fibrosis of single hepatitis B virus infection: a systematic review and meta-analysis of a diagnostic test. Hepatol Int. Oct 2015;9(4):558-566.
73. Abdel Alem S, Elsharkawy A, El Akel W et al. Liver stiffness measurements and FIB-4 are predictors of response to sofosbuvir-based treatment regimens in 7256 chronic HCV patients. Expert Rev Gastroenterol Hepatol, 2019 Aug 17;1-8:1-8. PMID 31418303.
74. Guo Y, Parthasarathy S, Goyal P, et al. Magnetic resonance elastography and acoustic radiation force impulse for staging hepatic fibrosis: a meta-analysis. Abdom Imaging. Apr 2015;40(4):818-834.
75. Hu X, Qiu L, Liu D, et al. Acoustic Radiation Force Impulse (ARFI) Elastography for noninvasive evaluation of hepatic fibrosis in chronic hepatitis B and C patients: a systematic review and meta-analysis. Med Ultrason. Jan 31, 2017;19(1):23-31. PMID 28180193
76. Lin Y, Li H, Jin C, et al. The diagnostic accuracy of liver fibrosis in non-viral liver diseases using acoustic radiation force impulse elastography: A systematic review and meta-analysis. PLoS One. 2020; 15(1): e0227358. PMID 31940395
77. Liu H, Fu J, Hong R, et al. Acoustic radiation force impulse elastography for the non-invasive evaluation of hepatic fibrosis in non-alcoholic fatty liver disease patients: a systematic review & meta-analysis. PLoS One. Jul 2015;10(7):e0127782. PMID 26131717
78. Nierhoff J, Chavez Ortiz AA, Herrmann E, et al. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. Eur Radiol. Nov 2013;23(11):3040-3053.
79. Singh S, Venkatesh SK, Wang Z, et al. Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. Clin Gastroenterol Hepatol. Mar 2015;13(3):440-451 e446.
80. Singh S, Venkatesh SK, Loomba R, et al. Magnetic resonance elastography for staging liver fibrosis in non-alcoholic fatty liver disease: a diagnostic accuracy systematic review and individual participant data pooled analysis. Eur Radiol. May 2016;26(5):1431-1440.
81. Xiao G, Zhu S, Xiao X et al. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. Hepatology, 2017 Jun 7;66(5): 1486-501. PMID 28586172.
82. Kobayashi K, Nakao H, Nishiyama T, et al. Diagnostic accuracy of real-time tissue elastography for the staging of liver fibrosis: a meta-analysis. Eur Radiol. Jan 2015;25(1):230-238. PMID 25149296.

83. Hong H, Li J, Jin Y, et al. Performance of real-time elastography for the staging of hepatic fibrosis: a meta-analysis. PLoS One. 2014;9(12):e115702.
84. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology. Jan 2018; 67(1): 328-357. PMID 28714183
85. Singh S, Muir AJ, Dieterich DT, et al. American Gastroenterological Association Institute Technical Review on the Role of Elastography in Chronic Liver Diseases. *Gastroenterology.* May 2017;152(6):1544-1577. PMID 28442120
86. National Institute for Health and Care Excellence (NICE). Hepatitis B (chronic): diagnosis and management [CG165]. 2017; https://www.nice.org.uk/guidance/CG165.
87. American Association for the Study of Liver Diseases and Infectious Diseases Society of America. Recommendations for Testing, Managing, and Treating Hepatitis C Last updated September 29, 2021; https://www.hcvguidelines.org/evaluate/testing-and-linkage.
88. Bashir MR, Horowitz JM, Kamel IR, et al. ACR Appropriateness Criteria(R) Chronic Liver Disease. J Am Coll Radiol. May 2020; 17(5S): S70-S80. PMID 32370979
89. Owens DK, Davidson KW, Krist AH, et al. Screening for Hepatitis C Virus Infection in Adolescents and Adults: US Preventive Services Task Force Recommendation Statement. JAMA. Mar 02 2020. PMID 32119076
90. Lim JK, Flamm SL, Singh S, et al. American Gastroenterological Association Institute Guideline on the Role of Elastography in the Evaluation of Liver Fibrosis. Gastroenterology. 2017;152(6):1536-43.
91. Horowitz JM, Kamel IR, Arif-Tiwari H, et al. ACR Appropriateness Criteria((R)) Chronic Liver Disease. Journal of the American College of Radiology : JACR. 2017;14(11s):S391-s405.
92. Su LN, Guo SL, Li BX, et al. Diagnostic value of magnetic resonance elastography for detecting and staging of hepatic fibrosis: a meta-analysis. Clin Radiol. 2014;69(12):e545-52.
93. Kanwal F, Shubrook JH, Adams LA, et al. Clinical Care Pathway for the Risk Stratification and Management of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. Nov 2021; 161(5): 1657-1669. PMID 34602251
94. Cusi K, Isaacs S, Barb D, et al. American Association of Clinical Endocrinology Clinical Practice Guideline for the Diagnosis and Management of Nonalcoholic Fatty Liver Disease in Primary Care and Endocrinology Clinical Settings: Co-Sponsored by the American Association for the Study of Liver Diseases (AASLD). Endocr Pract. May 2022; 28(5): 528-562. PMID 35569886
95. McDonald, Natasha et al. “Multiparametric magnetic resonance imaging for quantitation of liver disease: a two-centre cross-sectional observational study.” Scientific reports vol. 8,1 9189. 15 Jun. 2018, doi:10.1038/s41598-018-27560-5

The articles reviewed in this research include those obtained in an Internet based literature search for relevant medical references through 5/12/23, the date the research was completed.

Joint BCBSM/BCN Medical Policy History

Policy Effective Date	BCBSM Signature Date	BCN Signature Date	Comments
1/1/15	6/16/15	7/16/15	Joint policy established
9/1/16	6/21/16	7/25/16	• Routine maintenance • MPS and Inclusion criteria changed to include XL probe
9/1/17	6/20/17	6/20/17	• Routine maintenance • References and rationale updated • Added Medicare noncoverage article regarding procedure codes 0001M-0003M
9/1/18	6/19/18	6/19/18	• Routine maintenance • New LCD added
9/1/19	6/18/19		• Routine maintenance • Code update - 0001M replaced with 81596 and 0346T replaced with 76981, 76982 and 76983 per AMA
9/1/20	8/18/20		• Routine maintenance • Updated policy stance to cover FibroSure testing for HCV • 81596 moved to EST based on above • Code update – 0014M EI
9/1/21	6/15/21		• Routine maintenance • MRE and US elastography changed to EST
9/1/22	6/21/22		Routine maintenance FibroSure tests: ASH and NASH added as covered. MPS, inclusions, exclusions edited. Ref added: 32,33,34,43,46,47,65
9/1/23	6/26/23		(BCBSA policy last updated Dec 2022; Updated JUMP policy based on BCBSA; JUMP DRAFT policy has multiparametric MRI as E/I. Last year's JUMP policy remained silent on multiparametric MRI.

			Minor editorial refinements to policy statements – patients changed to individuals intent unchanged Per code update recommendation added new code 87467 EFD 1/1/23 as payable. 5/4/23: Received request from Perspectum for JUMP policy to cover multiparametric MRI. Carelon allows multiparametric MRI as an alternative to MRE. Our JUMP policy covers MRE. Vendor: Carelon manages code 76391 (MRE). Codes 0648T and 0649T multiparametric MRI (LiverMultiScan): are not on the list of Carelon procedures for BCBSM. Aligned with Carelon: Added to MPS: Multiparametric MRI (LiverMultiScan) is considered a useful option for nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH) as an alternative to MR elastography (MRE) for diagnosis and management of advanced hepatic fibrosis/cirrhosis. Added to Inclusions: Multiparametric Liver may be considered established in the following scenario: As an alternative to MR elastography for diagnosis and management of advanced hepatic fibrosis/cirrhosis PostJUMP: • Add codes 0648T and 0649T under established codes. • Updated MPS to read: Multiparametric MRI (LiverMultiScan) is considered a useful option for diagnosis and management of advanced hepatic fibrosis/cirrhosis. • Updated Inclusion under statement: Multiparametric MRI (LiverMultiScan) is considered a useful option for the diagnosis and management of advanced hepatic fibrosis/cirrhosis when diagnostic

			testing such as an ultrasound is inconclusive or non-diagnostic (ky)
5/1/24	2/20/24		This policy is coming early as code update – informational to add code 81517 effective 1/1/24 per code update as E/I. Code 0014M is removed as this code is deleted eff 1/1/24. This policy will go back to its original date of June 2024 JUMP. Vendor: Carelon (ky)

Next Review Date: 2nd Qtr, 2024

BLUE CARE NETWORK BENEFIT COVERAGE
POLICY: NONINVASIVE TECHNIQUES FOR THE EVALUATION AND MONITORING OF PATIENTS WITH CHRONIC LIVER DISEASE

I. Coverage Determination:

Commercial HMO (includes Self-Funded groups unless otherwise specified)	Covered; criteria apply
BCNA (Medicare Advantage)	Refer to the Medicare information under the Government Regulations section of this policy.
BCN65 (Medicare Complementary)	Coinsurance covered if primary Medicare covers the service.

II. Administrative Guidelines:

- The member's contract must be active at the time the service is rendered.
- Coverage is based on each member's certificate and is not guaranteed. Please consult the individual member's certificate for details. Additional information regarding coverage or benefits may also be obtained through customer or provider inquiry services at BCN.
- The service must be authorized by the member's PCP except for Self-Referral Option (SRO) members seeking Tier 2 coverage.
- Services must be performed by a BCN-contracted provider, if available, except for Self-Referral Option (SRO) members seeking Tier 2 coverage.
- Payment is based on BCN payment rules, individual certificate and certificate riders.
- Appropriate copayments will apply. Refer to certificate and applicable riders for detailed information.
- CPT - HCPCS codes are used for descriptive purposes only and are not a guarantee of coverage.