Multianalyte Assays with Algorithmic Analysis for the Evaluation and Monitoring of Patients with Chronic Liver Disease

Effective: July 1, 2018

Next Review: May 2019
Last Review: May 2018

IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

Multianalyte serum assays with algorithmic analysis are being evaluated as a substitute for biopsy in the screening, evaluation, and monitoring of patients with chronic liver disease.

MEDICAL POLICY CRITERIA

Multianalyte assays with algorithmic analyses, including but not limited to the following tests are considered investigative for the evaluation and monitoring of patients with chronic liver disease:

A. HCV FibroSure™ (FibroTest™)
B. Elasto-FibroTest®
C. FibroSpect II
D. ASH FibroSURE™ (ASH Test)
E. NASH FibroSURE™ (NASH Test)

NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.
CHRONIC LIVER DISEASES

Hepatitis C

Infection with the hepatitis C virus can lead to permanent liver damage. Liver biopsy is typically recommended prior to 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-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

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 0-4, 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 spectrum of the disease, there is usually no appreciable inflammation, hepatocyte death, or fibrosis. In contrast, non-alcoholic 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 histological 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 0-4, with 4 being cirrhosis.

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 can 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 0-4 (with 0 being no or minimal inflammation and 4 being severe) and fibrosis from 0-4 (with 0 being no fibrosis and 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.

MULTIANALYTE ASSAYS

A variety of noninvasive laboratory tests are being evaluated as alternatives 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 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 and FibroTest**

**HCV FibroSURE**

HCV FibroSURE (FibroTest) uses a combination of six 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 (i.e., fibrosis) and grade (i.e., necroinflammatory activity). The measures are combined using a linear regression equation to produce a score between 0 and 1, with higher values corresponding to more severe disease. The biochemical markers include the readily available measurements of α2-macroglobulin, haptoglobin, bilirubin, γ-glutamyl transpeptidase (GGT), ALT, and apolipoprotein AI. 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 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 three 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.

EVIDENCE SUMMARY

Validation of the clinical use of any diagnostic test focuses on three main principles:

1. Analytic validity of the test;
2. Clinical validity of the test (i.e., sensitivity, specificity, and positive and negative predictive values in relevant populations of patients and compared to the gold standard); and
3. Clinical utility of the test (i.e., how the results of the diagnostic test will be used to improve the management of the patient).

LIVER BIOPSY IS AN IMPERFECT REFERENCE STANDARD
As mentioned in the Background, 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] This will bias estimates of performance characteristics of the noninvasive tests to which it is compared and must be considered in appraising the body of evidence. Mehta estimated that, under the best scenario where sensitivity and specificity of liver biopsy are 90% and the prevalence of significant disease (Metavir ≥ F2) is 40%, even a perfect alternative marker would have calculated area under the receiver operating characteristic (AUROC) curve of 0.90.[3] Therefore, 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.

**FIBROSURE AND FIBROTEST**

**Hepatitis C Virus**

**Analytic Validity**

Measurement of the serum levels of liver function tests (ie, α₂-macroglobulin, haptoglobin, γ-glutamyl transpeptidase [GGT], total bilirubin, apolipoprotein AI) are readily available biochemical tests. However, measurement of serum factors that directly measure fibrogenesis are relatively novel, and not readily available. Studies to formally validate the parameters used to calculate the HCV FibroSURE scores reported acceptable levels of intralaboratory and intrapatient variability.[4,5]

**Clinical Validity**

Initial research into the HCV FibroSURE algorithm involved testing an initial panel of 11 serum markers in 339 patients with liver fibrosis who had undergone liver biopsy. From the original group of 11 markers, five were selected as the most informative, based on logistic regression, neural connection, and receiver operating characteristic (ROC) curves. Markers included α₂-macroglobulin, haptoglobin, γ-globulin, apolipoprotein AI, GGT, and total bilirubin.[6] Using an algorithm-derived scoring system ranging from 0 to 1.0, authors reported that a score of less than 0.10 was associated with a NPV of 100% (ie, absence of fibrosis, as judged by liver biopsy scores of Metavir F2-F4). A score greater than 0.60 was associated with a 90% positive predictive value (PPV) of fibrosis (ie, Metavir F2-F4). Authors concluded that liver biopsy might be deferred in patients with a score less than 0.10.

The next step in the development of this test was further evaluation of the algorithm in a cross-section of patients, including patients with HCV participating in large clinical trials before and after the initiation of antiviral therapy. One study focused on patients with HCV who were participating in a randomized study of pegylated interferon and ribavirin.[7] 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% PPV for the diagnosis of Metavir F2-F4. The specificity was 36%, and the NPV was 40%.

Poynard 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.[8] In this study, cutoff values were used for
individual Metavir scores (ie, F0-F4) and for combinations of Metavir scores (ie, F0-F1, F1-F2). The definition of a significant discordance between FibroTest and ActiTest and biopsy scores was at least two stages or grades in the Metavir system. Discordance was observed in 29% of patients. Risk factors for failure of HCV FibroSURE scoring system were 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 and to the biopsy in 18% and nonattributed 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.[9,10]

One Australian study attempted to independently replicate the results of FibroSURE in 125 patients with hepatitis C.[11] Using the cutoff of less than 0.1 to identify lack of bridging fibrosis (ie, Metavir stages F0-F1) and greater than 0.6 to identify fibrosis (ie, Metavir stages F2-F4), the NPV for a score of less than 0.1 was 89%, and the PPV of a score greater than 0.6 was 78%.

In 2012, Poynard assessed the relative accuracy of FibroTest and FibroScan using a method to estimate performance characteristics when no perfect reference standard exists.[12] The study included 1893 subjects retrospectively extracted from four prospective cohorts: three cohorts with HCV (n=1289) and one cohort of healthy volunteers (n=604). Four different tests (FibroTest, FibroScan, alanine aminotransferase [ALT], liver biopsy) were performed on all patients with HCV. Latent class models with random effects were used to combine the test results to construct a reference standard. When compared to biopsy as the reference standard, the sensitivity and specificity for the diagnosis of advanced fibrosis were 85% and 66% for FibroTest and 93% and 48% for FibroScan. However, when compared to the latent class reference standard, the specificity and sensitivity for the diagnosis of advanced fibrosis were 93% and 70% for FibroTest and 96% and 45% for FibroScan.

In the Crossan (2015) systematic review, FibroTest was the most widely validated commercial serum test.[13] 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.

Clinical Utility

The effect on patient outcomes of a test depends on a demonstration that the test can be used to improve patient management. The primary benefit of the FibroSURE (FibroTest) for HCV is the ability to avoid liver biopsy in patients without significant fibrosis. Thus, empiric data are needed that demonstrate that the FibroSURE test impacts clinician decision making on whether a biopsy should be performed and that the net effect is to reduce the overall number of biopsies while achieving similar clinical outcomes. There are currently no such published studies to demonstrate effect on patient outcomes. However, FibroTest has been used as an alternative to biopsy to establish trial eligibility in terms of fibrosis or cirrhosis in several trials (ION-1,-3; VALENCE; ASTRAL-2, -3, -4) that established efficacy of HCV treatments.[14-19] For
example, in the ASTRAL-2 and -3 trials, cirrhosis could be defined by liver biopsy, FibroScan, or FibroTest score of more than 0.75 and 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 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 the HCV FibroSURE could be used as an interval test in patients receiving therapy to determine whether an additional liver biopsy was necessary.

Alcoholic Liver Disease and Alcoholic Steatohepatitis

Analytic Validity

As above (see the Technical Performance: Hepatitis C Virus section).

Clinical Validity

The diagnostic value of FibroSURE (FibroTest) has also been evaluated for the prediction of liver fibrosis in patients with ALD and NAFLD.[20,21] In 2006, Thabut reported the development of a panel of biomarkers (ASH FibroSURE [ASH Test]) for the diagnosis of alcoholic steatohepatitis (ASH) in patients with chronic ALD.[22] Biomarkers were initially assessed with a training group consisting of 70 patients, and a panel was constructed using a combination of the six biochemical components of the FibroTest-ActiTest plus aspartate aminotransferase (AST). The algorithm was subsequently studied in two validation groups (one prospective study for severe ALD, one retrospective study for nonsevere 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 to be 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, one study was identified that described diagnostic accuracy of FibroTest for significant fibrosis (stage ≥ F2) or cirrhosis in ALD.[13] 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%) and for cirrhosis were 100% (95% CI, 95% to 100%) and 50% (95% CI, 42% to 58%), respectively.

Clinical Utility

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

Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis

Analytic Validity

As above (see the Technical Performance: Hepatitis C Virus section).

Clinical Validity

In 2006, Poynard reported the development of a panel of biomarkers (NASH FibroSURE [NASH Test]) for the prediction of nonalcoholic steatohepatitis (NASH) in patients with NAFLD.[23] Biomarkers were initially assessed with a training group consisting 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 three classes for NASH (NASH, borderline NASH, no NASH). The main end point was steatohepatitis, defined as a histologic NASH score (NAS) 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 PPVs and NPVs of 66% and 81%, respectively. For borderline NASH or NASH, there was a sensitivity of 88%, specificity of 50%, PPV of 74%, and NPV of 72%. Clinically significant discordance (two class difference) was observed in eight (8%) patients. None of the 383 controls was considered to have NASH by NASH FibroSURE (NASH Test). Authors proposed that this test would be suitable for mass screening for NAFLD in patients with obesity and diabetes.

An independent study from France attempted to prospectively validate the NASH Test (along with the FibroTest, SteatoTest, and ActiTest) in a cohort of 288 patients treated with bariatric surgery.[24] Included were patients with severe or morbid obesity (body mass index [BMI], >35 kg/m²), at least one comorbidity for at least five 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 κ reliability test (0.14). In 183 patients 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, seven (47%) were no NASH and four (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, four studies were included in the pooled estimate of the diagnostic accuracy of FibroTest for advanced fibrosis (stage ≥ 3) in NAFLD. The summary sensitivities and specificities were 40% (95% CI, 24% to 58%) and 96% (95% CI, 91% to 98%). Only one 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.

Clinical Utility

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

Hepatitis B Virus

Analytic Validity

As above (see the Technical Performance: Hepatitis C Virus section).

Clinical Validity

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. In a 2013 study, Park compared liver biopsy and the FibroTest results obtained on the same day from 330 patients with chronic HBV. Discordance was found in 30 (9.1%) patients for whom the FibroTest underestimated fibrosis in 25 patients and overestimated it in five 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<0.001). The only independent factor for discordance on multivariate analysis was a Metavir stages F3 or F4 on liver biopsy (p<0.001).

In 2014 Salkic conducted a meta-analysis of studies on the diagnostic accuracy of FibroTest in chronic HBV. Included in the meta-analysis were 16 studies (2494 patients) on liver fibrosis diagnosis and 13 studies (1754 patients) 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-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 odds ratio 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.

In 2014 Xu reported on a systematic review and meta-analysis of studies on biomarkers to detect fibrosis in HBV. Included in the analysis on FibroTest were 11 studies (total 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, six studies were included in the pooled estimate of the diagnostic accuracy of FibroTest for significant fibrosis (stage ≥ F2) in HBV. 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 cirrhosis in HBV was based on four 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%).

Clinical Utility

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.[30]

Section Summary: FibroSURE and FibroTest

FibroSURE (FibroTest) is the most widely validated of the noninvasive commercial serum tests. It has been studied in populations with viral hepatitis, NAFLD, and ALD. Although there are established cutoffs for positivity for FibroTest, they were not consistently used in validation studies. The methodologic quality of the validation studies was generally poor. There is no direct evidence that FibroSURE (FibroTest) improves health outcomes. However, FibroTest has been allowed as an alternative to biopsy to establish trial eligibility in terms of fibrosis or cirrhosis in several randomized controlled trials (RCTs) that established the efficacy of HCV treatments.

FIBROSPECT II

Analytic Validity

As previously noted, the FIBROSpect test consists of measurements of hyaluronic acid, tissue inhibitors of metalloproteinase–1 (TIMP-1), and α2-macroglobulin. In a 2004 review, Lichtinghagen and Bahr noted that the lack of standardization of assays of matrix metalloproteinases and TIMP limited the interpretation of studies.[10]

Clinical Validity

Patel 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.[31] The algorithm was designed to distinguish between no or mild fibrosis (F0-F1) and moderate-to-severe fibrosis (F2-F4). With the prevalence of F2-F4 disease of 52% and a cutoff value of 0.36, the PPVs and NPVs were 74.3% and 75.8%, respectively. Using a FIBROSpect II cutoff score of 0.42, Christensen reported a sensitivity of 93%, specificity of 66%, overall accuracy of 76%, and a NPV of 94% for advanced fibrosis in 136 patients with HCV.[32]

The published studies for this combination of markers continue to focus on test characteristics such as sensitivity, specificity, and accuracy.[33-35] In Crossan (2015), the summary diagnostic accuracy for detecting significant fibrosis (stage ≥ F2) in five 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%).[13]

Clinical Utility

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

Section Summary: FIBROSpect II
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.

OTHER MULTIANALYTE SCORING SYSTEMS

Other scoring systems have been developed. For example, the APRI requires only the serum level of AST and the number of platelets, and uses a simple nonproprietary formula that can be calculated at the bedside to produce a score for the prediction of fibrosis.[36] 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 (2015), APRI was frequently evaluated and has been tested in HCV, HBV, NAFLD, and ALD.[13] The summary diagnostic accuracies are in Table 1.

Table 1. Diagnostic Accuracy for APRI from Crossan (2015)

<table>
<thead>
<tr>
<th>Disease</th>
<th>Metavir Fibrosis Stage</th>
<th>Cutoff</th>
<th>Studies</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCV</td>
<td>≥ F2 (significant)</td>
<td>Low: 0.4 to 0.7</td>
<td>47</td>
<td>82% (77% to 86%)</td>
<td>57% (49% to 65%)</td>
</tr>
<tr>
<td>HCV</td>
<td>≥ F2 (significant)</td>
<td>High: 1.5</td>
<td>36</td>
<td>39% (32% to 47%)</td>
<td>92% (89% to 95%)</td>
</tr>
<tr>
<td>HCV</td>
<td>F4 (cirrhosis)</td>
<td>Low: 0.75 to 1</td>
<td>24</td>
<td>77% (73% to 81%)</td>
<td>78% (74% to 81%)</td>
</tr>
<tr>
<td>HCV</td>
<td>F4 (cirrhosis)</td>
<td>High: 2</td>
<td>19</td>
<td>48% (41% to 56%)</td>
<td>94% (91% to 95%)</td>
</tr>
<tr>
<td>HBV</td>
<td>≥ F2 (significant)</td>
<td>Low: 0.4 to 0.6</td>
<td>8</td>
<td>80% (68% to 88%)</td>
<td>65% (52% to 77%)</td>
</tr>
<tr>
<td>HBV</td>
<td>≥ F2 (significant)</td>
<td>High: 1.5</td>
<td>6</td>
<td>37% (22% to 55%)</td>
<td>93% (85% to 97%)</td>
</tr>
<tr>
<td>HBV</td>
<td>F4 (cirrhosis)</td>
<td>Low: 1</td>
<td>4</td>
<td>58% (49% to 66%)</td>
<td>76% (70% to 81%)</td>
</tr>
<tr>
<td>HBV</td>
<td>F4 (cirrhosis)</td>
<td>High: 2</td>
<td>3</td>
<td>24% (8% to 52%)</td>
<td>91% (83% to 96%)</td>
</tr>
<tr>
<td>NAFLD</td>
<td>≥ F3 (significant)</td>
<td>0.5 to 1.0</td>
<td>4</td>
<td>40% (7% to 86%)</td>
<td>82% (78% to 60%)</td>
</tr>
<tr>
<td>NAFLD</td>
<td>F4 (cirrhosis)</td>
<td>0.54 and NA</td>
<td>2</td>
<td>78% (71% to 99%)</td>
<td>71% (30% to 93%)</td>
</tr>
<tr>
<td>ALD</td>
<td>≥ F2 (significant)</td>
<td>Low: 0.5</td>
<td>2</td>
<td>72% (60% to 82%)</td>
<td>46% (33% to 60%)</td>
</tr>
<tr>
<td>ALD</td>
<td>≥ F2 (significant)</td>
<td>High: 1.5</td>
<td>2</td>
<td>54% (42% to 66%)</td>
<td>78% (64% to 88%)</td>
</tr>
<tr>
<td>ALD</td>
<td>F4 (cirrhosis)</td>
<td>High: 2.0</td>
<td>1</td>
<td>40% (22% to 61%)</td>
<td>62% (41% to 79%)</td>
</tr>
</tbody>
</table>

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.

Rosenberg developed a scoring system based on an algorithm combining hyaluronic acid, amino terminal propeptide of type III collagen, and TIMP-1.[37] 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%.
Giannini reported that use of the AST/ALT ratio and platelet counts in a diagnostic algorithm would have avoided liver biopsy in 69% of their patients and would have correctly identified the absence/presence of significant fibrosis in 80.5% of these cases. In Crossan (2015), the cutoffs for positivity of AST/ALT ratio for diagnosis of significant fibrosis (stage ≥ F2) varied from 0.6 to 1 in seven studies. 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 ROC analysis. For example, Bourliere reported 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. A 2012 prospective multicenter study from France compared nine of the best-evaluated blood tests in 436 patients with hepatitis C and found similar performance for HCV FibroSURE (FibroTest), FibroMeter, and HepaScore (ROC curve, 0.84, 0.86, 0.84, respectively). These three tests were significantly superior to the six 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. One of these, the sequential algorithm for fibrosis evaluation (SAFE), combines the APRI and FibroTest. Crossan (2015) reported that the algorithm has been assessed in four studies of HCV for diagnosing both significant fibrosis (stage ≥ F2) and cirrhosis. 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.

PRACTICE GUIDELINE SUMMARY

AMERICAN ASSOCIATION FOR THE STUDY OF LIVER DISEASES (AASLD), THE AMERICAN COLLEGE OF GASTROENTEROLOGY (ACG), AND THE AMERICAN GASTROENTEROLOGICAL ASSOCIATION (AGA)

The 2012 AASLD/ACG/AGA practice guideline on the Diagnosis and Management of Non-Alcoholic Fatty Liver Disease (NAFLD) delineates when subsequent biopsy is recommended following unsuspected hepatic steatosis detected on imaging (strong and high to moderate recommendations). Regarding non-invasive assessment of steatohepatitis and advanced fibrosis in NAFLD, discussion mentions that prediction models and biomarkers are limited by the cross-sectional study designs utilized to assess their clinical utility.

AMERICAN ASSOCIATION FOR THE STUDY OF LIVER DISEASES (AASLD) AND THE INFECTIOUS DISEASES SOCIETY OF AMERICA (IDSA)

The AASLD/IDSA Guidance on Hepatitis (updated in 2017) recommends evaluation for advanced fibrosis in those with current (active) HCV infection. Evaluation includes using liver biopsy, imaging, and/or noninvasive markers to facilitate an appropriate decision regarding HCV treatment strategy and to determine the need for initiating additional measures for the management of cirrhosis (e.g., hepatocellular carcinoma screening). This recommendation has a rating of Class I: Conditions for which there is evidence and/or general
agreement that a given diagnostic evaluation, procedure, or treatment is beneficial, useful, and effective; and Level A: Data derived from multiple randomized clinical trials, meta-analyses, or equivalent.

DEPARTMENT OF VETERANS AFFAIRS (VA) AND THE NATIONAL VIRAL HEPATITIS PROGRAM IN THE OFFICE OF PATIENT CARE SERVICES

The VA and National Viral Hepatitis Program in the Office of Patient Care Services treatment considerations (updated in 2017) for chronic hepatitis C virus infection state that cirrhosis can be diagnosed by a variety of non-invasive means; liver biopsy should be reserved for situations in which the risks and limitations of the procedure are outweighed by the benefits of obtaining information via this technique.[44]

UNITED STATES PREVENTIVE SERVICES TASK FORCE (USPSTF)

The USPSTF June 2013 Hepatitis C: Screening Final Recommendation Statement specifies there is adequate evidence that various noninvasive tests have good to very good diagnostic accuracy in diagnosing fibrosis or cirrhosis.[45] However, a research gap exists with regard to the clinical utility of “noninvasive” assessment of cirrhosis and fibrosis; patient populations may vary from those who were enrolled in trial cohorts. Specific multianalyte assays with algorithmic analysis are not mentioned in the recommendation.

SUMMARY

There is not enough research to know if multianalyte assays with algorithmic analyses improve health outcomes for people with chronic liver disease. Therefore, the use of multianalyte assays to evaluate or monitor people with chronic liver disease is considered investigational.

REFERENCES


**CODES**

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0001M</td>
<td>HCV FibroSURE™, LabCorp. Infectious disease, HCV, 6 biochemical assays (ALT, A2-macroglobulin, apolipoprotein A-1, total bilirubin, GGT, and haptoglobin) utilizing serum, prognostic algorithm reported as scores for fibrosis and necroinflammatory activity in liver</td>
</tr>
<tr>
<td></td>
<td>0002M</td>
<td>ASH FibroSURE™, LabCorp. Liver disease, 10 biochemical assays (ALT, A2-macroglobulin, apolipoprotein A-1, total bilirubin, GGT, haptoglobin, AST, glucose, total cholesterol and triglycerides) utilizing serum, prognostic algorithm reported as quantitative scores for fibrosis, steatosis, and alcoholic steatohepatitis (ASH)</td>
</tr>
<tr>
<td></td>
<td>0003M</td>
<td>NASH FibroSURE™, LabCorp. Liver disease, 10 biochemical assays (ALT, A2-macroglobulin, apolipoprotein A-1, total bilirubin, GGT, haptoglobin, AST, glucose, total cholesterol and triglycerides) utilizing serum, prognostic algorithm</td>
</tr>
<tr>
<td>Codes</td>
<td>Number</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>reported as quantitative scores for fibrosis, steatosis, and non-alcoholic steatohepatitis (NASH)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>81599</td>
<td>Multianalyte assay with algorithmic analysis</td>
</tr>
<tr>
<td></td>
<td>83520</td>
<td>Immunoassay, analyte, quantitative; not otherwise specified] tissue inhibitor of metalloproteinase (TIMP-1)</td>
</tr>
<tr>
<td></td>
<td>83883</td>
<td>Alpha-2 macroglobulin– Nephelometry, each analyte not elsewhere specified</td>
</tr>
<tr>
<td></td>
<td>84999</td>
<td>Unlisted chemistry procedure</td>
</tr>
</tbody>
</table>

**Date of Origin:** June 2013