HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: NONINVASIVE TESTING FOR LIVER FIBROSIS IN PATIENTS WITH CHRONIC HEPATITIS C

Approved 10/6/2016

HERC Coverage Guidance

If a fibrosis score of ≥F2 is the threshold for antiviral treatment of hepatitis C, the following are recommended for coverage *(weak recommendation)*:

Imaging tests:

- Transient elastography (FibroScan®)
- Acoustic radiation force impulse imaging (ARFI) (Virtual Touch™ tissue quantification, ElastPQ)
- Shear wave elastography (SWE) (Aixplorer®)

Blood tests (only if imaging tests are unavailable):

- Enhanced Liver Fibrosis (ELF™)
- Fibrometer™
- FIBROSpect® II

If a fibrosis score of ≥F3 is the threshold for antiviral treatment of hepatitis C, one or more of the following are recommended for coverage *(strong recommendation*):f

Imaging tests:

- Transient elastography (FibroScan®)
- Acoustic radiation force impulse imaging (ARFI)
- Shear wave elastography (SWE)

Magnetic resonance elastography is recommended for coverage for ≥F2 or ≥F3 only when at least one imaging test (FibroScan, ARFI, and SWE) has resulted in indeterminant results, a second one is similarly indeterminant, contraindicated or unavailable, and MRE is readily available *(weak recommendation).*

Noninvasive tests should be performed no more often than once per year *(weak recommendation).*

The following tests are not recommended for coverage for the detection of liver fibrosis to guide treatment decisions with antivirals in chronic hepatitis C *(strong recommendation):*

Imaging tests

- Real time tissue elastography

Blood tests (proprietary):

- Hepascore® (FibroScore®)
- FibroSure® (FibroTest®)

Blood tests (non-proprietary):

- Age-platelet index
- AST-platelet ratio index (APRI)
- AST-ALT ratio
- Cirrhosis discriminant score (Bonacini index)
- FIB-4
- Fibro-α score
- FibroIndex
- Fibronectin discriminant score
- FibroQ
- Fibrosis–cirrhosis index
- Fibrosis index
- Fibrosis probability index (Sud index)
- Fibrosis–protein index
- Fibrosis Routine Test
- Forns index
- Globulin–albumin ratio
- Göteborg University Cirrhosis Index (GUCI)
- HALT-C model (Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis)
- King's score
- Lok index
- MP3 score
- Pohl index
- Sabadell NIHCED index (Non-Invasive Hepatitis-C–Related Cirrhosis Early Detection)
- Significant fibrosis index
- Zeng index

Note: Definitions for strength of recommendation are provided in Appendix A *GRADE Informed Framework Element Description*.

RATIONALE FOR DEVELOPMENT OF COVERAGE GUIDANCES AND MULTISECTOR INTERVENTION REPORTS

Coverage guidances are developed to inform coverage recommendations for public and private health plans in Oregon as they seek to improve patient experience of care, population health and the cost-effectiveness of health care. In the era of the Affordable Care Act and health system transformation, reaching these goals may require a focus on population-based health interventions from a variety of sectors as well as individually-focused clinical care. Multisector intervention reports will be developed to address these population-based health interventions or other types of interventions that happen outside of the typical clinical setting.

HERC selects topics for its reports to guide public and private payers based on the following principles:

- Represents a significant burden of disease or health problem
- Represents important uncertainty with regard to effectiveness or harms
- Represents important variation or controversy in implementation or practice
- Represents high costs or significant economic impact
- Topic is of high public interest

Our reports are based on a review of the relevant research applicable to the intervention(s) in question. For coverage guidances, which focus on clinical interventions and modes of care, evidence is evaluated using an adaptation of the GRADE methodology. For more information on coverage guidance methodology, see Appendix A.

Multisector interventions can be effective ways to prevent, treat or manage disease at a population level. For some conditions, the HERC has reviewed evidence and identified effective interventions, but has not made coverage recommendations, as many of these policies are implemented in settings beyond traditional healthcare delivery systems.

GRADE-INFORMED FRAMEWORK

The HERC develops recommendations by using the concepts of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system. GRADE is a transparent and structured process for developing and presenting evidence and for carrying out the steps involved in developing recommendations. There are several elements that determine the strength of a recommendation, as listed in the table below. The HERC reviews the evidence and makes an assessment of each element, which in turn is used to develop the recommendations presented in the coverage guidance box. Estimates of effect are derived from the evidence presented in this document. The level of confidence in the estimate is determined by the Commission based on assessment of two independent reviewers from the Center for Evidence-based Policy. Unless otherwise noted, estimated resource allocation, values and preferences, and other considerations are assessments of the Commission.

Coverage question: Should noninvasive testing for liver fibrosis for chronic hepatitis C be recommended for coverage?				
Outcomes	**Estimate of Effect for Outcome/ *Confidence in Estimate***	**Resource allocation**	**Values and Preferences**	**Other considerations**
Hepatitis-related morbidity/ progression *(Critical outcome)*	Diagnostic strategies have not been directly compared to assess the effect on hepatitis-related morbidity or progression.	Non-invasive imaging tests are generally less costly than liver biopsy, but more costly than serum tests. Given that both serum and noninvasive tests are less invasive that biopsy, it is likely that more patients will be referred for, and receive treatment with noninvasive testing. Some	Most patients would strongly prefer to have a noninvasive test over a liver biopsy in order to avoid the procedural risks associated with the biopsy. Policy makers will need to balance the value of this greater access to less	Guidelines are mixed in their recommendations about the use of serum biomarker testing as an adjunct or alternative to imaging. Many of the serum biomarkers are commonly obtained and inexpensive.
Need for liver biopsy *(Critical outcome)*	No studies directly addressed whether the use of noninvasive tests reduce the need for liver biopsy. However, in clinical practice, these tests are used to replace liver biopsy. Therefore, their diagnostic operating characteristics, in comparison to liver biopsy, are reported here as AUROC for ≥F2, and tests with adequate diagnostic performance may be indirectly assumed to reduce the use of liver biopsy: Magnetic Resonance Elastography AUROC 0.88 (95%CI 0.84 to 0.91)			

Coverage question: Should noninvasive testing for liver fibrosis for chronic hepatitis C be recommended for coverage?				
Outcomes	**Estimate of Effect for Outcome/ *Confidence in Estimate***	**Resource allocation**	**Values and Preferences**	**Other considerations**
	●●●○ *(Moderate confidence)* Transient Elastography AUROC 0.89 (95% CI 0.86 to 0.91) ●●●○ *(Moderate confidence)* Acoustic Radiation Force Impulse Imaging AUROC 0.88 (95% CI 0.81 to 0.96) ●●○○ *(Low confidence)* Shear Wave Elastography AUROC 0.88 (95% CI 0.85 to 0.91) ●●○○ *(Low confidence)* Real-time Tissue Elastography AUROC 0.69 (95% CI NR) ●○○○ *(Very low confidence)* Platelet count Median AUROC 0.71 (range 0.38 to 0.94) ●○○○ *(Very low confidence)* Platelet count Median AUROC 0.71 (range 0.38 to 0.94) ●○○○ *(Very low confidence)*	patients who have noninvasive tests may also still require additional testing if findings are inconclusive. In cases where treatment decisions are based on the results of these tests, false positives may lead to high treatment costs; false negatives may lead to undertreatment or delayed treatment. MRE is much more expensive than the other imaging tests.	sensitive/specific tests with the potential undertreatment or overtreatment that could occur as a result of the inferior accuracy of these tests compared to liver biopsy.	Many institutions may only have one type of imaging modality available. It could be equally appropriate to do a second imaging test versus going straight to liver biopsy depending on the institution and availability of nearby alternatives.

Coverage question: Should noninvasive testing for liver fibrosis for chronic hepatitis C be recommended for coverage?				
Outcomes	**Estimate of Effect for Outcome/ *Confidence in Estimate***	**Resource allocation**	**Values and Preferences**	**Other considerations**
	Hyaluronic acid Median AUROC 0.75 (range 0.65 to 0.88) ●○○○ *(Very low confidence)* Age-platelet index Median AUROC 0.74 (range 0.64 to 0.79) ●●○○ *(Low confidence)* APRI Median AUROC 0.77 (range 0.58 to 0.95) ●○○○ *(Very low confidence)* AST-ALT ratio Median AUROC 0.59 (range 0.50 to 0.82) ●○○○ *(Very low confidence)* Bonacini index Median AUROC 0.66 (range 0.58 to 0.71) ●●○○ *(Low confidence)* ELF™ Median AUROC 0.81 (range 0.72 to 0.87) ●○○○ *(Very low confidence)*			

Coverage question: Should noninvasive testing for liver fibrosis for chronic hepatitis C be recommended for coverage?				
Outcomes	**Estimate of Effect for Outcome/ *Confidence in Estimate***	**Resource allocation**	**Values and Preferences**	**Other considerations**
	FIB-4 Median AUROC 0.74 (range 0.61 to 0.81) ●○○○ *(Very low confidence)* FibroIndex Median AUROC 0.76 (0.58 to 0.86) ●○○○ *(Very low confidence)* FibroMeter™ Median AUROC 0.82 (range 0.78 to 0.85) ●○○○ *(Very low confidence)* FIBROSpect® II Median AUROC 0.86 (range 0.77 to 0.95) ●○○○ *(Very low confidence)* FibroTest® Median AUROC 0.79 (range 0.70 to 0.89) ●○○○ *(Very low confidence)* Forns index Median AUROC 0.76 (0.60 to 0.86) ●○○○ *(Very low confidence)*			

Coverage question: Should noninvasive testing for liver fibrosis for chronic hepatitis C be recommended for coverage?				
Outcomes	**Estimate of Effect for Outcome/ *Confidence in Estimate***	**Resource allocation**	**Values and Preferences**	**Other considerations**
	Hepascore® Median AUROC 0.79 (range 0.69 to 0.82) ●○○○ *(Very low confidence)* Pohl index Median AUROC 0.52 (range 0.52 to 0.53) ●●○○ *(Low confidence)*			
Quality of life *(Critical outcome)*	No data identified			
Testing-related adverse events *(Important outcome)*	No data identified			
Change in treatment plan *(Important outcome)*	No data identified			
Balance of benefits and harms: Given the good (F2) and excellent (F3) performance of the recommended imaging tests and the potential harms of liver biopsy, the balance is strongly in favor of offering these tests as an option for patients for whom hepatitis C direct-acting antiviral therapy is being considered. Because these tests sometimes return inconclusive results, additional testing including liver biopsy may still be required for some patients. Though they are inferior to the recommended imaging tests, blood tests also have a good performance at the F2 threshold and have a favorable balance when imaging tests are unavailable and biopsy is not required.				

Rationale: The diagnostic operating characteristic of the recommended imaging tests are good to excellent (defined as an AUROC ≥0.8). Patient-oriented health outcomes are not available. However, given the characteristics of the tests, the strong values and preferences for noninvasive tests when results are comparable, and the improved individual-level resource allocation, these tests are recommended for coverage. The strong recommendation for imaging tests when the cutoff is F3 is due to the excellent performance at this level of cutoff (defined as an AUROC ≥0.9) and the other factors in favor of their use. The weak recommendation at the F2 cutoff is based on "good" but not "excellent" performance, and the high societal cost of treating patients at levels of fibrosis who are not at short-term risk.

The diagnostic operating characteristics of the blood tests are variable. Though tests recommended at the F2 threshold can accurately assess the fibrosis stage F2 or higher, they are inferior to the imaging tests at this level, and expert input suggests less clinically reliable, and so are recommended only when imaging tests are unavailable. No existing blood test can accurately distinguish between F2 and F3. Therefore, blood tests cannot be recommended (alone or in combination with noninvasive imaging tests) when the treatment planning revolves around an accurate diagnosis of F3. Many of the non-recommended blood tests have fair to poor operating characteristics regardless of the treatment threshold.

MRE is much more expensive than the other imaging tests and thus is only recommended when available after two other imaging tests fail to return useful results.

Recommendation:

If a fibrosis score of ≥F2 is the threshold for antiviral treatment of hepatitis C, the following are recommended for coverage *(weak recommendation)*:

Imaging tests:

- Transient elastography (FibroScan®)
- Acoustic radiation force impulse imaging (ARFI) (Virtual Touch™ tissue quantification, ElastPQ)
- Shear wave elastography (SWE) (Aixplorer®)

Blood tests (only if imaging tests are unavailable):

- Enhanced Liver Fibrosis (ELF™)
- Fibrometer™

- FIBROSpect® II

If a fibrosis score of ≥F3 is the threshold for antiviral treatment of hepatitis C, one or more of the following are recommended for coverage *(strong recommendation*):

Imaging tests:

- Transient elastography (FibroScan®)
- Acoustic radiation force impulse imaging (ARFI)
- Shear wave elastography (SWE)

Magnetic resonance elastography is recommended for coverage for ≥F2 or ≥F3 only when at least one imaging test (FibroScan, ARFI, and SWE) has resulted in indeterminant results, a second one is similarly indeterminant, contraindicated or unavailable, and MRE is readily available *(weak recommendation).*

Noninvasive tests should be performed no more often than once per year *(weak recommendation).*

Other imaging and blood tests are not recommended for coverage *(strong recommendation).*

*The Quality of Evidence rating was assigned using information from the editing sources and judgments made by CEbP staff based on direction from the subcommittee.

Note: GRADE framework elements are described in Appendix A. A GRADE Evidence Profile is provided in Appendix B.

EVIDENCE OVERVIEW

Clinical background

Hepatitis C virus (HCV) is a major cause of liver disease in the United States, and chronic hepatitis C infection is the leading indication for liver transplantation (Centers for Disease Control and Prevention [CDC], 2016). The CDC estimates that 3.5 million people in the United States are currently infected with HCV, though the precise number is not known. One study cited by the CDC estimated that around 15,000 deaths were attributable to HCV in 2007. Well established modes of transmission for HCV infection include injection drug use and receipt of blood products prior to 1992. According to the CDC, the prevalence of HCV infection among injection drug users ranges from about 30% for younger users (aged 18 to 30) to 70-90% for older injection drug users.

The natural history of HCV infection is variable, and 15-25% of people will clear the infection and not develop chronic hepatitis C. Between 5% and 20% of those with HCV infection will develop cirrhosis, generally over the course of 20 to 30 years, and between 1% and 5% will die from HCV-related liver disease (CDC, 2016). There are no highly accurate tools to predict which individuals with chronic hepatitis C will go on to develop cirrhosis.

The United States Preventive Services Task Force recommends birth-cohort screening for hepatitis C for anyone born between 1945 and 1965. HCV testing is also recommended for those in high risk groups included people with a history of injection drug use, those who received blood products before 1992, those with HIV infection, and those born to HCV-positive mothers (CDC, 2016).

Before 2013, treatment for chronic hepatitis C relied on interferon and ribavirin, sometimes with the addition of a protease inhibitor in the case of genotype 1 infections. These treatments were long (24 to 48 weeks), entailed a high burden of adverse effects, and response rates were highly variable. The advent of direct-acting antiviral treatments (i.e. sofosbuvir, simeprevir, and others) appears to have improved the success rates (as measured by the surrogate marker of sustained virologic response at 12 weeks) and acceptability of treatment, though at considerable cost.

Traditionally, staging of chronic hepatitis C infection was done by examining histologic specimens from liver biopsies of the liver for evidence of fibrosis. The METAVIR fibrosis stage is the most commonly used measure for assessing the histologic degree of hepatic fibrosis:

- F0 = No fibrosis
- F1 = Portal fibrosis without septa
- F2 = Portal fibrosis with few septa
- F3 = Portal fibrosis with numerous septa without cirrhosis
- F4 = Cirrhosis

Progression from fibrosis to cirrhosis is associated with complications of end-stage liver disease including portal hypertension, portosystemic encephalopathy, and hepatocellular carcinoma.

Noninvasive tests of liver fibrosis and cirrhosis have developed as an alternative to biopsy for staging chronic hepatitis C infection.

Indications

In patients with chronic hepatitis C infection, the likelihood of progression is closely correlated with the presence and severity of liver fibrosis (Chou et al., 2013). Thus, tests to diagnose the presence and ascertain the degree of fibrosis are indicated in the staging of patients with chronic hepatitis C, particularly when that information is relevant to decisions about HCV treatment. For instance, accurate determination of fibrosis stage is essential when treatment eligibility decisions are made on the basis of fibrosis severity. Beyond decisions about HCV treatment, tests to determine the presence of cirrhosis may be indicated in order to ensure appropriate supportive care and screening for complications of cirrhosis for these patients.

Until recently, the only options for staging fibrosis in hepatitis C patients was histological examination of the liver by percutaneous, transjugular, transfemoral, or laparoscopic surgical biopsy. However, biopsy entails procedural risks (including bleeding, infection, and pain), and the results are prone to sampling and interpretation errors. Despite these drawbacks, liver biopsy remains the "gold standard" for the diagnosis of fibrosis and cirrhosis (Chou et al., 2013).

The accuracy of noninvasive tests of liver fibrosis are measured against the reference standard of the results from a liver biopsy, using these definitions:

- **Sensitivity** refers to the proportion of patients who actually have the condition in question who have a positive test result.
- **Specificity** refers to the proportion of patients who really do not have the condition in question who have a negative test result.
- **Positive likelihood ratio** is the ratio of the probability of a positive test result in a patient with the condition to the probability of a positive test result in a patient without the condition. Likelihood ratios are most useful when the pre-test probability of the condition is known and the post-test probability at which treatment would be recommended is well established.
- **Negative likelihood ratio** is the ratio of the probability of a negative test in a patient with the condition to the probability of a negative test in a patient without the condition.
- The receiver operating curve **(ROC)** is a graphical illustration of the trade-off between sensitivity and specificity for an index diagnostic test (specifically for a test that has continuous rather than binary, or yes/no results) compared to a reference standard. The "index" test refers to the test that we are looking at to see how good it is. The reference standard has sometimes been referred to as the "gold standard," but given that some reference standards are not themselves perfectly accurate the terminology has shifted to "reference standard."

- The area under the receiver operating curve **(AUROC)** is an overall measure of how well the index test compares to the reference standard across a range of possible cutoffs. An index test that has cutoff value that allows perfect sensitivity and specificity (i.e. perfect classification of those with and without the condition) would have an AUROC of 1.0, while an AUROC of 0.5 represents a useless test (no better than a coin flip, on average). A test with an AUROC of 0.80-0.89 is generally regarded as a good test, while tests with an AUROC >0.90 are regarded as excellent tests. These distinctions are conventional, but arbitrary.

Technology description

Noninvasive techniques for staging liver fibrosis include imaging and blood tests. Five types of imaging tests are available: transient elastography (TE), acoustic radiation force impulse imaging (ARFI), shear wave elastography (SWE), magnetic resonance elastography (MRE), and real-time tissue elastography (RTE).

Transient Elastography (FibroScan®) measures the velocity of a low-frequency (50 Hz) elastic shear wave propagating through the liver. The velocity of the wave indicates the tissue stiffness, with the stiffer the tissue, the faster the shear wave propagates. The patient lies supine during the procedure, which takes less than five minutes.

Acoustic radiation force impulse imaging (Virtual Touch™ tissue quantification, ElastPQ) measures the speed of short-duration acoustic pulses that propagate shear waves and generate localized displacements in liver tissue. Commercial ultrasound machines can be easily modified to implement ARFI.

Shear wave elastography (Aixplorer® Supersonic Imagine) creates ultrasonic beams that are focused on liver tissues, and a very high frame rate ultrasound imaging sequences monitors the transient propagation of the shear waves in real time. This procedure can be implemented on commercial ultrasound machines.

Magnetic resonance elastography images the propagation characteristics of a shear wave in the liver using a modified phase-contrast method. Almost the entire liver can be analyzed with MRE, and it can be used effectively in patients with obesity or ascites. This procedure is more costly and more time consuming than the other imaging techniques.

Real-time tissue elastography constructs elasticity images of the liver by measuring the tissue strain induced by compression from a high-frequency ultrasound scanner. Tissue compression produces strain in the tissue, where the strain is smaller in harder tissue than in softer tissue.

Five proprietary blood testing protocols are available in the U.S., which use a combination of biochemical markers and patented algorithms to determine fibrosis stage. There are 25 additional blood tests that are not proprietary. The components of these blood tests are shown in Table 1 below. The most common components of the blood tests are platelet count, aspartate aminotransferase (AST), and alanine aminotransferase (ALT). About half of the tests include patient's age in the algorithm.

Table 1: Blood Tests for Measuring Liver Fibrosis in Patients with Hepatitis C

Blood tests	Components of test/algorithm
Proprietary tests	
ELF™ Test (Enhanced Liver Fibrosis)	Hyaluronic acid, tissue inhibitor of metalloproteinase 1, and procollagen III amino terminal peptide
FibroMeter™	Alanine aminotransferase (ALT), α_2-macroglobulin, gamma-glutamyl transpeptidase (GGT), aspartate aminotransferase (AST), platelet count, prothrombin index, urea, and patient's age and gender
FIBROSpect® II	Hyaluronic acid, tissue inhibitor of metalloproteinase, and α_2-macroglobulin
FibroSure® (FibroTest®)	α2-macroglobulin, haptoglobin, apolipoprotein A1, total bilirubin and gamma-glutamyl transpeptidase (GGT), and patient's age and gender ActiTest® is similar, with the addition of alanine aminotransferase (ALT)
Hepascore® (FibroScore®)	α_2-macroglobulin, hyaluronic acid, gamma-glutamyl transferase (GGT), bilirubin, and patient's age and gender
Non-proprietary tests	
Age–platelet index	Platelet count and patient's age
AST–platelet ratio index (APRI)	Platelet count and aspartate aminotransferase (AST)
AST–ALT ratio	Aspartate aminotransferase (AST) and alanine aminotransferase (ALT)
Cirrhosis discriminant score (Bonacini index)	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), prothrombin index, presence of ascites, and presence of spider angiomata
FIB-4	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and patient's age
Fibro-α score	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and α-Fetoprotein
FibroIndex	Platelet count, aspartate aminotransferase (AST), and gamma globulin
Fibronectin discriminant score	Platelet count, aspartate aminotransferase (AST), albumin, and fibronectin
FibroQ	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), prothrombin index, and patient's age
Fibrosis–cirrhosis index	Platelet count, Alkaline phosphatase, bilirubin, and albumin

Blood tests	Components of test/algorithm
Fibrosis index	Platelet count and albumin
Fibrosis probability index (Sud index)	Aspartate aminotransferase (AST), total cholesterol, insulin resistance, alcohol intake, and patient's age
Fibrosis–protein index	α_2-macroglobulin and hemopexin
Fibrosis Routine Test	Platelet count, aspartate aminotransferase (AST), α-Fetoprotein, albumin, and patient's age
Forns index	Platelet count, gamma-glutamyl transpeptidase (GGT), cholesterol, and patient's age
Globulin–albumin ratio	Globulin and albumin
Göteborg University Cirrhosis Index (GUCI)	Platelet count, aspartate aminotransferase (AST), and prothrombin index
HALT-C model (Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis)	Platelet count, tissue metalloproteinase inhibitor 1 (TIMP-1), and hyaluronic acid
King's score	Platelet count, aspartate aminotransferase (AST), international normalized ratio (INR), and patient's age
Lok index	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and international normalized ratio (INR)
MP3 score	Matrix metalloproteinase-1 (MMP-1) and procollagen III propeptide
Pohl index	Platelet count, aspartate aminotransferase (AST), and alanine aminotransferase (ALT)
Sabadell NIHCED index (Noninvasive Hepatitis-C–Related Cirrhosis Early Detection)	Platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), prothrombin time, right hepatic lobe atrophy, splenomegaly, caudate lobe hypertrophy, and patient's age
Significant fibrosis index	Haptoglobin, α_2-macroglobulin, tissue metalloproteinase inhibitor 1 (TIMP-1), matrix metalloproteinase-2 (MMP-2), and gamma-glutamyl transpeptidase (GGT)
Zeng index	α_2-macroglobulin, gamma-glutamyl transpeptidase (GGT), hyaluronic acid, and patient's age

Adapted from Chou & Wasson (2013)

Key Questions and Outcomes

The following key questions (KQ) guided the evidence search and review described below. For additional details about the review scope and methods please see Appendix C.

1. What is the comparative effectiveness of noninvasive tests for the diagnosis and management of hepatic fibrosis in patients with chronic hepatitis C?
2. Does the comparative effectiveness of noninvasive tests of liver fibrosis in patients with chronic hepatitis C vary based on:
    a. Duration of infection
    b. Fibrosis score
    c. Body habitus
    d. Operator/interpreter training or experience
    e. Co-existence of other etiologies of liver disease (e.g., non-alcoholic steatohepatitis)
3. What are the comparative diagnostic operating characteristics of tests of liver fibrosis?
4. What is the evidence for the timing of the initial testing for fibrosis and intervals for subsequent reassessment of fibrosis?

Critical outcomes selected for inclusion in the GRADE table were hepatitis-related morbidity/progression, need for liver biopsy, and quality of life. Important outcomes selected for inclusion in the GRADE table were testing-related adverse events and change in treatment plan (especially a decision to begin antiviral therapy).

Evidence Review

We identified no randomized controlled evidence on the use of noninvasive tests of liver fibrosis compared to liver biopsy with respect to clinical outcomes in hepatitis C infection.

We identified a poor quality systematic review and meta-analysis of six studies reporting on the relative prognostic value of liver biopsy, FibroTest®, FIB-4, and APRI for predicting overall survival. All of the tests offered statistically significant prognostic value for overall survival with AUROCs of 0.58 for APRI (95% CI 0.53 to 0.63), 0.68 for FIB-4 (95% CI 0.58 to 0.78), 0.77 for biopsy (95% CI 0.62 to 0.93), and 0.80 for FibroTest® (95% CI 0.76 to 0.95). The authors did not describe the methodologic rigor of the included studies. There was significant heterogeneity in the included studies (for example, in one study of APRI and FIB-4 in HCV patients, 68% of the patients had HIV co-infection). Lastly, the review was authored by the inventor of the FibroTest® and two employees of the company that market the test.

A more recent study (Vergniol et al., 2014) examined the prognostic value of evolving measurements of liver stiffness. In this study, about 1,025 people with chronic hepatitis C and two recorded measurements of liver stiffness (separated by >1,000 but <1,500 days) recorded between 2004 and 2008 were included. The average age of included patients was 52 years, half were men, the average BMI was 25 kg/m^2, and about 12% reported excessive alcohol consumption. During the mean follow-up period of three years (after the second measurement of liver stiffness), 16% of patients achieved sustained

virologic response from HCV treatment. Survival data was available for 95% of patients; of those, 35 patients had died and 7 had undergone liver transplantation. Twenty-one of the deaths were from liver-related causes. In the univariate analysis, several factors were associated with statistically significantly increased hazard ratios for death: age (HR 1.03, 95% CI 1.01 to 1.06), male sex (HR 2.25, 95% CI 1.17 to 4.43), baseline liver stiffness measurement (HR 4.27, 95% CI 2.94 to 6.22), follow-up liver stiffness measurement (HR 5.47, 95% CI 3.82 to 7.84), and change in liver stiffness measurement (HR 1.25, 95% CI 1.16 to 1.36). Unusually, alcohol abuse appeared to have a protective effect in this study (HR 0.42, 95% CI 0.18 to 0.97). In the multivariate analysis, baseline liver stiffness measurement (HR 5.76, 95% CI 3.74 to 8.87), change in liver stiffness measurement (HR 1.19, 95% CI 1.11 to 1.28), and achievement of SVR (HR 0.19, 95% CI 0.05 to 0.80) were statistically significant independent predictors of death. Overall, the authors concluded that patients with low-baseline liver stiffness measurements, those who achieve SVR, and those with non-cirrhotic baseline liver stiffness measurements and stable or decreasing measurements at follow-up all have an excellent prognosis. Conversely, patients with cirrhotic baseline liver stiffness measurement or those with advancing significant fibrosis have a poorer prognosis.

Cross-sectional data has correlated liver stiffness measurements by TE with the presence of portal hypertension (Kim et al., 2013), but TE has not been demonstrated in prospective studies to predict clinical outcomes related to portal hypertension in hepatitis C patients. A prospective cohort study of nearly 900 Japanese patients with HCV investigated the correlation between liver stiffness measurements by TE and the development of hepatocellular carcinoma (HCC) over a mean follow-up of 3 years (Masuzaki et al., 2009). Compared to a reference value of less than 10 kilopascals (kPa), various cut-offs of liver stiffness were associated with relative risk of HCC ranging from 16 to 45.

The remainder of the identified systematic reviews summarized diagnostic accuracy studies of various tests compared to a reference standard of liver biopsy. Most of these studies report diagnostic performance by way of sensitivity, specificity, and AUROC. A test that perfectly matches the diagnoses assigned by the reference test would have an AUROC of 1. Conventionally, tests with an AUROC of 0.9 to 1 are considered excellent, 0.8-0.89 are good, 0.7-0.79 are fair, and below 0.7 are poor, and though widely used, these distinctions are arbitrary.

Magnetic Resonance Elastography

Singh et al., 2015

This is a good quality systematic review and meta-analysis of patient-level data to determine the diagnostic performance of magnetic resonance elastography (MRE) compared to liver biopsy as the reference standard. The use of patient-level data in the meta-analysis allowed them to perform stratified analyses to determine if the diagnostic performance of MRE varied based on sex, obesity, or the etiology of the liver disease, and also allowed the authors to reduce the risk of spectrum bias and standardize diagnostic cut-offs for various fibrosis stages. The authors included 12 studies that met inclusion criteria and for which they were able to obtain the individual participant data (n=697). Overall, the included studies were judged to be at low to moderate risk of bias. Three of the studies did not adequately report on blinding procedures, raising the possibility of review bias.

Among the included patients, the average age was 55 years old, the majority were males (60%), and the average BMI was 27. Nearly half of the participants had HCV-related liver disease (47%), with smaller numbers of patients with HBV, NAFLD, ALD, AIH, or other miscellaneous etiologies. The distribution of fibrosis level on biopsy was 19.5% F0, 19.4% F1, 15.5% F2, 15.9% F3, and 29.7% F4.

The diagnostic operating characteristics of MRE from the meta-analysis, including both positive and negative likelihood ratios, are reported in Table 2 below.

Table 2: Diagnostic Operating Characteristics of MRE

Fibrosis Stage	AUROC (95% CI)	Sensitivity	Specificity	Positive LR	Negative LR
Any: ≥F1	0.84 (0.76 - 0.92)	0.73	0.79	3.48	0.34
Significant: ≥F2	0.88 (0.84 - 0.91)	0.79	0.81	4.16	0.26
Advanced: ≥F3	0.93 (0.90 - 0.95)	0.85	0.85	5.67	0.18
Cirrhosis: F4	0.92 (0.90 - 0.94)	0.91	0.81	4.79	0.11

In the subgroup and sensitivity analysis, the diagnostic performance of MRE did not significantly vary based on sex, presence of obesity, or etiology of liver disease. In this review, MRE had a failure rate of about 4%, and this was most commonly due to interference from hepatic iron overload.

Overall, the authors concluded that MRE was highly accurate for diagnosing fibrosis and cirrhosis regardless of BMI or the etiology of chronic liver disease.

Transient Elastography

Steadman et al., 2013

This is a good-quality, comprehensive technology assessment of transient elastography (TE) for the diagnosis of significant fibrosis in adults with chronic liver disease. Overall, 57 studies reporting diagnostic performance of TE compared with liver biopsy were included. The results were stratified by the etiology of liver disease, and 13 of the included studies were in patients with HCV. The included studies were methodologically rigorous with the authors rating nearly 80% of them as high quality.

The diagnostic operating characteristics of TE (in HCV patients only) from the meta-analysis are reported in Table 3 below.

Table 3: Diagnostic Operating Characteristics of Transient Elastography

Fibrosis Stage	AUROC (95% CI)	Sensitivity	Specificity	Positive LR	Negative LR
Significant: ≥F2	0.89 (0.86 - 0.91)	0.76	0.86	5.43	0.28
Advanced: ≥F3	0.92 (0.89 - 0.94)	0.88	0.91	9.7	0.13
Cirrhosis: F4	0.94 (0.92 - 0.96)	0.85	0.91	9.4	0.16

The authors also performed a basic economic analysis to calculate the incremental cost per correct diagnosis gained by liver biopsy over TE. In the subgroup of patients with HCV, the incremental cost per correct diagnosis using biopsy ranged from $1,861 for patients with F2 disease to $3,260 for patients with F3 disease. The authors were careful to note that their economic modeling does not account for the practice of monitoring progression of liver fibrosis and observe that the common practice in Alberta, Canada is yearly TE and biopsy every 3-5 years.

Overall, the authors concluded that TE was an accurate method for diagnosing fibrosis or cirrhosis and was less costly than liver biopsy.

Acoustic Radiation Force Impulse Imaging

Nierhoff et al., 2013

This is a good-quality systematic review and meta-analysis of the diagnostic operating characteristics of ARFI in patients with chronic liver disease using liver biopsy as the reference standard. The authors included 36 studies (both published manuscripts and abstracts) of nearly 4,000 patients. Among the included studies, 7 examined only patients with HCV as the etiology of their liver disease while another 18 studies reported on populations with mixed etiologies of chronic liver disease, including HCV. The methodologic quality of the included studies was mixed, and about half of the studies had potential flaws related to spectrum bias (bias introduced because the range and distribution of disease severity in the study is not representative of the overall population of people with the condition) and review bias (bias introduce when the interpreter of the index test is already aware of the result of the reference test, or vice-versa). The main reported measure of diagnostic performance was AUROC. The results of the meta-analysis of the HCV only and mixed etiology studies are reported in Table 4 below.

Table 4: AUROC of Acoustic Radiation Force Impulse (ARFI) Imaging Tests

Fibrosis Stage	AUROC – HCV only studies (95% CI)	AUROC – Mixed studies (95% CI)
Significant: ≥F2	0.88 (0.81 - 0.96)	0.83 (0.80 - 0.86)
Advanced: ≥F3	0.93 (0.89 - 0.97)	0.87 (0.85 - 0.90)
Cirrhosis: F4	0.92 (0.85 - 0.99)	0.91 (0.89 - 0.93)

One possible explanation for the poorer diagnostic performance in the mixed studies is the finding in subgroup analysis that higher BMI is associated with reduced diagnostic accuracy and a higher failure rate for testing.

Overall, the authors concluded that the diagnostic performance of ARFI is good to excellent for detecting fibrosis and cirrhosis. The authors also note that their findings are consistent with those of an earlier, smaller meta-analysis of ARFI using individual participant data.

Acoustic Radiation Force Impulse (ARFI) vs. Transient Elastography (TE)

Bota et al., 2013

This is a good-quality systematic review and meta-analysis of studies comparing ARFI and TE to a reference standard of liver biopsy for the evaluation of fibrosis. The authors included 13 trials; 10 of the trials reported diagnostic accuracy of ARFI and TE for the diagnosis of significant fibrosis (≥F2), and all the trials reported diagnostic accuracy for cirrhosis (F4). The etiology of liver disease in each study was variable, and all but one study included patients with chronic hepatitis C. The authors observed that failure rates (i.e. inability to obtain any valid measurements) were higher for TE (6.6%) than ARFI (2.1%), and five of the trials only included patients with valid ARFI and TE. The authors' risk of bias assessment for most studies was low. The results of the meta-analysis are reported in Table 5 below.

Table 5: Diagnostic Operating Characteristics of ARFI and TE

Test and Fibrosis Stage	AUROC (95% CI)	Sensitivity	Specificity	Positive LR	Negative LR
ARFI: ≥F2	0.85 (0.82 - 0.88)	0.74	0.83	4.29	0.31
TE: ≥F2	0.87 (0.83 - 0.89)	0.78	0.84	4.79	0.26
ARFI: F4	0.93 (0.91 - 0.95)	0.87	0.87	6.48	0.15
TE: F4	0.93 (0.91 - 0.95)	0.89	0.87	6.79	0.13

Overall, the authors concluded that there were no significant differences in the diagnostic accuracy of ARFI and TE. They note that while the higher failure rate for TE is concerning, new and more sensitive probes may mitigate this limitation.

Blood Tests

Dozens of blood tests and related interpretive indices or scores have been proposed for the diagnosis of fibrosis or cirrhosis in patients with HCV. The components of these tests are discussed in detail in the technology description section of this report.

Chou & Wasson, 2013

This is a good-quality systematic review of blood tests for the diagnosis of fibrosis and cirrhosis in patients with HCV. The authors did not perform a meta-analysis but present results for measures of diagnostic accuracy as medians and ranges. The number of studies for each test and the authors' GRADE assessment of the strength of evidence are provided in Table 6 below.

The results of the review of these tests are also summarized in Table 6. Because of the large number of tests as well as the various cut-offs used for each test, only the AUROC (median and range) are presented in this table.

Table 6: Studies of Blood Tests for Liver Fibrosis

Test	Number of studies	Strength of evidence	Fibrosis (≥F2) AUROC median (range)	Cirrhosis AUROC median (range)
Platelet count	18	Moderate	0.71 (0.38 - 0.94)	0.89 (0.64 - 0.99)
Hyaluronic acid	8	Moderate	0.75 (0.65 - 0.88)	0.90 (0.80 - 0.97)
Age-platelet index	11	Moderate	0.74 (0.64 - 0.79)	0.86 (0.64 - 0.91)
AST-platelet ratio index	7	High	0.77 (0.58 - 0.95)	0.84 (0.54 - 0.97)
AST-ALT ratio	32	High	0.59 (0.50- 0.82)	0.72 (0.52 - 0.91)
Bonacini index	12	Moderate	0.66 (0.58 - 0.71)	0.74 (0.61 - 0.91)
ELF™	8	Moderate	0.81 (0.72 - 0.87)	0.88 (0.78 - 0.91)
FIB-4	19	Moderate	0.74 (0.61 - 0.81)	0.87 (0.83 - 0.92)
FibroIndex	9	Moderate	0.76 (0.58 - 0.86)	0.86 (0.78 - 0.92)
Fibrometer™	8	Moderate	0.82 (0.78 - 0.85)	0.91 (0.89 - 0.94)
FIBROSpect® II	7	Low	0.86 (0.77 - 0.90)	NR
FibroTest®	32	High	0.79 (0.70 - 0.89)	0.86 (0.71 - 0.92)
Forns index	22	High	0.76 (0.60 - 0.86)	0.87 (0.85 - 0.91)
GUCI	5	Low	NR	0.82 (0.78 - 0.86)
Hepascore®	12	High	0.79 (0.69 - 0.82)	0.89 (0.88 - 0.94)
Lok index	10	Moderate	NR	0.80 (0.61 - 0.91)
Pohl index	12	Low	0.52 (0.52 - 0.53)	0.65 (0.64 - 0.66)

The Chou & Wasson review also summarized the results of trials making direct comparisons between APRI or FibroTest® and various other blood tests. Very few of these direct comparisons showed substantial differences in the median AUROC for fibrosis, but median differences in excess of 0.05 are reported in Table 7 below. Only one of the direct comparisons (APRI vs. AST-ALT ratio) for the diagnosis of cirrhosis exceed a median difference in AUROC of greater than 0.05; in those studies APRI was more accurate than the AST-ALT ratio.

Table 7. Studies of Direct Comparisons between Two Blood Tests

Number of studies	Test A AUROC median	Test B AUROC median	Median difference (range)
13	APRI 0.76	AST-ALT ratio 0.58	0.17 (-0.06 to 0.23)
4	APRI 0.74	Bonacini index 0.66	0.08 (0.07 to 0.09)
8	APRI 0.79	Fibrometer™ 0.84	-0.06 (-0.07 to -0.02)
8	APRI 0.76	Platelet count 0.67	0.08 (-0.06 to 0.53)
3	APRI 0.69	Pohl index 0.52	0.17 (0.13 to 0.23)
3	FibroTest® 0.78	FibroIndex 0.72	0.08 (0.02 to 0.10)

The authors also include 9 studies that report on the use of combinations of blood tests or indices. Four studies reported on diagnostic performance of the Sequential Algorithm for Fibrosis Evaluation that combines results from APRI and FibroTest®. In two studies of patients with fibrosis (≥F2), the algorithm had an AUROC of 0.90 and 0.94. In 3 studies of cirrhosis, the algorithm had a median AUROC of 0.87. The remaining combinations of tests or indices were only studied in single trials.

The authors point out several limitations of the review, the most important of which is the binary interpretation of presence or absence of clinically significant fibrosis. As they note, "Measures that incorporate the accuracy of tests at each fibrosis stage would therefore be more informative than estimates based on dichotomized classifications." Additionally, because nearly all the included studies grouped patients with both lesser stages of fibrosis and cirrhosis, it was not possible to ascertain the diagnostic performance of blood tests for less severe fibrosis independent from the diagnostic accuracy of the full spectrum of significant fibrosis, and distinguishing between F2 and F3 is not possible. Overall, the authors conclude that a variety of blood tests are moderately useful for the identification of clinically significant fibrosis in patients with HCV.

Shear Wave Elastography

Li et al., 2016

This is a good-quality systematic review and meta-analysis of diagnostic accuracy studies of real-time shear wave elastography (SWE) for staging liver fibrosis. The authors identified eight studies with a total of 934 patients comparing SWE to a reference standard of liver biopsy. Most patients in the included studies had chronic viral hepatitis, but the precise breakdown was not provided. The included studies were generally at low risk of bias, though three were judged to be susceptible to disease progression

bias because of the time difference between the two tests. The diagnostic operating characteristics from the meta-analysis are reported in Table 8 below.

Table 8. Diagnostic Operating Characteristics for Shear Wave Elastography

Fibrosis Stage	AUROC (95% CI)	Sensitivity	Specificity	Positive LR	Negative LR
Significant: ≥F2	0.88 (0.85 - 0.91)	0.85	0.81	4.47	0.18
Advanced: ≥F3	0.94 (0.92 - 0.96)	0.90	0.81	4.73	0.12
Cirrhosis: F4	0.92 (0.89 - 0.94)	0.87	0.88	7.25	0.15

The authors note that the primary limitations of their review include the small number of studies and the inability to perform subgroup analysis by etiology of chronic liver disease.

The authors observe that compared with reported diagnostic accuracy of other modalities, SWE is comparable to TE and ARFI for diagnosis of cirrhosis, and comparable to ARFI but better than TE for the diagnosis of significant fibrosis (≥F2). Overall, the authors conclude that the diagnostic accuracy of SWE for fibrosis staging is good.

Real-Time Tissue Elastography

Kobayashi et al., 2014

This is a good-quality systematic review and meta-analysis of diagnostic accuracy studies of real-time tissue elastography (RT-TE) compared to a reference standard of liver biopsy. The authors identified 15 trials including over 1,600 patients. Ten of 15 studies included patients with HCV. The authors expressed concerns over the risk of bias in several included studies related to patient selection bias and the absence of pre-specified cut-off values for the index tests. They also identified possible publication bias in their funnel plots. The meta-analytic results for sensitivity and specificity are reported in Table 9 below.

Table 9. Diagnostic Operating Characteristics for Real-Time Tissue Elastography

Fibrosis Stage	AUROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)
Significant: ≥F2	0.69 (NR)	0.79 (0.75 - 0.83)	0.76 (0.68 - 0.82)	3.29 (NR)	0.27 (NR)
Advanced: ≥F3	0.86 (NR)	0.82 (0.75 - 0.88)	0.81 0.72 - 0.88)	4.31 (NR)	0.22 (NR)
Cirrhosis: F4	0.72 (NR)	0.74 (0.63 - 0.82)	0.84 0.79 - 0.88)	4.6 (NR)	0.30 (NR)

Overall, the authors conclude that, "RTE is not highly accurate for any cut-off stage of fibrosis."

Direct Comparisons of FibroTest®, FIB-4, APRI, and TE

Houot et al., 2016

This is a poor-quality systematic review and meta-analysis of trials making direct comparisons between FibroTest®, APRI, FIB-4, and TE compared to a reference standard of liver biopsy. The authors identified 71 trials, of which 37 included only patients with HCV. The main purpose of the review was to determine whether there were differences between the AUROC of these tests for the diagnosis of advanced fibrosis (defined here as ≥F2) or cirrhosis. The review did not provide information on the methodologic quality of the included studies. The authors applied three meta-analytic methods to ascertain whether the differences in test performance were statistically significant: an indirect pooled AUROC difference, a standard pooled AUROC difference, and a Bayesian pooled AUROC difference. Among the HCV-only studies, the differences in AUROC for most comparisons were generally small (<0.05). In the indirect pooled analysis, only one comparison showed a statistically significant difference in favor of TE over APRI for diagnosis of cirrhosis. In the standard pooled analysis FibroTest® was favored over TE and APRI for diagnosis of fibrosis; TE and FIB-4 were favored over APRI for the diagnosis of cirrhosis. In the Bayesian pooled analysis, FibroTest® was favored over APRI for the diagnosis of fibrosis and TE and FIB-4 were favored over APRI for the diagnosis of cirrhosis. This review is subject to potential conflict of interest as the senior author is the inventor of FibroTest® and the study was funded in part by BioPredictive, the company that markets FibroTest®.

Factors Influencing Accuracy of TE

Perazzo et al., 2015

This is a narrative review article that summarizes research on various factors that influence the accuracy and interpretation of transient elastography. The authors identify four factors that are associated with overestimation of fibrosis by TE: heightened necroinflammatory activity as denoted by alanine transaminases greater than 10 times the upper limit of normal, extrahepatic cholestasis and hepatic

congestion, non-fasting status, and the presence of severe steatosis. The authors also note that the reliability of TE measurements is modified by operator experience and propose a definition of an experienced operator as greater than 100 examinations. Similarly, large ranges of inter-observer variability are reported in the literature and discrepancies between assessments of adjacent fibrosis stages are more common. The authors suggest that longitudinal follow-up and examination by the same experienced operator may prove most accurate.

We did not identify any evidence that addresses the question of initial timing of staging or the appropriate intervals for re-staging using non-invasive tests. The systematic review of TE did observe that the common practice in Alberta, Canada is to perform non-invasive tests to assess fibrosis stage every 3 to 5 years.

EVIDENCE SUMMARY

Although an imperfect test itself, liver biopsy remains the reference standard by which noninvasive tests of liver fibrosis and cirrhosis are judged. There is no direct comparative evidence that examines the effects of different diagnostic strategies on the predetermined clinical outcomes:

- Hepatitis-related morbidity/progression
- Need for liver biopsy
- Quality of life
- Testing-related adverse events
- Change in treatment plan

Furthermore, there is only sparse evidence on the value and reliability of prognostic information obtained from noninvasive tests. However, there are a large number of studies comparing the diagnostic accuracy of noninvasive tests of liver fibrosis to the reference standard of liver biopsy. Many of these studies (see Appendix D) demonstrated good or excellent performance of non-invasive tests for the detection of various levels of fibrosis; in general, imaging studies appear to have greater ability to distinguish between intermediate stages of fibrosis (i.e. between F2 and F3), while blood tests appear to be suitable for establishing the presence of significant fibrosis (≥F2) or cirrhosis (F4).

OTHER DECISION FACTORS

Resource Allocation

The price of noninvasive tests is generally significantly less than liver biopsy and avoids the costs associated with harms from liver biopsy. However, noninvasive testing is likely to be done at a higher frequency than liver biopsy and the increased number of total procedures may somewhat reduce the cost-savings associated with avoiding liver biopsy. The more significant cost driver is the impact noninvasive testing may have on determining the eligible population for treatment with hepatitis C. Health plans have prioritized treatment of hepatitis C patients with the newer expensive medications both because of the high cost of these medications and the prevalence of chronic hepatitis C infection in

the general population. The cutoff point for some plans in Oregon include only treating persons with a score of F3 or above. This requires testing that can accurately distinguish between the cutoff points for treatment. If a test has a high false positive rate, that would lead more people into a hepatitis C treatment pathway (increasing overall costs of the population in the near term). If a test has a high false negative rate, then people with more advanced fibrosis who may particularly benefit from treatment would not qualify for treatment (decreasing health system costs, but at the expense of fewer eligible people receiving appropriate treatment).

Values and preferences

Patients would highly value avoiding an invasive procedure as long as the information provided by a noninvasive test was comparable. There would be minimal variability in this preference. From a population perspective, it would be very important that these tests can accurately distinguish between those persons who would benefit the most from the very expensive treatment versus others who may be able to delay or avoid treatment altogether.

POLICY LANDSCAPE

Quality measures

No quality measures were identified when searching the National Quality Measures Clearinghouse.

Payer coverage policies

The Oregon Medicaid fee-for-service Approval Criteria for Hepatitis C Direct-Acting Antivirals requires liver fibrosis staging by either:

- A biopsy, transient elastography (FibroScan®), or serum test (FibroSure®) to indicate advanced fibrosis (METAVIR F3) or cirrhosis (METAVIR F4)
- Radiologic, laboratory (APRI score >1.5 or FIB-4 score >3.25), or clinical evidence (ascites, portal hypertension) of cirrhosis

The Washington Health Care Authority outlines the treatment policy for patients with HCV, with the accepted diagnostic tests for liver damage including imaging procedures (FibroScan®, ARFI, SWE) and blood tests (FibroSure®, APRI). The Table 10 below shows the allowed tests and cutoffs used to stage liver fibrosis to determine hepatitis C treatments.

Table 10: Washington Health Care Authority Accepted Diagnostic Tests and Procedures to Stage Liver Damage in Patients with Chronic HCV Infection

METAVIR Score	Biopsy	FibroScan®	Elastography (ARFI/PSWE)	FibroSure®	APRI	Other Imaging
F4	F4	≥ 12.5 kPa	≥ 2.34 m/s	≥ 0.75	≥ 2.0	Cirrhosis
F3	F3	9.6 - 12.4 kPa	2.01 - 2.33 m/s	0.58 - 0.74	1.5 - 1.9	
F2	F2	7.1 - 9.5 kPa	1.38 - 2.0 m/s	0.49 - 0.57	1.0 - 1.4	
F1/0	F1/0	≤ 7.0 kPa	≤ 1.37 m/s	≤ 0.48	≤ 0.9	

On May 27, 2016, a United States District Court issued a preliminary injunction requiring the Washington Medicaid program to cover direct-acting antiviral medications for Medicaid clients with hepatitis C, regardless of the extent of liver fibrosis.

Coverage policies for noninvasive tests of liver fibrosis were searched for four commercial payers: Aetna, Cigna, Moda, and Regence. Transient elastography (FibroScan®) is covered by three of these payers: Aetna, Cigna, and Moda. MRE for staging liver fibrosis is covered by only Moda. None of the other imaging tests are covered by these payers. Three of the four payers do not cover the blood tests for staging liver fibrosis. Moda Health covers the blood tests FibroSure®, FIBROSpect®, APRI, ActiTest®, and Hepascore®.

Aetna's precertification criteria for direct-acting antivirals require the staging of liver disease by liver biopsy, METAVIR scores, FibroScan® score, APRI score, radiological imaging consistent with cirrhosis (i.e., evidence of portal hypertension), or physical findings or clinical evidence consistent with cirrhosis as attested by the prescribing physician. The Regence Medical Policy Manual states that, "Liver biopsy is typically recommended prior to the initiation of antiviral therapy." Coverage policies for direct-acting antivirals for Cigna and Moda do not indicate specific methods for staging of liver fibrosis.

For Medicare, no National Coverage Determinations or Local Coverage Determinations related to noninvasive tests for liver fibrosis were identified.

Professional society guidelines

American Association for the Study of Liver Disease (AASLD) and Infectious Disease Society of America (IDSA) Guideline, 2016

The AASLD and IDSA guideline endorses the use of biopsy, imaging, and/or noninvasive markers to evaluate advanced fibrosis in HCV patients for treatment planning and to ascertain whether additional screening and management of cirrhosis is needed (Class I, Level A). It also endorses the continued monitoring of liver disease in those who defer treatment, but does not specify the use of noninvasive tests or provide an optimal interval for re-assessment.

Regarding noninvasive tests, the AASLD and IDSA guideline makes the following statements:

- “No single method is recognized to have high accuracy alone and each test must be interpreted carefully. A recent publication of the Agency for Healthcare Research and Quality found evidence in support of a number of blood tests; however, at best, they are only moderately useful for identifying clinically significant fibrosis or cirrhosis.”
- “Vibration-controlled transient liver elastography is a noninvasive way to measure liver stiffness and correlates well with measurement of substantial fibrosis or cirrhosis in patients with chronic HCV infection. The measurement range does overlap between stages.”
- “The most efficient approach to fibrosis assessment is to combine direct biomarkers and vibration-controlled transient liver elastography. A biopsy should be considered for any patient who has discordant results between the 2 modalities that would affect clinical decision making. For example, one shows cirrhosis and the other does not. The need for liver biopsy with this approach is markedly reduced.”
- “Alternatively, if direct biomarkers or vibration-controlled transient liver elastography are not available, the AST-to-platelet ratio index (APRI) or FIB-4 index score can help, although neither test is sensitive enough to rule out substantial fibrosis. Biopsy should be considered in those in whom more accurate fibrosis staging would impact treatment decisions. Individuals with clinically evident cirrhosis do not require additional staging (biopsy or noninvasive assessment).”

European Association for the Study of the Liver (EASL) and Asociación Latinoamericano para el Estudio del Hígado (ALEH), 2015

This is a comprehensive clinical practice guideline on the use of noninvasive tests for evaluating liver disease across a variety of etiologies. In general, EASL/ALEH endorse the use of noninvasive tests of liver fibrosis. Specific recommendations and statements include:

- "Non-invasive tests should always be interpreted by specialists in liver disease, according to the clinical context, considering the results of other tests (biochemical, radiological and endoscopic) and taking into account the recommended quality criteria for each test and its possible pitfalls (A1)."
- "TE is a fast, simple, safe and easy to learn procedure that is widely available. Its main limitation is the impossibility of obtaining results in case of ascites or morbid obesity and its limited applicability in case of obesity and limited operator experience (A1)."
- "TE should be performed by an experienced operator (>100 examinations) following a standardized protocol with the patient, fasting for at least 2 hours, in the supine position, right arm in full abduction, on the midaxillary line with the probe-tip placed in the 9th to 11th intercostal space with a minimum of 10 shots (A1)."
- "Although alternative techniques, such as pSWE/ARFI or 2D-SWE seem to overcome limitations of TE, their quality criteria for correct interpretation are not yet well defined (A1)."
- "MR elastography is currently too costly and time consuming for routine clinical practice use and seems more suited for research purposes (A1)."
- "When compared in HCV patients, the different patented tests have similar levels of performance in diagnosing significant fibrosis and cirrhosis (A1). Although non-patented tests might have lower diagnostic accuracy than patented tests, they are not associated with additional costs, are easy to calculate, and are widely available (A2)."
- "Among the different available strategies, algorithms combining TE and serum biomarkers appear to be the most attractive and validated one (A2). In patients with viral hepatitis C, when TE and serum biomarkers results are in accordance, the diagnostic accuracy is increased for detecting significant fibrosis but not for cirrhosis. In cases of unexplained discordance, a liver biopsy should be performed if the results would change the patient management (A1)."

The EASL/ALEH guideline includes the following proposed algorithm for noninvasive testing in HCV patients.

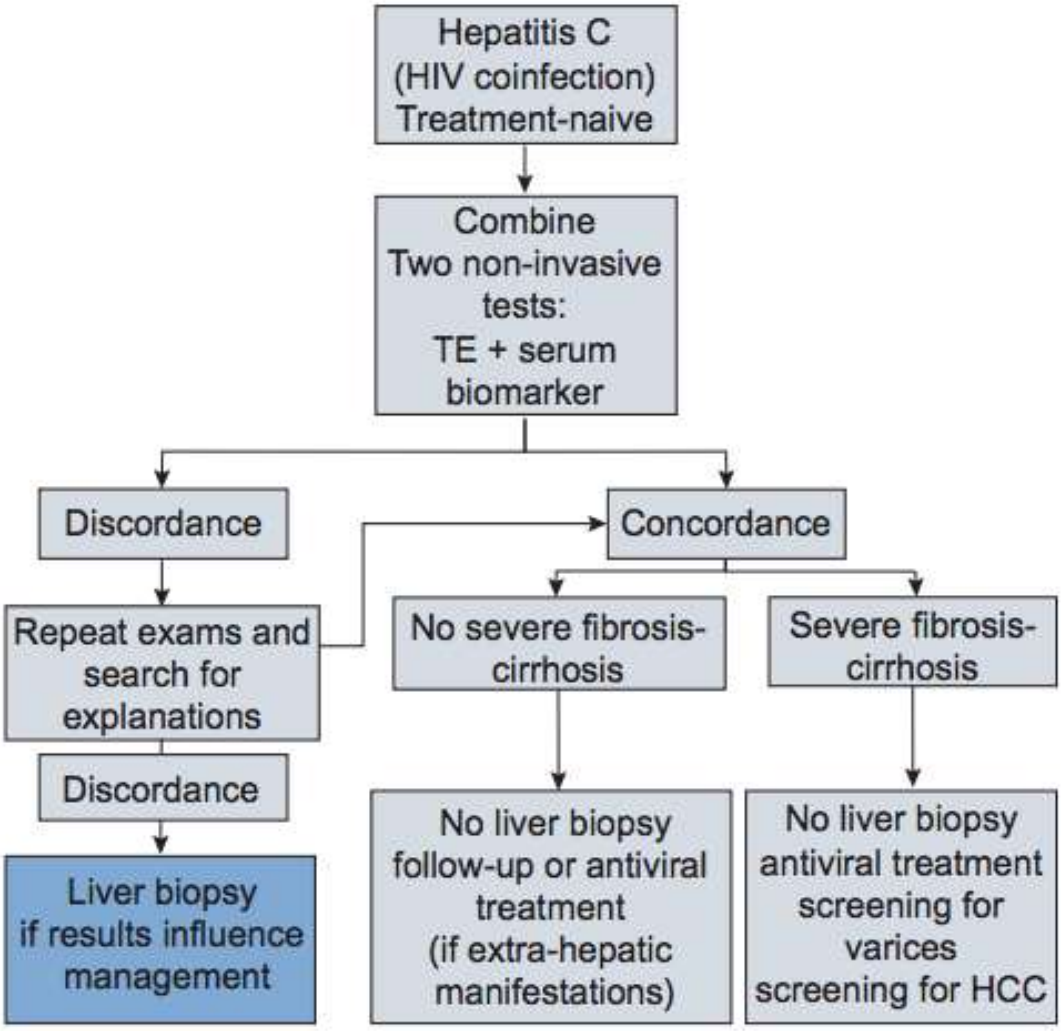

Fig. 1. Proposed algorithm for the use of non-invasive tests in treatment-naive patients with Hepatitis C with or without HIV coinfection.

National Institute for Health and Care Excellence (NICE), 2015

NICE issued medical technology guidance on the use of Virtual Touch™ Quantification (VTq, a proprietary system for performing ARFI) for diagnosing and monitoring liver fibrosis in chronic hepatitis B and C. The panel endorsed the use of VTq as an option for assessing liver fibrosis in chronic hepatitis B or C. They concluded that VTq is as accurate as transient elastography and cost modelling suggested that VTq would likely to be cost saving compared to transient elastography and liver biopsy.

Scottish Intercollegiate Guidelines Network (SIGN), 2013

SIGN published a comprehensive guideline on the management of hepatitis C in 2013 including recommendations regarding the use of noninvasive tests for diagnosing fibrosis and cirrhosis. The SIGN guideline states that while biochemical markers may be able to distinguish cirrhosis from less degrees of fibrosis, "intermediate stages are not distinguishable." Thus, SIGN recommends that biochemical markers should not be considered an alternative to biopsy for staging intermediate levels of fibrosis, but may be used in place of biopsy to diagnose cirrhosis (B recommendations, 2++ evidence). The guideline does offer that measurement of liver stiffness by noninvasive testing may be considered a "recommended best practice based on the clinical experience of the guideline development group."

Society of Radiologists in Ultrasound Consensus Conference Statement, 2015

This consensus conference statement (Barr et al., 2015) asserts that elastography (using either ultrasound or magnetic resonance techniques) can be used to diagnose liver fibrosis in patients "without overt decompensated cirrhosis." The panel stated that elastography should be used to group patients into three categories: those with minimal fibrosis (F0 or F1), those with a high likelihood of cirrhosis (F4), and those with values in between suggesting moderate to severe fibrosis (F2 and F3). The panel also proposed consensus diagnostic thresholds which are reproduced in Table 11.

Table 11: Consensus of Suggested Thresholds in Patients with Hepatitis C

Device	No Clinically Significant Fibrosis: METAVIR Stage < F2, Unlikely to Need Follow-up	Advanced Fibrosis and/or Cirrhosis: METAVIR Stage of F4 and Some Stages of F3 – Clinically Significant Fibrosis
TE FibroScan® (Echosens)	<7 kPa (1.5 m/sec)	>15 kPa (2.2 m/sec)
Siemens pSWE	1.2 m/sec (Siemens suggests <1.34 m/sec, <5.6 kPa)	>2.2 m/sec (>15 kPa)
Philips pSWE	<5.7 kPa (1.37 m/sec)	>2.2 m/sec (>15 kPa)
2D SWE (SuperSonic Imagine)	<7 kPa (1.5 m/sec)	>2.2 m/sec (>15 kPa)
MR elastography (GE, Siemens, Philips)	<3.0 kPa* (27–30)	>5.0 kPa*

*MR elastography is reported as shear modulus, while U.S. elastography techniques are reported in Young modulus. The Young modulus is three times the shear modulus.

World Health Organization, 2014

The WHO released a comprehensive guideline in 2014 focused on management of hepatitis C in resource limited settings. In general, the guideline states that noninvasive tests should be favored over liver biopsy and "in resource-limited settings, it is suggested that aminotransferase/platelet ratio index (APRI) or FIB4 be used for the assessment of hepatic fibrosis rather than other noninvasive tests that

require more resources such as elastography or Fibrotest." (Conditional recommendation, low quality evidence)

REFERENCES

Evidence Sources

Bota, S., Herkner, H., Sporea, I., Salzl, P., Sirli, R., Neghina, A. M., & Peck-Radosavljevic, M. (2013). Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. *Liver International, 33*(8), 1138-1147. DOI: 10.1111/liv.12240.

Chou, R., & Wasson, N. (2013). Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: A systematic review. *Annals of Internal Medicine, 158*(11), 807-820. DOI: 10.7326/0003-4819-158-11-201306040-00005.

Houot, M., Ngo, Y., Munteanu, M., Marque, S., & Poynard, T. (2016). Systematic review with meta-analysis: Direct comparisons of biomarkers for the diagnosis of fibrosis in chronic hepatitis C and B. *Alimentary Pharmacology & Therapeutics, 43*(1), 16-29. DOI: 10.1111/apt.13446.

Kim, B. K., Fung, J., Yuen, M. F., & Kim, S. U. (2013). Clinical application of liver stiffness measurement using transient elastography in chronic liver disease from longitudinal perspectives. *World Journal of Gastroenterology, 19*(12), 1890-1900. DOI: 10.3748/wjg.v19.i12.1890.

Kobayashi, K., Nakao, H., Nishiyama, T., Lin, Y., Kikuchi, S., Kobayashi, Y., ... Yoneda, M. (2015). Diagnostic accuracy of real-time tissue elastography for the staging of liver fibrosis: A meta-analysis. *European Radiology, 25*(1), 230-238. DOI: 10.1007/s00330-014-3364-x.

Li, C., Zhang, C., Li, J., Huo, H., & Song, D. (2016). Diagnostic accuracy of real-time shear wave elastography for staging of liver fibrosis: A meta-analysis. *Medical Science Monitor, 22*, 1349-1359. DOI: 10.12659/MSM.895662.

Masuzaki, R., Tateishi, R., Yoshida, H., Goto, E., Sato, T., Ohki, T., … Omata, M. (2009). Prospective risk assessment for hepatocellular carcinoma development in patients with chronic hepatitis C by transient elastography. *Hepatology, 49*(6), 1954-1961. DOI: 10.1002/hep.22870

Nierhoff, J., Chavez Ortiz, A. A., Herrmann, E., Zeuzem, S., & Friedrich-Rust, M. (2013). The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: A meta-analysis. *European Radiology, 23*(11), 3040-3053. DOI: 10.1007/s00330-013-2927-6.

Perazzo, H., Veloso, V. G., Grinsztejn, B., Hyde, C., & Castro, R. (2015). Factors that could impact on liver fibrosis staging by transient elastography. *International Journal of Hepatology, 2015*, 624596. DOI: 10.1155/2015/624596

Singh, S., Venkatesh, S. K., Wang, Z., Miller, F. H., Motosugi, U., Low, R. N., ... Ehman, R. L. (2015). Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: A systematic review and meta-analysis of individual participant data. *Clinical Gastroenterology and Hepatology, 13*(3), 440-451.e446. DOI: 10.1016/j.cgh.2014.09.046.

Steadman, R., Myers, R. P., Leggett, L., Lorenzetti, D., Noseworthy, T., Rose, S., … Clement, F. (2013). A health technology assessment of transient elastography in adult liver disease. *Canadian Journal of Gastroenterology, 27*(3), 149–158. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3732152/

Vergniol, J., Boursier, J., Coutzac, C., Bertrais, S., Foucher, J., Angel, C., ... Cales, P. (2014). Evolution of noninvasive tests of liver fibrosis is associated with prognosis in patients with chronic hepatitis C. *Hepatology, 60*(1), 65-76. DOI: 10.1002/hep.27069.

Other Citations

American Association for the Study of Liver Diseases and Infectious Diseases Society of America. (2014). *Recommendations for testing, managing and treating hepatitis C*. Retrieved from http://www.hcvguidelines.org/

Barr, R. G., Ferraioli, G., Palmeri, M. L., Goodman, Z. D., Garcia-Tsao, G., Rubin, J., … Levine, D. (2015). Elastography assessment of liver fibrosis: Society of Radiologists in Ultrasound consensus conference statement. *Radiology, 276*(3), 845-861. DOI: 10.1148/radiol.2015150619.

Centers for Disease Control and Prevention. (2016). *Hepatitis C FAQs for health professionals*. Retrieved from http://www.cdc.gov/hepatitis/hcv/hcvfaq.htm#section1

European Association for the Study of the Liver (EASL). (2014). *EASL recommendations on the treatment of hepatitis C*. Retrieved from http://www.easl.eu/assets/application/files/bdb06ff135c7ccb_file.pdf

National Institute for Health and Care Excellence (2015). *Virtual Touch Quantification to diagnose and monitor liver fibrosis in chronic hepatitis B and C*. Retrieved from https://www.nice.org.uk/guidance/mtg27

Scottish Intercollegiate Guidelines Network (SIGN). (2013). *Management of hepatitis C.* Retrieved from http://www.sign.ac.uk/guidelines/fulltext/133/index.html

World Health Organization (WHO). (2014). *Guidelines for the screening, care and treatment of persons with hepatitis C infection*. Geneva, Switzerland: WHO. Retrieved from http://apps.who.int/iris/bitstream/10665/111747/1/9789241548755_eng.pdf?ua=1

Coverage guidance is prepared by the Health Evidence Review Commission (HERC), HERC staff, and subcommittee members. The evidence summary is prepared by the Center for Evidence-based Policy at Oregon Health & Science University (the Center). This document is intended to guide public and private purchasers in Oregon in making informed decisions about health care services.

The Center is not engaged in rendering any clinical, legal, business or other professional advice. The statements in this document do not represent official policy positions of the Center. Researchers involved in preparing this document have no affiliations or financial involvement that conflict with material presented in this document.

Suggested citation: Obley, A., Mosbaek, C., King, V., & Livingston, C. (2016). *Coverage guidance: Noninvasive testing for liver fibrosis in patients with chronic hepatitis C*. Portland, OR: Center for Evidence-based Policy, Oregon Health & Science University.

APPENDIX A. GRADE INFORMED FRAMEWORK – ELEMENT DESCRIPTIONS

Element	Description
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Resource allocation	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted
Other considerations	Other considerations include issue about the implementation and operationalization of the technology or intervention in health systems and practices within Oregon.

Strong recommendation

In Favor: The subcommittee is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Against: The subcommittee is confident that the undesirable effects of adherence to a recommendation outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Weak recommendation

In Favor: The subcommittee concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Against: The subcommittee concludes that the undesirable effects of adherence to a recommendation probably outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Quality or strength of evidence rating across studies for the treatment/outcome[1]

High: The subcommittee is very confident that the true effect lies close to that of the estimate of the effect. Typical sets of studies are RCTs with few or no limitations and the estimate of effect is likely stable.

Moderate: The subcommittee is moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Typical

[1] Includes risk of bias, precision, directness, consistency and publication bias

sets of studies are RCTs with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.

***Low*:** The subcommittee's confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Typical sets of studies are RCTs with serious limitations or nonrandomized studies without special strengths.

***Very low*:** The subcommittee has very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect. Typical sets of studies are nonrandomized studies with serious limitations or inconsistent results across studies.

APPENDIX B. GRADE EVIDENCE PROFILE

Quality Assessment for MRE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
12	Diagnostic accuracy studies (cross-sectional or cohort designs)	Low	Not serious	Serious	Not serious		Moderate confidence ●●●○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for TE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
2	Prospective prognostic studies	Moderate to high	Not serious	Serious	Serious		Very low confidence ●○○○
Need for liver biopsy (Critical outcome)							
57	Diagnostic accuracy studies (cross-sectional or	Low	Not serious	Serious	Not serious		Moderate confidence ●●●○

Quality Assessment for TE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
	cohort designs)						
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for ARFI (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
36	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Not serious		Low confidence ●●○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for SWE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
8	Diagnostic accuracy studies (cross-sectional or cohort designs)	Low to Moderate	Not serious	Serious	Not serious		Low confidence ●●○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for RT-TE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
15	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Unclear	Possible publication bias	Very low confidence ●○○○
Quality of life (Critical outcome)							

Quality Assessment for RT-TE (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Platelet count (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
18	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Hyaluronic acid (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
8	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Age-platelet index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
11	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Not Serious		Low confidence ●●○○
Quality of life (Critical outcome)							
0							Insufficient

Quality Assessment for Age-platelet index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for APRI (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
6	Retrospective prognostic studies	High	Not serious	Serious	Not serious		Very low confidence ●○○○
Need for liver biopsy (Critical outcome)							
7	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for AST-ALT ratio (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
32	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Bonacini index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
12	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Not serious		Low confidence ●●○○
Quality of life (Critical outcome)							

Quality Assessment for Bonacini index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for ELF™ (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
8	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for FIB-4 (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
6	Retrospective prognostic studies	High	Not serious	Serious	Not serious		Very low confidence ●○○○
Need for liver biopsy (Critical outcome)							
19	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for FibroIndex (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
9	Diagnostic accuracy studies (cross-sectional	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○

Quality Assessment for FibroIndex (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
	or cohort designs)						
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for FibroMeter™ (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
8	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for FIBROSpect® II (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
7	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for FibroTest® (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
6	Retrospective prognostic studies	High	No serious	Serious	Serious		Very low confidence ●○○○
Need for liver biopsy (Critical outcome)							
32	Diagnostic accuracy studies (cross-sectional or	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○

Quality Assessment for FibroTest® (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
	cohort designs)						
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Forns index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
7	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Hepascore® (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
12	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Serious		Very low confidence ●○○○
Quality of life (Critical outcome)							
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

Quality Assessment for Pohl index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
Hepatitis related morbidity/progression (Critical outcome)							
0							Insufficient
Need for liver biopsy (Critical outcome)							
12	Diagnostic accuracy studies (cross-sectional or cohort designs)	Moderate	Not serious	Serious	Not serious		Low confidence ●●○○
Quality of life (Critical outcome)							

Quality Assessment for Pohl index (Confidence in Estimate of Effect)							
No. of Studies	**Study Design(s)**	**Risk of Bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other Factors**	**Quality**
0							Insufficient
Testing related adverse events (Important outcome)							
0							Insufficient
Change in treatment plan (Important outcome)							
0							Insufficient

APPENDIX C. METHODS

Scope Statement

Populations

Adults and children with chronic hepatitis C infection

Population scoping notes: *None*

Interventions

Noninvasive tests of liver fibrosis (e.g., acoustic radiation force impulse imaging, transient elastography, magnetic resonance elastography, biochemical tests with predictive algorithms)

Intervention exclusions: *None*

Comparators

Liver biopsy, other interventions listed above

Outcomes

Critical: Hepatitis-related morbidity/progression, need for liver biopsy, quality of life

Important**:** Testing-related adverse events, change in treatment plan (especially decision to begin antiviral therapy)

Considered but not selected for the GRADE table: *None*

Key Questions

1. What is the comparative effectiveness of noninvasive tests for the diagnosis and management of hepatic fibrosis in patients with chronic hepatitis C?

2. Does the comparative effectiveness of noninvasive tests of liver fibrosis in patients with chronic hepatitis C vary based on:
   a. Duration of infection
   b. Fibrosis score
   c. Body habitus
   d. Operator/interpreter training or experience
   e. Co-existence of other etiologies of liver disease (e.g., non-alcoholic steatohepatitis)

3. What are the comparative diagnostic operating characteristics of tests of liver fibrosis?

4. What is the evidence for the timing of the initial testing for fibrosis and intervals for subsequent reassessment of fibrosis?

Search Strategy

A full search of the core sources was conducted to identify systematic reviews, meta-analyses, technology assessments, and clinical practice guidelines using terms for each of the studied interventions. Searches of core sources were limited to citations published after 2010.

The core sources searched included:

Agency for Healthcare Research and Quality (AHRQ)
Blue Cross/Blue Shield Health Technology Assessment (HTA) program
BMJ Clinical Evidence
Canadian Agency for Drugs and Technologies in Health (CADTH)
Cochrane Library (Wiley Interscience)
Hayes, Inc.
Institute for Clinical and Economic Review (ICER)
Medicaid Evidence-based Decisions Project (MED)
National Institute for Health and Care Excellence (NICE)
Tufts Cost-effectiveness Analysis Registry
Veterans Administration Evidence-based Synthesis Program (ESP)
Washington State Health Technology Assessment Program

A MEDLINE search was then conducted to identify randomized control trials, systematic reviews, meta-analyses, and technology assessments published after the end search date of the most recent SR for each studied intervention.

Searches for clinical practice guidelines were limited to those published since 2010. A search for relevant clinical practice guidelines was also conducted, using the following sources:

Australian Government National Health and Medical Research Council (NHMRC)
Centers for Disease Control and Prevention (CDC) – Community Preventive Services
Choosing Wisely
Institute for Clinical Systems Improvement (ICSI)
National Guidelines Clearinghouse
New Zealand Guidelines Group
NICE
Scottish Intercollegiate Guidelines Network (SIGN)
United States Preventive Services Task Force (USPSTF)
Veterans Administration/Department of Defense (VA/DOD)

Inclusion/Exclusion Criteria

Studies were excluded if they were not published in English or did not address the scope statement.

APPENDIX D: TEST CHARACTERISTICS

Noninvasive Tests with Good or Excellent Accuracy by Pooled or Median AUROC

Test	Pooled/Median AUROC ≥F2 (95% CI/Range)	Pooled/Median AUROC ≥F3 (95% CI/Range)
MRE	0.88 (0.84 - 0.91)	0.93 (0.90 - 0.95)
TE	0.89 (0.86 - 0.91)	0.92 (0.89 - 0.94)
ARFI	0.88 (0.81 - 0.96)	0.93 (0.89 - 0.97)
SWE	0.88 (0.85 - 0.91)	0.94 (0.92 - 0.96)
RT-TE		0.86 (NR)
ELF™	0.81 (median) (Range 0.72 - 0.87)	
Fibrometer™	0.82 (median) (Range 0.78 - 0.85)	
FIBROSpect® II	0.86 (median) (Range 0.77 - 0.90)	

Noninvasive Tests with Fair or Poor Accuracy by Median AUROC

Test	Median AUROC ≥F2 (Range)
Platelet count	0.71 (0.38 - 0.94)
Hyaluronic acid	0.75 (0.65 - 0.88)
Age-platelet index	0.74 (0.64 - 0.79)
APRI	0.77 (0.58 - 0.95)
AST-ALT ratio	0.59 (0.50 - 0.82)
Bonacini index	0.66 (0.58 - 0.71)
FIB-4	0.74 (0.61 - 0.81)
FibroIndex	0.76 (0.58 - 0.86)
FibroTest®	0.79 (0.70 - 0.89)
Forns index	0.76 (0.60 - 0.86)
Hepascore®	0.79 (0.69 - 0.82)
Pohl index	0.52 (0.52 - 0.53)

Illustrative Effects of Reported Cut-Offs on Sensitivity and Specificity

MRE (Singh et al., 2015)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	3.66 kPa	0.79	0.81
≥F3	4.11 kPa	0.85	0.85

TE (Steadman et al., 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	7.4 (SD ±1.5) kPa	0.80	0.81
≥F3	9.9 (SD ±2.4) kPa	0.84	0.87

ARFI (selected individual studies included in Nierhoff et al., 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	1.22 m/s	1.0	0.71
	1.37 m/s	0.69	0.92
	1.63 m/s	0.59	1.0
≥F3	1.71 m/s	1.0	0.73
	1.73 m/s	0.93	0.85

SWE (selected individual studies included in Li et al., 2016)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	7.2 kPa	0.86	0.86
	8.6 kPa	0.78	0.93
≥F3	9.1 kPa	0.92	0.85
	10.46 kPa	0.89	0.80

APRI (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	≥0.5 to >0.55	0.81	0.55
	≥1.5	0.37	0.95
F4	≥1.0	0.77	0.75
	≥2.0	0.48	0.94

ELF™ (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	>8.75	0.86	0.62
	>9.78	0.84	0.80

FIB-4 (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	≥1.45	0.64	0.68
	≥3.25	0.5	0.79
F4	≥1.45	0.90	0.58
	≥3.25	0.55	0.92

Fibrometer™ (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	>0.419 to >0.59	0.69	0.81

FIBROSpect® II (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	>0.36	0.95	0.66
	≥0.42	0.67	0.74

FibroTest® (Chou & Wasson, 2013)

Fibrosis Stage	Cut-off	Sensitivity	Specificity
≥F2	>0.10 to >0.22	0.92	0.38
	>0.70 to >0.80	0.22	0.96
F4	>0.56	0.85	0.77
	>0.73 to >0.862	0.56	0.81

APPENDIX E. APPLICABLE CODES

CODES	DESCRIPTION
ICD-10 Diagnosis Codes	
B18.2	Chronic viral hepatitis C
CPT Codes	
0346T	Ultrasound elastography (with diagnosis code)
91200	Liver elastography, mechanically induced shear wave (e.g. vibration), without imaging, with interpretation and report
91299	Other diagnostic gastroenterology procedures
0001M	Infectious disease, HCV, six biochemical assays (ALT, A2-macroglobulin, apolipoprotein A-1, total bilirubin, GGT, and haptoglobin) utilizing serum, prognostic algorithm reported as scores of fibrosis and necroinflammatory activity in liver
81599	Unlisted multianalyte assay with algorithm
82172	Apolipoprotein
82246	Bilirubin
82977	Glutamiltransferase, gamma (GGT)
83010	Hepatoglobin; quantitative
83519	Immunoassay, analyte quantitative by radiopharmaceutical technique
83520	Immunoassay NOS
83883	Nephelometry, each analyte not elsewhere specified
84450	Transferase; aspartate amino (AST) (SGOT)
84460	Transferase; alanine amino (ALT) (SGPT)

Note: Inclusion on this list does not guarantee coverage