Canadian Agency for Drugs and Technologies in Health

RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

**TITLE: Acoustic Radiation Force Impulse Imaging for Diagnosis and Monitoring of Liver Fibrosis in Patients with Hepatitis C: A Review of Diagnostic Accuracy, Clinical Effectiveness, Cost-Effectiveness, and Guidelines**

**DATE:** 18 April 2016

## CONTEXT AND POLICY ISSUES

Hepatitis C is a disease that affects the liver. The hepatitis C virus (HCV) is most commonly spread through contact with contaminated blood or blood products, or contaminated needles.[1] HCV causes inflammation of the liver, which can lead to liver damage. The early stages of inflammation, known as fibrosis, causes scarring on the liver and can affect how it functions.[1] When liver scarring worsens, it can become cirrhosis, and can increase the chances of developing liver cancer.[1] Hepatitis C is reported to infect 130 million to 150 million people globally. In Canada, the reported rate in 2013 was 29.55 cases per 100,000 people.[1] It is estimated that 250,000 Canadians have hepatitis C and many of them are unaware they are infected.[2]

Hepatitis C is a treatable disease, and there are drug therapies available for use in Canada.[2] Treating hepatitis C can prevent further liver damage, improve health outcomes, and cure the disease.[2] Testing for liver fibrosis in patients with hepatitis C can be informative for defining the treatment time and guiding appropriate treatment.[3] Treatment tends to be less successful in patients who have more advanced fibrosis.[4]

The current accepted method of testing for liver fibrosis is liver biopsy. Liver biopsy is an invasive method which evaluates 1/50000 of the total volume of liver.[5] Moreover, the specimen obtained from the liver biopsy must meet certain quality criteria, which is not always possible in a clinical setting.[5] Liver biopsies can also be quite painful and have potential complications including risk of death.[6] These issues have led to the development of non-invasive methods for the diagnosis of liver fibrosis.

Acoustic Radiation Force Imaging (ARFI) is an emerging non-invasive procedure that is a potential alternative to liver biopsy.[7] One significant advantage of ARFI imaging is that it is integrated into a conventional ultrasonographic system.[8] This also allows for a sonographic evaluation of the liver to be performed simultaneously with ARFI; this provides patients an ideal

'one-stop shop' for noninvasive liver evaluation, even in patients with a significant amount of ascites.[9] ARFI is a technology designed to measure shear wavefront at multiple locations to calculate tissue stiffness.[7] The wave velocity determines tissue stiffness through a simple method: the stiffer the tissue, the greater velocity.[7]

The purpose of this Rapid Response report is to review the clinical effectiveness, diagnostic accuracy, cost-effectiveness, and evidence-based guidelines regarding the use of ARFI for detecting and grading liver fibrosis in patients with hepatitis C. This report represents an update to a previous 2012 report that reviewed the clinical and cost-effectiveness evidence of diagnosis and monitoring of liver fibrosis among patients with hepatitis C; that report found moderate to high accuracy for FibroTest, transient elastography (known as FibroScan), and aminotransferase-to-platelet ratio index (APRI) with generally higher accuracy for cirrhosis compared with earlier fibrosis stages.[10]

## RESEARCH QUESTIONS

1. What is the clinical effectiveness of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C?

2. What is the diagnostic accuracy of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C?

3. What is the cost-effectiveness of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C?

4. What are the evidence-based guidelines regarding use of acoustic radiation force impulse imaging for detecting and grading liver fibrosis in patients with hepatitis C?

## KEY FINDINGS

Seventeen studies were identified regarding the diagnostic accuracy of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C. The studies identified in this report suggest that acoustic radiation force impulse imaging (ARFI) is a comparable method to liver biopsy to evaluate liver fibrosis and cirrhosis in patients with hepatitis C. One study was identified on the cost-effectiveness of ARFI for detecting and grading liver fibrosis in patients with hepatitis C, where liver biopsy and ARFI were found to be dominated by less costly and more effective options for chronic hepatitis C patients. However, the economic model did not include costs for more recent hepatitis C treatment options and may not reflect current practice. No literature on guidelines and clinical effectiveness were identified.

## METHODS

### Literature Search Methods

A limited literature search was conducted on key resources including PubMed, The Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to

the human population. The search was also limited to English language documents published between January 1, 2012 and March 17, 2016.

Rapid Response reports are organized so that the evidence for each research question is presented separately.

### Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

**Table 1:** Selection Criteria

| | |
|---|---|
| **Population** | Children, adolescents, or adults with hepatitis C |
| **Intervention** | Acoustic radiation force impulse imaging (ARFI), also known as point shear wave elastography (PSWE) and shear wave elastoplasty (SWE); or studies examining both ARFI and Transient Elastography (FibroScan) |
| **Comparator** | Liver biopsy |
| **Outcomes** | Comparative clinical effectiveness (e.g. clinical benefit, harms), diagnostic accuracy (e.g. sensitivity, specificity), cost-effectiveness, and guidelines |
| **Study Designs** | Health technology assessment, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies, evidence-based guidelines |

### Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2012. Articles were also excluded if they examined mixed populations without reporting hepatitis C findings separately from other etiologies.

### Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised using the AMSTAR checklist,[11] the non-randomized studies were critically appraised using the QUADAS-2 tool,[12] and the economic evaluation was critically appraised using the Drummond checklist.[13] Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations of each included study were described narratively.

## SUMMARY OF EVIDENCE

### Quantity of Research Available

A total of 224 citations were identified in the literature search. Following screening of titles and abstracts, 193 citations were excluded and 32 potentially relevant reports from the electronic search were retrieved for full-text review. Two potentially relevant publications were retrieved from the grey literature search. Of these potentially relevant articles, 16 publications were excluded for various reasons, while 18 publications met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection.

Additional references of potential interest are provided in Appendix 5.

**Summary of Study Characteristics**

Details of the study characteristics are located in Appendix 2.

*Study Design*

Four systematic reviews (SRs)[9,14-16] regarding the diagnostic accuracy of acoustic radiation force impulse imaging (ARFI) for patients with Hepatitis C were identified. Each of the SRs combined the results of the individual studies in a meta-analysis. One SR[14] examined fourteen systematic reviews, but the findings from only one of these reviews met the inclusion criteria and is reported in this report because it analyzed the diagnostic accuracy of ARFI compared to liver biopsy in patients with hepatitis C in isolation from other etiologies. Another SR[15] reviewed thirty-six non-randomized studies; four original articles and six abstracts examined the diagnostic accuracy of ARFI compared to liver biopsy in patients with hepatitis C in isolation from other etiologies and are reported in this report. The third SR[9] analyzed thirteen non-randomized studies or abstracts, with three of these reviews being reported in this report because they analyzed the diagnostic accuracy of ARFI/TE compared to liver biopsy in patients with hepatitis C in isolation from other etiologies. The final SR[16] examined eight non-randomized studies. All of the studies in this SR evaluated the overall performance of ARFI for the diagnosis of liver fibrosis.

Thirteen non-randomized studies (NRSs),[3,5,8,17-26] designed as cross-sectional, were also identified regarding the diagnostic accuracy of ARFI for patients with Hepatitis C.

One HTA[4] included a systematic review and meta-analysis and economic evaluation. The clinical findings are captured in the systematic review of reviews.[14] The economic portion represented a cost-utility analysis conducted from a United Kingdom Ministry of Health perspective. A lifetime horizon was used. A decision tree model was constructed which incorporated data from multiple sources including the meta-analysis. Long-term costs and health outcomes were estimated from a series of Markov models.

*Country of Origin*

One SR[14] was conducted in Canada. Two SRs[15,16] and one NRS[18] were conducted in Germany. One SR[9] and three NRSs[5,25,26] were conducted in Romania. One NRS[17] was conducted in Italy. Two NRSs[3,8] were conducted in Brazil. One NRS[19] was conducted in Korea. One NRS[20] was conducted in Taiwan. Three NRSs[21-23] were conducted in Japan. One NRS[24] was conducted in China. The economic evaluation[4] reflected a United Kingdom Ministry of Health perspective.

*Patient Population*

Four SRs[9,14-16] examined all causes of liver disease, which included hepatitis C virus, hepatitis B virus, nonalcoholic fatty liver disease, alcoholic liver disease, or cholestatic diseases. Ten of the NRSs[5,8,17-19,21-25] examined hepatitis C patients. Two of the NRSs[20,26] examined hepatitis C and B patients. One of the NRS[3] examined hepatitis C and nonalcoholic fatty liver disease patients.

The economic evaluation[4] generated separate estimates of costs and effects for populations with chronic hepatitis C. The model cohort was intended to mirror the distribution of chronic HCV genotype 1, 2 and 3, and 4 infections.

*Interventions and Comparators*

All of the studies[3,5,8,9,14-26] used liver biopsy as the reference standard and ARFI as the index test. Cut-off values for the index test ranged from 1.035 m/s to 2.11 m/s. Some of the studies[9,18,26] also examined transient elastography (TE) as an index test.

The economic evaluation[4] investigated several non-invasive tests identified in their systematic review, including ARFI and TE, for patients with chronic hepatitis C. These included direct and indirect non-invasive methods as well as imaging methods. Some of these methods included simple or indirect serum marks, direct serum markers, and imaging modalities.[4] In the primary analysis, tests were compared individually, while secondary analyses explored combination testing with different modalities. The analytic model was based on interferon-based treatments and did not include recently approved, interferon-free regimens.

*Outcomes*

All of the studies[3,5,8,9,14-26] examined the diagnostic accuracy of ARFI as a primary outcome. One study[21] also examined factors (such as BMI and hyaluronic acid) other than fibrosis stage correlating with ARFI in CHC. One study[22] developed a new index for assessment of liver fibrosis.

In the economic evaluation by Crossan et al.,[4] the mean costs were estimated, as well as the effects for the various non-invasive methods and liver biopsy in terms of the pound sterling (£) in 2012 and quality adjusted life years (QALYS). The economic evaluation assumed that patients could not regress to an earlier health state, that a METAVIR score <2 indicated a negative test outcome, that there was not difference in SVR rates based on mild, moderate, or cirrhotic health states, and that there was no disutility associated with non-invasive testing (a disutility was applied to biopsy testing due to the potential morbidity, mortality, and patient discomfort).

**Summary of Critical Appraisal**

Details of the critical appraisal are located in Appendix 3.

Most of the SRs[9,14,15] contained studies that were moderate to high quality. Two of the SRs[14,15] used an "a priori" design. Three of the SRs[9,14,15] conducted a comprehensive literature search, including multiple databases. Study selection was done in duplicate in two of the reviews.[9,15] One SR[14] did study selection by a single reviewer and it was unclear if the other SR[16] performed study selection in duplicate. None of the SRs[9,14-16] provided a list of the excluded studies. Three SRs[9,14,15] used validated tools to assess the quality of evidence. One of the SRs[16] did not provide any evidence of assessing the quality of the included studies. All of the reviews used appropriate methods to combine the findings of the studies. Three of the SRs[9,15,16] declared no conflicts of interest; the fourth SR[14] did not include a declaration of conflicts of interest.

All of the NRSs[3,5,8,17-26] avoided inappropriate exclusions and a case-control design. Seven of the NRSs[8,18-22,24] selected their patients consecutively; it was unclear how the remaining NRSs[3,5,17,23,25,26] selected their patients. Nine of the NRSs[8,18-22,24-26] blinded the results of the

index test when interpreting the results of the reference standard and vice versa. Two of the NRSs[5,17] blinded the results of the index test when interpreting the reference standard, but it was unclear if the index test results were interpreted without knowledge of the results of the reference standard. One NRS[23] interpreted the index test results without knowledge of the reference standard, however, it was unclear if the reference standard was interpreted without the results of the index test. It was unclear if one of the NRSs[3] blinded the results of the index test and/or the reference standard. Nine of NRSs[5,17-22,24,26] performed the index test and the reference standard within the same session. All of the NRSs[3,5,8,17-26] used the same reference standard, which was likely to classify patients appropriately. Some of the NRSs[18,24-26] excluded patients from the final analysis for one of two reasons: either they could not retrieve a good sample from the patient or because the patient met pre-determined exclusion criteria upon further review. A threshold was pre-specified in the index test in four of the NRSs[5,17,18,26]. In the other nine NRSs,[3,8,19-25] it was unclear if a threshold was pre-specified.

The economic evaluation[4] had several strengths. Both the perspective of the analysis and the alternatives being compared were explicit. The decision analytic model comparing the various non-invasive tests relative to liver biopsy for patients with chronic hepatitis C was explained in detail. The sources of evidence and methods of synthesis used to inform the model parameters were explicit. The time horizon for the model and the associated discount rates for the costs and effects were stated. Incremental analysis was reported and the conclusions were justified based on the data. Limitations of the analysis included the failure to explicitly state that the analysis being conducted was a cost-utility analysis. In addition, quantities of resource use were not reported separately from their unit costs and the distributions chosen from the probabilities sensitive analysis were not justified. Moreover, the treatments assessed in the model don't reflect the newest treatments available to hepatitis C, so costs are not likely reflective of current practice.

**Summary of Findings**

Details of the study findings are located in Appendix 4.

*What is the clinical effectiveness of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C?*

One SR[14] examined the diagnostic accuracy of ARFI and transient elastography (TE) compared to liver biopsy. The study concluded that ARFI was similar to TE for assessing liver fibrosis and that TE was an accurate method to detect moderate fibrosis or cirrhosis. Three SRs[9,15,16] all suggested that ARFI was a good method for assessing liver fibrosis and cirrhosis. One SR[15] reported that ARFI had a higher diagnostic accuracy for patients infected with HCV than than those with other liver diseases.

One NRS[17] concluded that ARFI measurement in patients with HCV was an easy and accurate non-invasive tool to identify patients with a benign course of HCV recurrence. Another NRS[5] reported that ARFI had a very good positive predictive value (93.2%) for predicting significant fibrosis and an excellent negative predictive value (97.8%) for excluding the presence of liver cirrhosis. One NRS[18] suggested that the diagnostic accuracy of ARFI was comparable to TE in testing liver fibrosis. One NRS[3] concluded that ARFI demonstrated good accuracy (area under the receiver operator curve [AUROC] 82.4%) in evaluating liver fibrosis and can replace liver biopsy in most cases; it recommends the use of liver biopsy to determine the degree of liver fibrosis when ARFI values are between 1.09 m/s and 1.70 m/s. Another NRS[19] reported that

ARFI demonstrated an acceptable diagnostic performance (sensitivity 79.5% to 85.0%, specificity 69.1% to 85.7% depending on fibrosis stage). One NRS[20] stated that ARFI was the modality of choice for predicting liver cirrhosis. One NRS[21] found ARFI correlated significantly with fibrosis stage, meaning that ARFI was able to distinguish between different stages of fibrosis. Another NRS[8] concluded that ARFI had very good accuracy for the assessment of liver fibrosis and cirrhosis (sensitivity 88.4% to 100%, specificity 75% to 95.2% depending on fibrosis stage). One NRS[22] stated that ARFI had excellent diagnostic accuracy when examining liver fibrosis (sensitivity 75% to 100%, specificity 40% to 90.9%). It also concluded that the VIA index (a combination of ARFI, alanine aminotransferase, and international normalized ratio) was potentially more useful for assessment of liver fibrosis than ARFI alone, as it easily and accurately measured liver fibrosis stage.[22] Another NRS[23] concluded that ARFI offers equivalent or higher diagnostic accuracy for liver fibrosis compared to TE. One NRS[24] found that ARFI may provide a promising alternative for evaluating liver fibrosis (sensitivity 74.1% to 88.9%, specificity 79.8% to 87%). Another NRS[25] reported that ARFI is a reliable method for predicting fibrosis severity (sensitivity 69.1% to 84.3%, specificity 76.3% to 81.5%). The final NRS[26] concluded that ARFI had similar predictive value with TE in patients with CHC and that ARFI might represent an alternative to TE.

*What is the cost-effectiveness of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C?*

The estimated costs of ARFI in patients with hepatitis C were £47,126 and the estimated QALYs were 14.25. For liver biopsy, the estimated costs were £47,900 and the estimated QALYs were 14.03. For FibroScan (TE), the estimated costs were £47,449 and the estimated QALYs were 14.28. In contrast the estimated costs and QALYs were £48,710 and 14.03 for liver biopsy, £54,878 and 12.45 for no treatment, and £51,241 and 14.73 for treating all patients. Liver biopsy, ARFI, and no treatment were dominated by less costly and more effective options when individual tests were compared, while transient elastography was extendedly dominated (incremental cost-effectiveness ratio [ICER] higher than the next most effective strategy). In this analysis, treating all patients irrespective of fibrosis stage was a cost-effective option (ICER £9,351 compared with magnetic resonance elastography which was the most cost-effective strategy when individual tests were compared).

**Limitations**

Three of the SRs[9,15,27] contained studies of moderate to high quality. One of the SRs[16] did not provide any evidence of quality appraisal. This affects the reliability of the study findings, as it is not known if the included studies are of low or high quality.

There were inconsistencies among all the studies in the various fibrosis stages when measuring ARFI and/or TE. Three of the SRs[14-16] used consistent fibrosis staging (F≥2, F≥3, F=4), whereas one of the of SRs[9] categorized fibrosis stages differently, depending on the study. There were also inconsistencies in all of the NRSs[3,5,8,17-26] on how they reported fibrosis. Four studies[8,21,23,25] used the stages ≥F1, ≥F2, ≥F3, F4 while others used other variations. These inconsistencies between studies may affect the way findings are interpreted.

Many of the NRSs[3,8,17-26] did not provide any indication that a pre-determined threshold was used to determine specific fibrosis stage. This represents a potential limitation in the interpretation of their results. This is also reflected in the inconsistencies in reporting the sensitivity and specificity in some of the studies. For example, one NRS[22] reported that ARFI

had excellent diagnostic accuracy when examining liver fibrosis even though the specificity for one of the fibrosis stages was reported as 40%.

Some of the NRSs[3,8,17] had sample sizes of fewer than 100 patients, including one[3] that included a total of 17 hepatitis C patients. These small sample sizes may impact the external validity of the findings, meaning they may not be a representative sample of the general population.

Most of the studies identified in this report[3,5,8,9,15-26] were developed and executed outside of Canada. Because these studies are not in a Canadian context, it may reduce the applicability of the findings in the studies.

## CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING

This report represents an update to a previous 2012 report that reviewed the clinical and cost-effectiveness evidence of diagnosis and monitoring of liver fibrosis among patients with hepatitis C; that report found moderate to high accuracy for FibroTest, transient elastography (known as FibroScan), and aminotransferase-to-platelet ratio index (APRI) with generally higher accuracy for cirrhosis compared with earlier fibrosis stages.[10] This report includes: four SRs[9,14-16] and thirteen NRSs[3,5,8,17-26] on the diagnostic accuracy of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis C. One economic evaluation[4] was identified regarding the cost-effectiveness of ARFI compared with liver biopsy. No literature was identified regarding evidence-based guidelines and clinical effectiveness on ARFI compared with liver biopsy in patients with Hepatitis C.

The studies identified in this report[3,5,8,9,14-26] were favourable to the use of ARFI in patients with hepatitis C. In four studies[14,18,23,26] that included both ARFI and TE, ARFI offered comparable or, in one study,[23] better diagnostic accuracy than transient elastography. ARFI has similar diagnostic accuracy to TE, which is also supported by studies[28,29] that have shown a high level of agreement between the two modalities. Ten studies[3,5,8,9,15,16,19,21,22,25] found that ARFI demonstrated good accuracy in evaluating liver fibrosis and/or liver cirrhosis. One study[20] concluded that ARFI was the modality of choice for predicting liver cirrhosis. One economic evaluation compared the cost-effectiveness of different options for the detection of liver fibrosis and treatment of hepatitis C. In this analysis, ARFI was dominated by less costly and more effective options among chronic hepatitis C patients. When tests were compared individually, treating all patients irrespective of fibrosis stage was a cost-effective option. However, the economic model did not include costs for more recent interferon-free hepatitis C treatment options and may not reflect current practice.

**PREPARED BY:**
Canadian Agency for Drugs and Technologies in Health
Tel: 1-866-898-8439
www.cadth.ca

## REFERENCES


1. Infectious diseases - Public Health Agency of Canada [Internet]. Ottawa: Public Health Agency of Canada. Hepatitis; 2015 Jul 16 [cited 2016 Apr 6]. Available from: http://www.phac-aspc.gc.ca/hep/index-eng.php

2. Canadian Liver Foundation: Hepatitis C [Internet]. Markham (ON): Canadian Liver Foundation. Can hepatitis C be treated? 2015 [cited 2016 Apr 6]. Available from: http://www.liver.ca/liver-disease/types/viral_hepatitis/Hepatitis_C.aspx#treatment

3. Guerra JA, Trippia M, Pissaia A, Jr., Teixeira BC, Ivantes CA. Acoustic radiation force impulse is equivalent to liver biopsy to evaluate liver fibrosis in patients with chronic hepatitis c and nonalcoholic fatty liver disease. Arq Gastroenterol [Internet]. 2015 Jul [cited 2016 Mar 23];52(3):234-8. Available from: http://www.scielo.br/pdf/ag/v52n3/0004-2803-ag-52-03-234.pdf

4. Crossan C, Tsochatzis EA, Longworth L, Gurusamy K, Davidson B, Rodriguez-Peralvarez M, et al. Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in patients with chronic liver disease: systematic review and economic evaluation. Health Technol Assess [Internet]. 2015 Jan [cited 2016 Mar 23];19(9):1-vi. Available from: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0082143/

5. Bota S, Sporea I, Sirli R, Popescu A, Gradinaru-Tascau O. How useful are ARFI elastography cut-off values proposed by meta-analysis for predicting the significant fibrosis and compensated liver cirrhosis? Med Ultrason [Internet]. 2015 Jun [cited 2016 Mar 23];17(2):200-5. Available from: http://www.medultrason.ro/assets/Magazines/Medultrason-2015-vol17-no2/articole-cu-DOI/11-Bota.pdf

6. Kleiner DE. The liver biopsy in chronic hepatitis C: a view from the other side of the microscope. In: Medscape Multispecialty [Internet]. New York: Medscape; 2015 [cited 2016 Apr 6]. Available from: http://www.medscape.com/viewarticle/501268_2

7. Steadman R, Leggett L, Lorenzetti D, Noseworthy T, Rose S, Sutherland L, et al. A health technology assessment of transient elastography in liver disease [Internet]. Edmonton (AB): Alberta Health Technologies Decision Process; 2012 Jan. [cited 2016 Mar 23]. Available from: http://www.health.alberta.ca/documents/AHTDP-ETG-UofC-STE.pdf

8. Silva RG Jr., Schmillevitch J, Nascimento MF, Miranda ML, Brant PE, Schulz PO, et al. Acoustic radiation force impulse elastography and serum fibrosis markers in chronic hepatitis C. Scand J Gastroenterol. 2014 Aug;49(8):986-92.

9. Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, et al. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int. 2013 Sep;33(8):1138-47.

10. Canadian Agency for Drugs and Technologies in Health. Diagnosis and monitoring of liver fibrosis in patients with chronic hepatitis C: a review of the clinical evidence and cost effectiveness [Internet]. Ottawa: CADTH; 2012 Mar 8. [cited 2016 Apr 6]. (Rapid response report: summary with critical appraisal). Available from: https://www.cadth.ca/sites/default/files/pdf/htis/mar-2012/RC0327_Hepatitis%20C_003_Final.pdf

11. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol [Internet]. 2007 [cited 2016 Apr 18];7:10. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1810543

12. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36.

13. Higgins JPT, editors. Cochrane handbook for systematic reviews of interventions [Internet]. Version 5.0.2. Oxford (U.K.): The Cochrane Collaboration; 2009. Figure 15.5.a: Drummond checklist (Drummond 1996). [cited 2016 Apr 18]. Available from: http://handbook.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm

14. Brener S. Transient elastography for assessment of liver fibrosis and steatosis: an evidence-based analysis. Ont Health Technol Assess Ser [Internet]. 2015 [cited 2016 Mar 23];15(18):1-45. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4664937

15. Nierhoff J, Chávez Ortiz AA, Herrmann E, Zeuzem S, Friedrich-Rust M. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: a meta-analysis. Eur Radiol. 2013 Nov;23(11):3040-53.

16. Friedrich-Rust M, Nierhoff J, Lupsor M, Sporea I, Fierbinteanu-Braticevici C, Strobel D, et al. Performance of acoustic radiation force impulse imaging for the staging of liver fibrosis: a pooled meta-analysis. J Viral Hepat. 2012 Feb;19(2):e212-e219.

17. Bignulin S, Falleti E, Cmet S, Cappello D, Cussigh A, Lenisa I, et al. Usefulness of acoustic radiation force impulse and fibrotest in liver fibrosis assessment after liver transplant. Ann Hepatol. 2016 Mar;15(2):200-6.

18. Friedrich-Rust M, Lupsor M, de Knegt R, Dries V, Buggisch P, Gebel M, et al. Point shear wave elastography by acoustic radiation force impulse quantification in comparison to transient elastography for the noninvasive assessment of liver fibrosis in chronic hepatitis C: a prospective international multicenter study. Ultraschall Med. 2015 Jun;36(3):239-47.

19. Joo SK, Kim JH, Oh S, Kim BG, Lee KL, Kim HY, et al. Prospective comparison of noninvasive fibrosis assessment to predict advanced fibrosis or cirrhosis in Asian patients with hepatitis C. J Clin Gastroenterol. 2015 Sep;49(8):697-704.

20. Tai DI, Tsay PK, Jeng WJ, Weng CC, Huang SF, Huang CH, et al. Differences in liver fibrosis between patients with chronic hepatitis B and C: evaluation by acoustic radiation force impulse measurements at 2 locations. J Ultrasound Med. 2015 May;34(5):813-21.

21. Nishikawa T, Hashimoto S, Kawabe N, Harata M, Nitta Y, Murao M, et al. Factors correlating with acoustic radiation force impulse elastography in chronic hepatitis C. World J Gastroenterol [Internet]. 2014 Feb 7 [cited 2016 Mar 29];20(5):1289-97. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3921510

22. Takaki S, Kawakami Y, Miyaki D, Nakahara T, Naeshiro N, Murakami E, et al. Non-invasive liver fibrosis score calculated by combination of virtual touch tissue quantification and serum liver functional tests in chronic hepatitis C patients. Hepatol Res. 2014 Mar;44(3):280-7.

23. Yamada R, Hiramatsu N, Oze T, Morishita N, Harada N, Miyazaki M, et al. Significance of liver stiffness measurement by acoustic radiation force impulse (ARFI) among hepatitis C patients. J Med Virol. 2014 Feb;86(2):241-7.

24. Chen SH, Li YF, Lai HC, Kao JT, Peng CY, Chuang PH, et al. Effects of patient factors on noninvasive liver stiffness measurement using acoustic radiation force impulse elastography in patients with chronic hepatitis C. BMC Gastroenterol [Internet]. 2012 [cited 2016 Mar 29];12:105. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3462151

25. Sporea I, Bota S, Peck-Radosavljevic M, Sirli R, Tanaka H, Iijima H, et al. Acoustic Radiation Force Impulse elastography for fibrosis evaluation in patients with chronic hepatitis C: an international multicenter study. Eur J Radiol. 2012 Dec;81(12):4112-8.

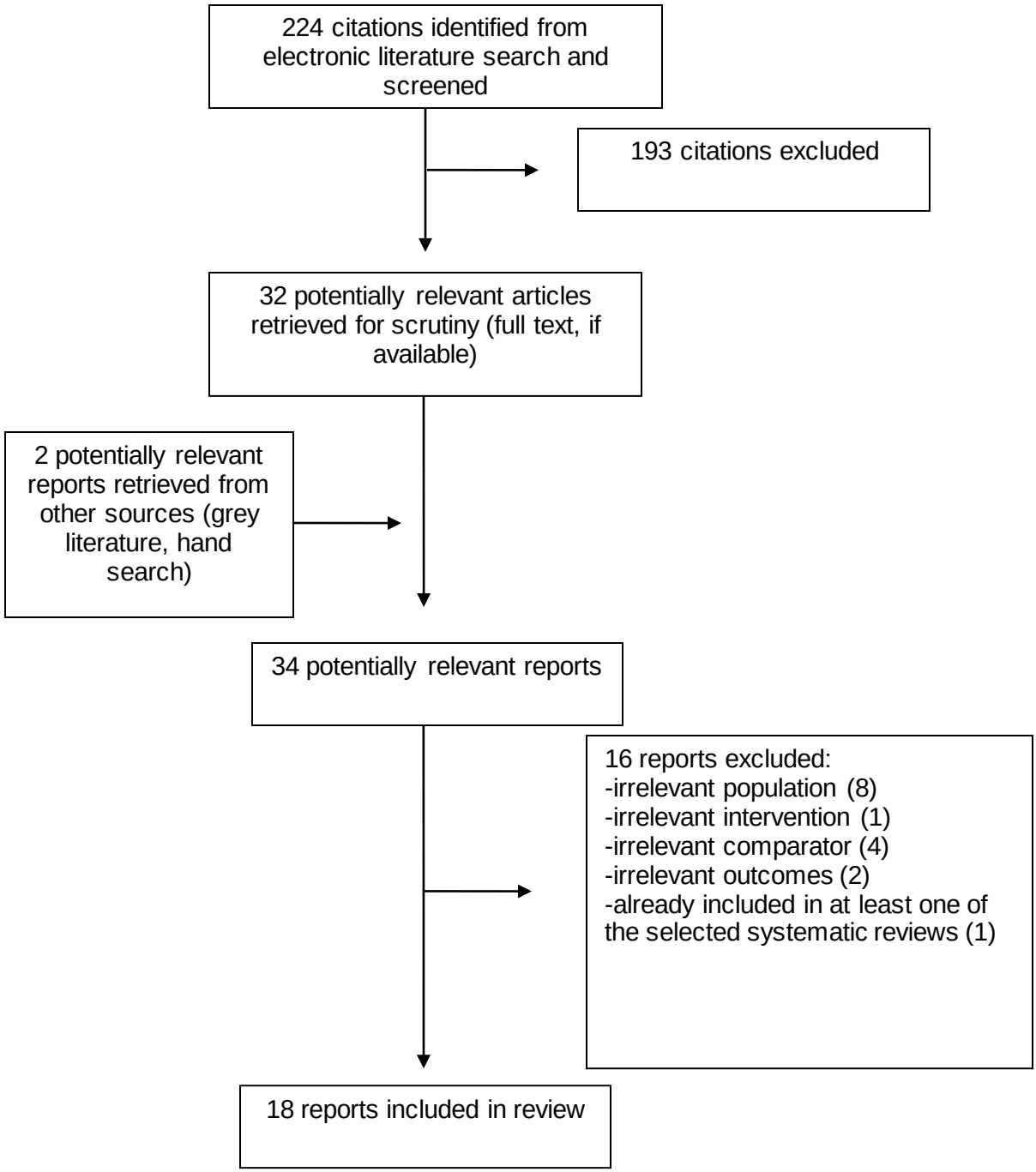
26. Sporea I, Sirli R, Bota S, Popescu A, Sendroiu M, Jurchis A. Comparative study concerning the value of acoustic radiation force impulse elastography (ARFI) in comparison with transient elastography (TE) for the assessment of liver fibrosis in patients with chronic hepatitis B and C. Ultrasound Med Biol. 2012 Aug;38(8):1310-6.

27. Trovato FM, Atzori S, Musumeci G, Tooley V, Marcinkowski H, Crossey MM, et al. Liver and spleen transient elastography and Acoustic Radiation Force Impulse Measurements. Performance and comparison of measurements in the same area concurrently assessed for liver fibrosis by biopsy. Adv Med Sci. 2015 Sep;60(2):300-6.

28. Kobayashi K, Nakao H, Nishiyama T, Lin Y, Kikuchi S, Kobayashi Y, et al. Diagnostic accuracy of real-time tissue elastography for the staging of liver fibrosis: a meta-analysis. Eur Radiol. 2015 Jan;25(1):230-8.

29. Frulio N, Trillaud H, Perez P, Asselineau J, Vandenhende M, Hessamfar M, et al. Acoustic Radiation Force Impulse (ARFI) and Transient Elastography (TE) for evaluation of liver fibrosis in HIV-HCV co-infected patients. BMC Infect Dis [Internet]. 2014 [cited 2016 Mar 23];14:405. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4223715

**APPENDIX 1: Selection of Included Studies**



## APPENDIX 2: Characteristics of Included Publications

**Table A1:** Characteristics of Included Systematic Reviews and Meta-Analyses

| First Author, Publication Year, Country | Types and numbers of primary studies included | Population Characteristics | Index Test | Reference Standard | Clinical Outcomes |
|---|---|---|---|---|---|
| Brener, 2015, Canada[14] | Fourteen SRs published between 2007 and 2014<br><br>Five of these SRs evaluated the overall performance of ARFI and TE for the diagnosis of LF | All causes of liver disease, n = NR<br><br>One of these reviews in this SR was reported on because it analyzed the DA of ARFI in patients with HCV | ARFI/TE | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Nierhoff, 2014, Germany[15] | Thirty-six studies published between 2007 to February 2012<br><br>All of these studies evaluated the overall performance of ARFI for the diagnosis of LF | All causes of liver disease, n = 3591<br><br>Four original articles and 6 abstracts in this SR were reported on because they analyzed the DA of ARFI in patients with HCV | ARFI | LB | Diagnostic accuracy of ARFI (measured by measured by AUROC, SE, SP, PPV, NPV) |
| Bota, 2013, Romania[9] | Thirteen studies (eleven full-length articles and 2 abstracts) published until May 31, 2012<br><br>All of these studies evaluated the overall performance of ARFI and TE for the diagnosis of LF | All causes of liver disease, n = 1163<br><br>Three of these reviews in this SR were reported on because they analyzed the DA of ARFI in patients with HCV | ARFI/TE | LB | Diagnostic accuracy of ARFI and TE (measured by measured by SROC, SE, SP, PPV, NPV) |

**Table A1:** Characteristics of Included Systematic Reviews and Meta-Analyses

| First Author, Publication Year, Country | Types and numbers of primary studies included | Population Characteristics | Index Test | Reference Standard | Clinical Outcomes |
|---|---|---|---|---|---|
| Friedrich-Rust, 2012, Germany[16] | Eight studies published until October 2010<br><br>All of these studies evaluated the overall performance of ARFI for the diagnosis of LF | All causes of liver disease, n = 518, HCV patients, n = 380 | ARFI | LB | Diagnostic accuracy of ARFI (measured by measured by SROC, SE, SP, PPV, NPV) |

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; DA = Diagnostic Accuracy; LB = Liver biopsy; LF = Liver fibrosis; NPV = Negative predictive value; PPV = Positive predictive value; SE = Sensitivity; SP = Specificity; SR = Systematic Review ; SROC = Summary receiver operating characteristic; TE = Transient elastography

**Table A2**: Characteristics of Included Clinical Studies

| First Author, Publication Year, Country | Study Design | Patient Characteristics | Intervention(s) | Comparator(s) | Clinical Outcomes |
|---|---|---|---|---|---|
| Bignulin, 2016, Italy [17] | Cross-sectional | 51 CHC patients | ARFI | LB | Diagnostic accuracy (measured by ROC curves, SE, SP, Youden index, PPV and NPV) |
| Bota, 2015, Romania[5] | Cross-sectional | 132 CHC patients | ARFI | LB | Diagnostic accuracy (measured by Kolmogrov-Smirnov test, parametric and/or non-parametric tests, PPV and NPV) |
| Friedrich-Rust, 2015 [18] | Cross-sectional | 241 CHC patients | ARFI/TE | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |

**Table A2**: Characteristics of Included Clinical Studies

| First Author, Publication Year, Country | Study Design | Patient Characteristics | Intervention(s) | Comparator(s) | Clinical Outcomes |
|---|---|---|---|---|---|
| Guerra, 2015, Brazil [3] | Cross-sectional | 24 CHC or NAFLD patients | ARFI | LB | Diagnostic accuracy (measured by ROC curves, SE, SP) |
| Joo, 2015, Korea [19] | Cross-sectional | 101 CHC patients who were antiviral treatment-naïve | ARFI | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Tai, 2015, Taiwan [20] | Cross-sectional | 204 CHC and CHB patients | ARFI | LB | Diagnostic accuracy (measured by ROC curves, AUROC, SE, SP) |
| Nishikawa, 2014, Japan [21] | Cross-sectional | 108 CHC patients | ARFI | LB | Factors correlating with ARFI in CHC, Diagnostic accuracy (measured by AUROC,SE, SP, PPV, NPV) |
| Silva Junior, 2014, Brazil [8] | Cross-sectional | 51 untreated CHC patients | ARFI | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Takaki, 2014, Japan [22] | Cross-sectional | 176 CHC patients | ARFI | LB | Develop a new index for assessment of liver fibrosis, diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Yamada, 2014, Japan [23] | Cross-sectional | 124 patients with CHC | ARFI | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |

**Table A2**: Characteristics of Included Clinical Studies

| First Author, Publication Year, Country | Study Design | Patient Characteristics | Intervention(s) | Comparator(s) | Clinical Outcomes |
|---|---|---|---|---|---|
| Chen, 2012, China [24] | Cross-sectional | 127 patients with CHC | ARFI | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Sporea, 2012, Romania [25] | Cross-sectional | 914 patients (10 centers, 5 countries) with CHC | ARFI | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |
| Sporea, 2012, Romania [26] | Cross-sectional | 160 patients with CHB and CHC | ARFI/TE | LB | Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV) |

ARFI = Acoustic radiation force impulse imaging ; AUROC = Areas under the respective receiver operator characteristics curves; CHB = Chronic hepatitis B; CHC = Chronic hepatitis C; HCV = Hepatitis C; LB = Liver biopsy; LF = Liver fibrosis; LS = Liver stiffness; NAFLD = Nonalcoholic fatty liver disease; NPV = Negative predictive value; PPV = Positive predictive value; PSWE = point shear wave elastography; ROC = Receiver operating characteristic; SE = Sensitivity; SP = Specificity; TE = Transient elastography; VTTQ = Virtual touch tissue quanitification

**Table A3** Characteristics of Included Cost Studies

| First author, Publication Year, Country | Type of Analysis, Perspective | Intervention, Comparator | Study Population | Time Horizon | Main Assumptions |
|---|---|---|---|---|---|
| Crossan, 2015, United Kingdom[4] | Cost-utility analysis United Kingdom Ministry of Health perspective (NHS) | Each non-invasive liver test identified from the systematic review was included in the analysis<br><br>For chronic hepatitis C there were 57 interventions considered: 22 indirect methods, 22 direct methods and 13 imaging methods<br><br>Liver biopsy | Patients with chronic hepatitis B or C who are suspected of having liver fibrosis (i.e., patients who a hepatologist would wish to biopsy to inform treatment decisions) | Lifetime | Markov model cycle length of one year<br><br>Patients were assumed to enter the model at between 30 and 40 years of age<br><br>Cost data were derived from the literature based on resource use information collected on inpatient and outpatient care, investigations, |

**Table A3** Characteristics of Included Cost Studies

| First author, Publication Year, Country | Type of Analysis, Perspective | Intervention, Comparator | Study Population | Time Horizon | Main Assumptions |
|---|---|---|---|---|---|
| | | | | | procedures, drug use and other services<br><br>Costs were inflated to 2012 levels using NHS inflation indices<br><br>Utility data were derived from the literature based on the EQ-5D preference measure within a UK population<br><br>Death was assumed to have a utility value of 0<br><br>Health outcomes were measured in quality adjusted life years<br><br>Costs and utilities were discounted at a rate of 3.5% |

NHS = National Health Services

## APPENDIX 3: Critical Appraisal of Included Publications

**Table A4:** Strengths and Limitations of Systematic Reviews and Meta-Analyses using the AMSTAR checklist[11]

| Strengths | Limitations |
|---|---|
| Brener[14] | |
| • An "a priori" design was provided.<br>• A comprehensive literature search was performed, including grey literature. A detailed search strategy and a flow diagram for the search results were provided.<br>• A list of the included studies was provided.<br>• The characteristics of the included studies were provided.<br>• The scientific quality of the included studies was assessed and documented, and the included studies were rated on their quality, using the QUADAS-2 tool and AMSTAR checklist.<br>• The scientific quality of the included studies was used appropriately in formulating conclusions.<br>• The methods used to combine the findings of studies were appropriate. | • Study selection was done by a single reviewer; it is unclear if data extraction was done in duplicate.<br>• A list of excluded studies was not provided.<br>• The likelihood of publication bias was not assessed.<br>• No conflicts of interest were mentioned by the authors. |
| Nierhoff[15] | |
| • An "a priori" design was provided.<br>• A comprehensive literature search was performed. A detailed search strategy and a flow diagram for the search results were provided.<br>• Study selection was done by two independent reviewers.<br>• A list of the included studies was provided.<br>• The characteristics of the included studies were provided.<br>• The likelihood of publication bias was assessed.<br>• The scientific quality of the included studies was assessed and documented, and the included studies were rated on their quality, using the QUADAS-2 tool.<br>• The scientific quality of the included studies was used appropriately in formulating conclusions.<br>• The methods used to combine the findings of studies were appropriate. | • A list of excluded studies was not provided.<br>• No conflicts of interest were mentioned by the authors.<br>• Abstracts were included in the meta-analysis, however, the authors of the abstracts were contacted if the data in the abstract were insufficient. |
| Bota[9] | |
| • A comprehensive literature search was performed. A detailed search strategy and a flow diagram for the search results were provided.<br>• Study selection was done by two independent reviewers. | • It is unclear if an "a priori" design was provided.<br>• A list of excluded studies was not provided.<br>• Abstracts were included in the meta-analysis. |

**Table A4:** Strengths and Limitations of Systematic Reviews and Meta-Analyses using the AMSTAR checklist[11]

| Strengths | Limitations |
|---|---|
| • A list of the included studies was provided.<br>• The characteristics of the included studies were provided.<br>• The likelihood of publication bias was assessed.<br>• The scientific quality of the included studies was assessed and documented, and the included studies were rated on their quality, using the QUADAS-2 tool.<br>• The scientific quality of the included studies was used appropriately in formulating conclusions.<br>• The methods used to combine the findings of studies were appropriate.<br>• No conflicts of interest were declared by the authors.<br>• The authors of the studies were contacted if the data were insufficient. | |
| Friedrich-Rust[16] | |
| • A comprehensive literature search was performed.<br>• A list of the included studies was provided.<br>• The characteristics of the included studies were provided.<br>• The likelihood of publication bias was assessed.<br>• The methods used to combine the findings of studies were appropriate.<br>• No conflicts of interest were declared by the authors.<br>• The authors of the studies were contacted for original patient data. | • It is unclear if an "a priori" design was provided.<br>• A flow diagram for the search results was not provided.<br>• A list of excluded studies was not provided.<br>• The scientific quality of the included studies was not assessed and documented, and the included studies were not rated on their quality.<br>• The scientific quality of the included studies was not used appropriately in formulating conclusions.<br>• It is unclear if study selection was done by two independent reviewers. |

**Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12]

| Strengths | Limitations |
|---|---|
| Bignulin[17] | |
| *Patient Selection*<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>*Reference Standard*<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>• The reference standard was likely to classify patients appropriately | *Patient Selection*<br>• Unclear if patient selection was consecutive<br>*Index Test*<br>• It is unclear if the index test results were interpreted without knowledge of the results of the reference standard<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• No limitations identified |

**Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12]

| Strengths | Limitations |
|---|---|
| *Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis<br>• There was an appropriate time interval between the index test and reference standard | |
| Bota[5] | |
| *Patient Selection*<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• A threshold was pre-specified<br>*Reference Standard*<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>• The reference standard was likely to classify patients appropriately<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• Unclear if patient selection was consecutive<br>*Index Test*<br>• It is unclear if the index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• No limitations identified |
| Friedrich-Rust[18] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• No limitations identified<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• Not all patients were included in the analysis |
| Guerra[3] | |
| *Patient Selection*<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• No limitations identified<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately | *Patient Selection*<br>• Unclear if patient selection was consecutive<br>*Index Test*<br>• Unclear if the index test results were interpreted without knowledge of the results of the reference standard<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• Unclear if the reference standard results were |

**Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12]

| Strengths | Limitations |
|---|---|
| *Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis | interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• There was a significant time interval between the index test and reference standard |
| Joo[19] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• No limitations identified<br>*Index Test*<br>• It is unclear if a threshold was pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• No limitations identified |
| Tai[20] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• No limitations identified<br>*Index Test*<br>• It is unclear if a threshold was pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• No limitations identified |
| Nishikawa[21] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without | *Patient Selection*<br>• Unclear if the study avoided inappropriate exclusions<br>*Index Test*<br>• A threshold was not pre-specified |

**Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12]

| Strengths | Limitations |
|---|---|
| knowledge of the results of the reference standard<br>• A threshold was pre-specified<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis | *Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• The index test and reference standard were completed at different times |
| Silva Junior[8] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis | *Patient Selection*<br>• No limitations identified<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• There was a significant time interval between the index test and reference standard |
| Takaki, 2014[22] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• It is unclear if the study avoided inappropriate exclusions<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• No limitations identified |

**Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12]

| Strengths | Limitations |
|---|---|
| Yamada, 2014[23] | |
| *Patient Selection*<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>*Flow and Timing*<br>• All patients received the reference standard<br>• All patients were included in the analysis | *Patient Selection*<br>• It is unclear if the patient selection was consecutive<br>*Index Test*<br>• It is unclear if a threshold was pre-specified<br>*Reference Standard*<br>• It is unclear if the reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• The reference standard was completed within one week of the index test |
| Chen[24] | |
| *Patient Selection*<br>• Patient selection was consecutive<br>• The study avoided inappropriate exclusions<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• There was an appropriate time interval between the index test and reference standard | *Patient Selection*<br>• No limitations identified<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• Not all patients were included in the analysis<br>• Not all patients received the reference standard |
| Sporea, 2012[25] | |
| *Patient Selection*<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard | *Patient Selection*<br>• Unclear if patient selection was consecutive<br>• Exclusion criteria was unclear<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• There was a significant time interval between the index test and reference standard<br>• Not all patients were included in the analysis |

| **Table A5:** Strengths and Limitations of Non-Randomized Studies using QUADAS-2[12] | |
|---|---|
| **Strengths** | **Limitations** |
| Sporea, 2012[26] | |
| *Patient Selection*<br>• A case-control design was avoided<br>*Index Test*<br>• The index test results were interpreted without knowledge of the results of the reference standard<br>*Reference Standard*<br>• The reference standard was likely to classify patients appropriately<br>• The reference standard results were interpreted without knowledge of the results of the index test<br>*Flow and Timing*<br>• All patients received the reference standard<br>• There was an appropriate interval between the index test and the reference standard | *Patient Selection*<br>• Unclear if patient selection was consecutive<br>• Exclusion criteria was unclear<br>*Index Test*<br>• A threshold was not pre-specified<br>*Reference Standard*<br>• No limitations identified<br>*Flow and Timing*<br>• Not all patients were included in the analysis |

| **Table A6:** Strengths and Limitations of Economic Studies using Drummond[13] | |
|---|---|
| **Strengths** | **Limitations** |
| Crossan[4] | |
| • The perspective of the analysis was explicit and the alternatives being compared were clearly described<br>• Details of the sources of evidence and methods of synthesis used to inform model parameters were given<br>• The models were described<br>• The time horizon for costs and effects and the associated discount rates were stated<br>• Incremental analysis was reported<br>• The conclusions followed from the reported data | • The form of the economic evaluation was not explicitly stated<br>• Quantities of resource use were not reported separately from their unit costs<br>• The choice of distributions for probabilistic sensitivity analysis were not justified |

## APPENDIX 4: Main Study Findings and Author's Conclusions

**Table A7:** Summary of Findings of Included SRs and MAs

<table>
<tr><th>Main Study Findings</th><th>Author's Conclusions</th></tr>
<tr><td colspan="2">Brener, 2015[14]</td></tr>
<tr><td>
<ul>
<li>One systematic review met the inclusion criteria for this report (Tsochatzis et al. 2015)</li>
<li>ARFI had similar value to TE for significant fibrosis and cirrhosis</li>
<li>ARFI had good diagnostic accuracy for staging liver fibrosis</li>
<li>Noninvasive tests (ie, ARFI and TE) were good at excluding advanced cirrhosis</li>
<li>TE was an accurate diagnostic method for moderate fibrosis or cirrhosis</li>
<li>The DA of LS measurement using TE or ARFI for the assessment of LF according to METAVIR stages in studies that examine <u>HCV only</u>:</li>
</ul>
<table>
<tr><th>Stage</th><th>Studies (Pts), N</th><th>SROC</th><th>SE (95% CI)</th><th>SP (95% CI)</th></tr>
<tr><td>ARFI</td><td></td><td></td><td></td><td></td></tr>
<tr><td>≥F2</td><td>3 (NR)</td><td>NR</td><td>0.79 (0.75-0.83)</td><td>0.89 (0.84-0.93)</td></tr>
<tr><td>≥F3</td><td>4 (NR)</td><td>NR</td><td>0.85 (0.69-0.94)</td><td>0.89 (0.72-0.97)</td></tr>
<tr><td>F=4</td><td>4 (NR)</td><td>NR</td><td>0.84 (0.72-0.91)</td><td>0.77 (0.50-0.92)</td></tr>
<tr><td>TE</td><td></td><td></td><td></td><td></td></tr>
<tr><td>≥F2</td><td>37 (NR)</td><td>0.87 (0.83-0.90)</td><td>0.79 (0.74-0.84)</td><td>0.83 (0.77-0.88)</td></tr>
<tr><td>≥F3</td><td>19 (NR)</td><td>0.94 (0.92-0.96)</td><td>0.88 (0.82-0.92)</td><td>0.90 (0.85-0.93)</td></tr>
<tr><td>F=4</td><td>36 (NR)</td><td>0.96 (0.94-0.97)</td><td>0.89 (0.84-0.92)</td><td>0.91 (0.89-0.93)</td></tr>
</table>
<u>Quality of Included Studies</u><br>
The study analyzing ARFI/TE with biopsy as a diagnostic technology was of higher quality.
</td><td>
<ul>
<li>The diagnostic accuracy of ARFI was not significantly different from TE for assessing liver fibrosis</li>
<li>"There was evidence that the diagnostic accuracy of FibroTest and ARFI were not significantly different from TE for assessing LF in the disease areas of interest" (page 33)</li>
</ul>
</td></tr>
<tr><td colspan="2">Nierhoff, 2013[15]</td></tr>
<tr><td>
<ul>
<li>10 of 36 studies (four original articles and six abstracts) met the inclusion criteria for this review</li>
<li>The DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages in studies that examine <u>HCV only</u>:</li>
</ul>
</td><td>
<ul>
<li>"In conclusion, the present meta-analysis including 36 studies with 3,951 patients revealed good DA of ARFI imaging for the staging of significant and severe fibrosis, and excellent diagnostic accuracy for the diagnosis of liver cirrhosis" (page 3051)</li>
</ul>
</td></tr>
</table>

**Table A7:** Summary of Findings of Included SRs and MAs

| Main Study Findings | Author's Conclusions |
|---|---|

| Stage | AUROC | Cutoff | SE | SP | DOR |
|---|---|---|---|---|---|
| **Original Articles** | | | | | |
| Fierbinteanu-Braticevici et al. | | | | | |
| ≥F2 | 0.91 | 1.22 | 100 | 71 | 315.83 |
| ≥F3 | 0.99 | 1.54 | 97 | 100 | 2295.67 |
| F=4 | 0.99 | 1.94 | 100 | 98 | 2009 |
| Lupsor et al. | | | | | |
| ≥F2 | 0.86 | 1.34 | 68 | 93 | 28.23 |
| ≥F3 | 0.91 | 1.61 | 79 | 95 | 71.48 |
| F=4 | 0.94 | 2.11 | 80 | 95 | 76 |
| Kuroda et al. | | | | | |
| ≥F2 | NA | NA | NA | NA | NA |
| ≥F3 | NA | NA | NA | NA | NA |
| F=4 | 0.93 | 1.59 | 95 | 83 | 92.76 |
| Rizzo et al. | | | | | |
| ≥F2 | 0.86 | 1.31 | 81 | 70 | 9.95 |
| ≥F3 | 0.94 | 1.71 | 91 | 86 | 62.11 |
| F=4 | 0.89 | 2.11 | 83 | 86 | 29.99 |
| **Abstracts** | | | | | |
| Hsu et al. | | | | | |
| ≥F2 | NA | NA | NA | NA | NA |
| ≥F3 | 0.87 | 1.81 | 80 | 83 | NA |
| F=4 | 0.92 | 2.31 | 75 | 92 | NA |
| Rossini et al. | | | | | |
| ≥F2 | NA | NA | NA | NA | NA |
| ≥F3 | 0.81 | 2.33 | 90 | 76 | 28.5 |
| F=4 | NA | NA | NA | NA | NA |
| Song et al. | | | | | |
| ≥F2 | 0.89 | NA | NA | NA | NA |
| ≥F3 | 0.94 | NA | NA | NA | NA |
| F=4 | 0.94 | NA | NA | NA | NA |
| Fierbinteanu-Braticevici et al. | | | | | |
| ≥F2 | 0.97 | NA | NA | NA | NA |
| ≥F3 | 0.98 | NA | NA | NA | NA |
| F=4 | NA | NA | NA | NA | NA |
| Sporea et al. | | | | | |
| ≥F2 | 0.81 | 1.29 | 72 | 80 | 10.29 |
| ≥F3 | 0.84 | 1.57 | 70 | 90 | 21 |
| F=4 | 0.85 | 1.59 | 84 | 80 | 21 |
| Yoshioka et al. | | | | | |
| ≥F2 | 0.58 | 1.26 | NA | NA | NA |
| ≥F3 | 0.87 | 1.65 | NA | NA | NA |
| F=4 | 0.78 | 2.03 | NA | NA | NA |

Quality of Included Studies

Most of the studies analyzing ARFI with liver biopsy as a diagnostic technology were of higher quality.

- "…there is a slight trend towards higher diagnostic accuracy in studies with patients only infected with HCV than in studies with patients infected with different liver diseases" (page 3050)

Bota, 2013[9]

- The statistical analysis performed showed similar diagnostic accuracy for ARFI and TE for the diagnosis of significant fibrosis and cirrhosis
- Comparison of the DA with ARFI and TE in studies that examine patients with **HCV only**:

- "[ARFI] elastography seems to be a good method for assessing LF (0.85 SROC curve for detecting significant fibrosis and 0.93 for diagnosing cirrhosis), especially in CHC patients and shows similar predictive value for significant fibrosis and cirrhosis compared to TE" (page 1143)

**Table A7:** Summary of Findings of Included SRs and MAs

| Main Study Findings | Author's Conclusions |
|---|---|
| • <br>**Stage** / **Cut-off ARFI (m/s) vs. cut-off TE (kPa)** / **AUROC ARFI vs. AUROC TE**<br>Lupsor et al.<br>F≥1: 1.19 vs. 5.2 / 0.709 vs. 0.902 (p=0.006)<br>F≥2: 1.34 vs. 8.1 / 0.851 vs. 0.941 (p=0.02)<br>F≥3: 1.61 vs. 9.6 / 0.869 vs. 0.926 (p=0.15)<br>F=4: 2 vs. 13.1 / 0.911 vs. 0.945 (p=0.3)<br>Sporea et al.<br>F≥2: 1.2 vs. 6.7 / 0.84 vs. 0.87<br>F=4: 1.8 vs. 12.2 / 0.91 vs 0.97<br>Rizzo et al.<br>F≥2: 1.3 vs. 6.5 / 0.86 vs. 0.78 (p=0.02)<br>F≥3: 1.7 vs. 8.8 / 0.94 vs. 0.83 (p=0.002)<br>F=4: 2 vs. 11 / 0.89 vs. 0.80 (p=0.09)<br><br>Quality of Included Studies<br>Most of the studies examining ARFI and TE with liver biopsy as a diagnostic technology were high quality. One study had an unclear risk of bias. | |
| Friedrich-Rust[16] | |
| • The DA of ARFI in studies that examine patients with **HCV only**:<br>**Stage** / **AUROC**<br>F≥2: 0.88 (0.83, 0.92)<br>F≥3: 0.90 (0.84, 0.97)<br>F=4: 0.92 (0.87, 0.98) | • "The presents meta-analysis reveals a good DA of ARFI for the diagnosis of significant LF and excellent DA for the diagnosis of liver cirrhosis" (page e218) |

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; CHC = Chronic hepatitis C; CI = Confidence Interval; DA = Diagnostic accuracy; HCV = Hepatitis C; LF = Liver fibrosis; LS = Liver stiffness; MA = Meta-analysis; SE = Sensitivity; SP = Sensitivity; SR = Systematic review ; SROC = Summary receiver operating characteristic; TE = Transient elastography

**Table A8:** Summary of Findings of Included NRSs

| Main Study Findings | Author's Conclusions |
|---|---|
| Bignulin, 2016[17] | |
| • ARFI had an excellent sensitivity and NPV (100%) in discriminating patients with significant fibrosis (using a cut off for LS of 1.365 m/s)<br>• The ARFI results in differentiating between Ishak fibrosis score ≥ 2 / ≤ 3:<br>**AUROC** / **p** / **Cutoff** / **SE** / **SP** / **PPV** / **NPV**<br>0.885 / <0.001 / 1.365 / 100 / 73.7 / 56.5 / 100 | • "ARFI measurement in HCV positive liver transplanted patients can be considered an easy and accurate non-invasive tool in identifying patients with a benign course of HCV recurrence" (page 205) |

**Table A8:** Summary of Findings of Included NRSs

<table>
<tr><th>Main Study Findings</th><th>Author's Conclusions</th></tr>
<tr><td colspan="2">Bota, 2015[5]</td></tr>
<tr><td>
<ul>
<li>In patients with LS lower than 1.35 m/s, neither ARFI nor LB were useful in predicting the presence of fibrosis</li>
<li>ARFI can be a useful method to select individuals who need to be treated as quickly as possible</li>
<li>If the LS values by ARFI are at least 1.35 m/s, the patient can receive antiviral treatment; if the values determined by ARFI are lower than 1.35 m/s, a LB should be performed</li>
</ul>
</td><td>
<ul>
<li>"ARFI had a very good PPV (93.2%) for predicting significant fibrosis and excellent NPV (97.8%) for excluding the presence of compensated liver cirrhosis" (page 204)</li>
</ul>
</td></tr>
<tr><td colspan="2">Friedrich-Rust, 2015[18]</td></tr>
<tr><td>
<ul>
<li>The diagnostic accuracy of PSWE and TE revealed no significant difference between the two methods in the study's analysis</li>
<li>Measurement failure was statistically significantly higher for TE than PSWE (p=0.03)</li>
<li>Cut-off values of ARFI for the diagnosis of significant LF and liver cirrhosis:</li>
</ul>
<table>
<tr><th>Stage</th><th>Cutoff</th><th>SE</th><th>SP</th><th>PPV</th><th>NPV</th></tr>
<tr><td>Exclusion of ≥F2</td><td><1.035</td><td>90.11</td><td>25.27</td><td>54.67</td><td>71.88</td></tr>
<tr><td>Diagnosis of ≥F2</td><td><1.435</td><td>64.84</td><td>90.11</td><td>86.76</td><td>71.93</td></tr>
<tr><td>Exclusion of F4</td><td><1.405</td><td>90.48</td><td>75.71</td><td>52.78</td><td>96.36</td></tr>
<tr><td>Diagnosis of F4</td><td>≥1.755</td><td>73.81</td><td>90.00</td><td>68.89</td><td>91.97</td></tr>
</table>
<ul>
<li>The area under ROC curve for ARFI and TE according to the METAVIR stages:</li>
</ul>
<table>
<tr><th>Method</th><th>F≥1</th><th>F≥2</th><th>F≥3</th><th>F=4</th></tr>
<tr><td colspan="5">Per protocol analysis (n=182)</td></tr>
<tr><td>ARFI</td><td>0.75 (0.67; 0.83)</td><td>0.81 (0.74; 0.88)</td><td>0.88 (0.82; 0.94)</td><td>0.89 (0.83; 0.96)</td></tr>
<tr><td>TE</td><td>0.80 (0.72; 0.88)</td><td>0.85 (0.80; 0.91)</td><td>0.92 (0.89; 0.97)</td><td>0.94 (0.90; 0.98)</td></tr>
<tr><td>p-value</td><td>0.28</td><td>0.15</td><td>0.11</td><td>0.19</td></tr>
<tr><td colspan="5">Intention to diagnose analysis (n = 235)</td></tr>
<tr><td>ARFI</td><td>0.77 (0.70; 0.84)</td><td>0.81 (0.76; 0.87)</td><td>0.86 (0.81;0.92)</td><td>0.87 (0.80; 0.93)</td></tr>
<tr><td>TE</td><td>0.77 (0.71; 0.84)</td><td>0.85 (0.80; 0.90)</td><td>0.88 (0.83; 0.93)</td><td>0.89 (0.84; 0.95)</td></tr>
<tr><td>p-value</td><td>0.90</td><td>0.25</td><td>0.66</td><td>0.48</td></tr>
</table>
</td><td>
<ul>
<li>The diagnostic accuracy of PSWE was comparable to TE for testing liver fibrosis in patients with CHC</li>
</ul>
</td></tr>
<tr><td colspan="2">Guerra, 2015[3]</td></tr>
<tr><td>
<ul>
<li>The performance of ARFI exhibited an accuracy of 82.4%, 71.4% SE, and 90% SP in CHC patients</li>
<li>The optimal cutoff value of ARFI electrography for distinguishing advanced fibrosis from initial fibrosis was 1.535 m/s with 83.3% SE and 91.7% SP</li>
<li>ARFI values between 1.09 m/s and 1.70 m/s were not able to differentiate between advanced fibrosis and initial fibrosis; therefore, a liver biopsy would be recommended to evaluate liver fibrosis</li>
</ul>
</td><td>
<ul>
<li>For CHC patients, ARFI demonstrated good accuracy in differentiating initial fibrosis from advanced fibrosis and can replace LB in most cases</li>
</ul>
</td></tr>
</table>

**Table A8:** Summary of Findings of Included NRSs

| Main Study Findings | Author's Conclusions |
|---|---|
| **Joo, 2015[19]** | |
| • LS measured by ARFI increased in parallel with METAVIR fibrosis stage<br>• The DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages:<br>**Stage / AUROC / Cutoff / SE / SP / PPV / NPV**<br>F0v≥F1 / 0.872 / 1.190 / 84 / 85.7 / 98.8 / 28.6<br>≤F1v≥F2 / 0.853 / 1.335 / 83.8 / 75.8 / 87.7 / 69.4<br>≤F2v≥F3 / 0.840 / 1.645 / 79.5 / 75.8 / 67.4 / 85.5<br>≤F3vF4 / 0.828 / 1.665 / 85.0 / 69.1 / 40.5 / 94.9 | • ARFI demonstrated acceptable diagnostic performance on the assessment of hepatic fibrosis in patients with CHC |
| **Tai, 2015[20]** | |
| • When the low ARFI cutoff value was set at 1.41 m/s, SE was 0.706 and SP was 0.803 for CHC<br>• Through logistic regression, ARFI demonstrated to be the best modality for predicting liver cirrhosis in CHC patients (OR = 18.469, 95% CI 3.064-111.3, p<0.001) | • "ARFI is the modality of choice for predicting liver cirrhosis in CHC" (page 819) |
| **Nishikawa, 2014[21]** | |
| • ARFI was affected by BMI, γ-glutamyltranspeptidase, and hyaluronic acid in each fibrosis stage; therefore, careful attention should be made to these factors when analyzing fibrosis stage by ARFI<br>• DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages:<br>**Stage / AUROC / Cutoff / SE / SP / PPV / NPV**<br>≥F1 / 0.810 / 1.28 / 69.1 / 85.7 / 97 / 29.3<br>≥F2 / 0.909 / 1.28 / 81.8 / 87.1 / 94 / 65.9<br>≥F3 / 0.869 / 1.44 / 88.9 / 82.5 / 78.4 / 91.2<br>F4 / 0.885 / 1.73 / 82.5 / 86.2 / 48 / 97.6 | • ARFI correlated significantly with fibrosis stage, which is consistent with previous findings |
| **Silva Junior, 2014[8]** | |
| • Liver propagation velocity obtained with ARFI was highly correlated with the METAVIR fibrosis stage<br>• DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages:<br>**Stage / AUROC / Cutoff / SE / SP / PPV / NPV**<br>≥F1 / 0.88 / 1.19 / 88.4 / 75 / 95 / 54.6<br>≥F2 / 0.90 / 1.31 / 89.3 / 87 / 89.2 / 87<br>≥F3 / 0.97 / 1.68 / 94.4 / 90.9 / 85 / 96.8<br>F4 / 0.98 / 1.95 / 100 / 95.2 / 81.8 / 100 | • ARFI had very good accuracy for the assessment of liver fibrosis<br>• "ARFI measurements increase as fibrosis stages increase and correlated best with severe liver fibrosis and cirrhosis" (page 991) |
| **Takaki, 2014[22]** | |
| • The diagnostic accuracy of ARFI alone was partially inferior to APRI and FIB-4 index<br>• DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages:<br>**Stage / AUROC / Cutoff / SE / SP / PPV / NPV**<br>≥F2 / 0.882 / 1.205 / 75.0 / 90.9 / 75.0 / 90.9<br>≥F3 / 0.858 / 1.595 / 94.3 / 81.8 / 94.3 / 81.8<br>F4 / 1.000 / 1.775 / 100 / 40 / 85.7 / 100 | • "The diagnostic accuracy of SWV [shear wave velocity, part of ARFI] of LF was excellent" (page 285) |
| **Yamada, 2014[23]** | |
| • The ARFI value increased with the progression of the fibrosis stage; the ARFI value had a strong correlation with the fibrosis stage(p<0.001, Pearson correlation coefficient=0.764) | • "ARFI offers equivalent or higher diagnostic accuracy for LF compared to FibroScan" (page 246) |

**Table A8:** Summary of Findings of Included NRSs

| Main Study Findings | Author's Conclusions |
|---|---|
| • DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages: | |

| Stage | AUROC | Cutoff | SE | SP | PPV | NPV |
|---|---|---|---|---|---|---|
| ≥F1 | NR | NR | NR | NR | NR | NR |
| ≥F2 | 0.890 | 1.26 | 92.5 | 76.2 | 64.9 | 95.5 |
| ≥F3 | 0.943 | 1.46 | 84.6 | 87.8 | 64.7 | 95.6 |
| F4 | NR | NR | NR | NR | NR | NR |

Chen, 2012[24]

| Main Study Findings | Author's Conclusions |
|---|---|
| • DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages: | • "Comparisons of concurrent FibroTest and ARFI LSM indicate that ARFI LSM may provide a promising alternative, or adjunctive predictive solution, for evaluating LF in patients with CHC" (page 11) |

| Stage | AUROC | Cutoff | SE | SP | PPV | NPV |
|---|---|---|---|---|---|---|
| F1v≥F2 | 0.847 | 1.55 | 74.1 | 87 | 90.9 | 65.6 |
| ≤F2v≥F3 | 0.902 | 1.81 | 90.2 | 89.5 | 80.4 | 95.0 |
| ≤F3vF4 | 0.831 | 1.98 | 88.9 | 79.8 | 42.1 | 97.8 |

Sporea, 2012[25]

| Main Study Findings | Author's Conclusions |
|---|---|
| • ARFI demonstrated a highly significant correlation (Spearman $r = 0.654$) when measuring fibrosis stage ($p<0.0001$)<br>• DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages: | • "LS measurement by means of ARFI is a reliable method for predicting fibrosis severity in CHC patients" (page 4118) |

| Stage | AUROC | Cutoff | SE | SP | PPV | NPV |
|---|---|---|---|---|---|---|
| ≥F1 | 0.779 | >1.19 | 69.9 | 80 | 95.4 | 16 |
| ≥F2 | 0.792 | >1.33 | 69.1 | 79.8 | 87.3 | 56.1 |
| ≥F3 | 0.829 | >1.43 | 74.8 | 81.5 | 76.3 | 79.8 |
| F4 | 0.842 | >1.55 | 84.3 | 76.3 | 53.1 | 93.7 |

• TE was better than ARFI for predicting the presence of liver cirrhosis and fibrosis

Sporea, 2012[26]

| Main Study Findings | Author's Conclusions |
|---|---|
| • LS measurements assessed by ARFI was statistically significant in patients with CHC ($r=0.490$, $p<0.0001$)<br>• LS measurements assessed by TE was statistically significant in patients with CHC ($r=0.660$, $p<0.0001$)<br>• ARFI and TE were able to differentiate between patients with no liver fibrosis (F0) and mild fibrosis (F1) and those with moderate fibrosis (F2) | • "ARFI and TE were statistically significantly correlated with the histologic fibrosis stage" (page 1315)<br>• "For the noninvasive assessment of liver fibrosis, ARFI had similar predictive value with TE in both CHC and CHB" (page 1315)<br>• "The conclusion of this study was that ARFI might represent an alternative to TE" (page 1315) |

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; CHB = Chronic hepatitis B; CHC = Chronic hepatitis C; DA = Diagnostic accuracy; FI= Forns' index; LB = Liver biopsy; LF = Liver fibrosis; LS = Liver stiffness; LSM= Liver stiffness measurement; NPV = Negative predictive value; NRS = Non-randomized Study; PSWE = Point shear w ave elastography ; PPV = Positive predictive value; ROC = Receiver operating characteristic; SE= Sensitivity; SP = Specificity; TE = Transient elastography

| **Table A9:** Summary of Findings of Included Economic Studies | |
|---|---|
| **Main Study Findings** | **Author's Conclusions** |
| Crossan, 2015[4] | |
| **Base-case analysis**<br>Liver biopsy:<br>Cost was £48,710, 14.03 QALYs<br><br>ARFI:<br>Cost was £47,126, 14.25 QALYs<br><br>TE (FibroScan):<br>Cost was £47,449, 14.28 QALYs<br><br>No treatment:<br>Cost was £54,878, 12.45 QALYs<br><br>Treat all:<br>Cost was £51,241, 14.73 QALYs | • "Liver biopsy was dominated by less costly and more effective non-invasive options among both chronic hepatitis B and C patients" (pages: 69, 73, 93) |

## APPENDIX 5: Additional References of Potential Interest

### Non-Randomized Studies – Mixed Population (HBV + HCV)

Gerber L, Kasper D, Fitting D, Knop V, Vermehren A, Sprinzl K, et al. Assessment of liver fibrosis with 2-D shear wave elastography in comparison to Transient elastography and acoustic radiation force impulse imaging in patients with chronic liver disease. Ultrasound Med Biol. 2015 Sep;41(9):2350-9.
PubMed: PM26116161

Goertz RS, Sturm J, Zopf S, Wildner D, Neurath MF, Strobel D. Outcome analysis of liver stiffness by ARFI (acoustic radiation force impulse) elastometry in patients with chronic viral hepatitis B and C. Clin Radiol. 2014 Mar;69(3):275-9.
PubMed: PM24309197

Trovato FM, Atzori S, Musumeci G, Tooley V, Marcinkowski H, Crossey MM, et al. Liver and spleen Transient elastography and Acoustic Radiation Force Impulse Measurements. Performance and comparison of measurements in the same area concurrently assessed for liver fibrosis by biopsy. Adv Med Sci. 2015 Sep;60(2):300-6.
PubMed: PM26143473

Yap WW, Kirke R, Yoshida EM, Owen D, Harris AC. Non-invasive assessment of liver fibrosis using ARFI with pathological correlation, a prospective study. Ann Hepatol. 2013 Jul;12(4):608-15.
PubMed: PM23813139

Chen SH, Li YF, Lai HC, Kao JT, Peng CY, Chuang PH, et al. Noninvasive assessment of liver fibrosis via spleen stiffness measurement using acoustic radiation force impulse sonoelastography in patients with chronic hepatitis B or C. J Viral Hepat. 2012 Sep;19(9):654-63.
PubMed: PM22863270

Kircheis G, Sagir A, Vogt C, Vom Dahl S, Kubitz R, Haussinger D. Evaluation of acoustic radiation force impulse imaging for determination of liver stiffness using Transient elastography as a reference. World J Gastroenterol [Internet]. 2012 Mar 14 [cited 2016 Apr 18];18(10):1077-84. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3296981
PubMed: PM22416182

Guo Y, Parthasarathy S, Goyal P, McCarthy RJ, Larson AC, Miller FH. Magnetic resonance elastography and acoustic radiation force impulse for staging hepatic fibrosis: a meta-analysis. Abdom Imaging. 2015 Apr;40(4):818-34.
PubMed: PM24711064

### Non-Randomized Studies – Alternate Reference Standard (Transient Elastography)

Potthoff A, Attia D, Pischke S, Kirschner J, Mederacke I, Wedemeyer H, et al. Influence of different frequencies and insertion depths on the diagnostic accuracy of liver elastography by acoustic radiation force impulse imaging (ARFI). Eur J Radiol. 2013 Aug;82(8):1207-12.
PubMed: PM23523513