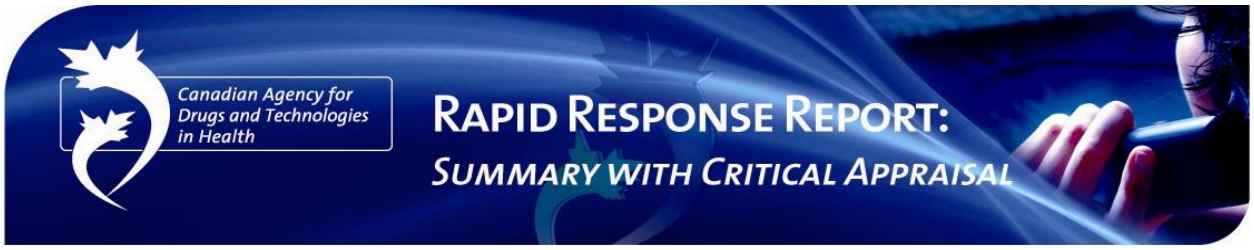


TITLE: Acoustic Radiation Force Impulse Imaging for Diagnosis and Monitoring of Liver Fibrosis in Patients with Hepatitis B: A Review of Diagnostic Accuracy, Clinical Effectiveness, Cost-Effectiveness, and Guidelines

DATE: 04 May 2016

CONTEXT AND POLICY ISSUES

Hepatitis B is an infectious disease that affects the liver. The hepatitis B virus (HBV) is transmitted through contact with body fluids and infected blood.[1] Globally, there are more than 300 million carriers of HBV around the world and 500,000 to 1.2 million die yearly from related liver disease.[1] The most recent figures, from 2011, report that less than 1% of Canada's population is infected with acute or chronic HBV.[1] HBV can be prevented by a vaccine, however, people who have not received the vaccine are unprotected against the virus. Most adults (~95%) who become infected recover within 6 months, known as acute hepatitis B; the remaining ~5% of adults go on to develop chronic hepatitis B (CHB).[1] CHB is often asymptomatic until the liver is severely damaged.[2] HBV is treatable and, if it is caught early enough, further liver damage can be prevented.

HBV causes inflammation and scarring on the liver, which can affect how it functions.[3] The early stages of inflammation are known as fibrosis and, if left untreated, the scarring worsens and becomes cirrhosis.[3] Cirrhosis can lead to permanent damage or scarring of the liver.[4] Cirrhosis may block blood flow though the liver, preventing normal metabolic and regulatory processes.[4] Measuring the fibrosis stage is crucial to the management and prognosis of CHB.[5]

The currently 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.[6] Moreover, the specimen obtained from the liver biopsy must meet certain quality criteria, which is not always possible in a clinical setting.[6] Liver biopsies can also be quite painful and have potential complications including risk of death.[7] 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.[8] One significant advantage of ARFI imaging is that it is integrated into a conventional ultrasonographic system.[9] This also allows for a sonographic

evaluation of the liver to be performed simultaneously with ARFI, this providing patients an ideal 'one-stop shop' for noninvasive liver evaluation, even in patients with a significant amount of ascites.[10] ARFI is a technology designed to measure shear wavefront at multiple locations to calculate tissue stiffness.[8] The wave velocity determines tissue stiffness through a simple method: the stiffer the tissue, the greater velocity.[8]

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 B. This report serves as a companion to a 2016 Rapid Response report[11] that reviewed the clinical and cost-effectiveness of ARFI among patients with hepatitis C. That report suggested that ARFI was a comparable method to liver biopsy to evaluate liver fibrosis and cirrhosis in patients with hepatitis C. An economic evaluation 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 treatment options and may not reflect current practice.

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 B?

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 B?

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 B?

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 B?

KEY FINDINGS

Eleven studies were identified regarding the diagnostic accuracy of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis with hepatitis B. There were mixed results on the use of ARFI to evaluate liver fibrosis in patients with hepatitis B. One study was identified on the cost-effectiveness of ARFI. ARFI was dominated by less costly and more effective options among hepatitis B e antigen-positive patients (high levels of the virus and greater infectiousness); however, ARFI was extendedly dominated in hepatitis B e antigen-negative patients (low to zero level of the virus and less infectious). No literature was found regarding the clinical effectiveness or evidence-based guidelines.

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 B
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 (clinical benefit, harms), diagnostic accuracy (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 B findings separately from other etiologies.

Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised using the AMSTAR checklist,[12] the non-randomized studies were critically appraised using the QUADAS-2 tool,[13] and the economic evaluation was critically appraised using the Drummond checklist.[14] 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, 199 citations were excluded and 25 potentially relevant reports from the electronic search were retrieved for full-text review. One potentially relevant publication was retrieved from the grey literature search. Of these potentially relevant articles, 14 publications were excluded for various reasons, while 12 publications met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection.

Summary of Study Characteristics

Details of the study characteristics are located in Appendix 2.

Study Design

Three systematic reviews (SRs)[2,15,16] regarding the diagnostic accuracy of acoustic force impulse imaging (ARFI) for patients with hepatitis B were identified. One of the SRs[2] pooled results from other SRs that were conducted; two of the SRs[15,16] combined the results of the individual studies in a meta-analysis. One SR[2] examined fourteen SRs; the findings from 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 B in isolation from other etiologies. Another SR[15] reviewed thirty-six non-randomized studies; one abstract analyzed patients with hepatitis B in isolation from other etiologies and is included in this report. This SR also conducted a subgroup analysis on patients with hepatitis B from all the studies, which is also described this report. The third SR[16] examined eight non-randomized studies, which also performed a subgroup analysis on patients with hepatitis B from all the studies.

Eight non-randomized studies (NRSs),[5,17-23] were also identified regarding the diagnostic accuracy of ARFI for patients with hepatitis B. Six of these NRSs[17-20,22,23] were designed as cross-sectional and two of these NRSs[5,21] were designed as case-control studies.

One HTA[24] included a systematic review and meta-analysis and an economic evaluation. The clinical findings are captured in the systematic review of reviews.[2] 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[2] was conducted in Canada. Two SRs[15,16] and one NRS[22] were conducted in Germany. One NRS[17] was conducted in Korea. Another NRS[18] was conducted in Taiwan. Four NRSs[5,19-21] were conducted in China. One NRS[23] was conducted in Romania. The economic evaluation[24] reflected a United Kingdom Ministry of Health perspective.

Patient Population

Three SRs[2,15,16] examined all cause of liver disease, which included hepatitis B virus, hepatitis C virus, nonalcoholic fatty liver disease, alcoholic liver disease or cholestatic diseases.

Six NRSs[5,17,19-22] examined patients with hepatitis B and two NRSs[18,23] examined patients with hepatitis B and C, with findings in hepatitis B patients reported separately.

The economic evaluation[24] generated separate estimates of costs and effects for populations with chronic hepatitis B.

Interventions and Comparators

All of the studies[2,5,15-24] used liver biopsy as the reference standard and ARFI as the index test. Cut-off values for the index test ranged from 1.03 to 1.98. Six of the studies[2,19-23] also examined transient elastography (TE) as an index test.

The economic evaluation[24] investigated several non-invasive tests identified in their systematic review, including ARFI and TE, for patients with chronic hepatitis B.

Outcomes

All of the studies[2,5,15-24] examined the diagnostic accuracy of ARFI as a primary outcome.

In the economic evaluation by Crossan et al.,[24] 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 a METAVIR score of ≥ F2 indicated a positive test outcome and that antiviral treatment would begin at this stage.

Summary of Critical Appraisal

Details of the critical appraisal are located in Appendix 3.

Two of the SRs[2,15] contained a study and an abstract of higher quality. An "a priori" design was provided in two of the SRs.[2,15] A comprehensive literature search, with a detailed search strategy, was provided in two of the SRs.[2,15] One of the SRs[16] did not indicate the use of an "a priori" design, but a comprehensive search strategy was provided. One SR[15] completed study selection by two independent reviewers, the second SR[2] completed study selection by a single reviewer, and it was unclear if the third SR[16] performed study selection by one or two reviewers. None of the SRs[2,15,16] provided a list of excluded studies. One SR[16] did not assess and document the quality of the included studies. Two of the SRs[2,15] used validated tools to assess the quality of the included studies. The methods used to combine the findings of studies were appropriate in all of the SRs.[2,15,16] One of the SRs[16] declared no conflicts of interest; two SRs[2,15] did not include a declaration of conflicts of interest.

Most of the NRSs[17-20,22,23] avoided a case-control design; two of the NRSs[5,21] did use a case-control design. All of the studies[5,17-23] avoided inappropriate exclusions. The patient selection was consecutive in five of the NRSs[17-20,22]; it was unclear if the patient selection was consecutive in three of the NRSs.[5,21,23] Five of the NRSs[17,18,20-23] blinded the results of the index test when interpreting the results of the reference standard and vice versa. It was unclear if the index test and/or reference standard was blinded in one of the NRSs.[19] Another study[5] blinded the results of the reference standard when interpreting the results of the index test, but it was unclear if the index test results were interpreted without knowledge of the results of the reference standard. Six of the NRSs[5,17-19,21,23] performed the index test and/or reference standard within three days of the other test. One NRSs[22] had an long interval (up to four weeks) between the index test and reference standard, which may not have been appropriate, and it was unclear if there was an appropriate time interval between the index test and reference standard in another NRS.[20] A threshold was pre-specified in one of the NRSs.[17] The other NRSs[5,18-23] either did not pre-specify a threshold or it was unclear if a there was threshold.

The economic evaluation[24] 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 B 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 cost-utility analysis. In addition, quantities of resource use were not reported separately from their unit costs and the distributions chosen from the probabilities sensitivity analysis were not justified.

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 B?

One SR[2] examined the diagnostic accuracy of ARFI and TE compared to liver biopsy. The study concluded that ARFI was similar to TE for assessing liver fibrosis patients with hepatitis B. The one systematic review that met the inclusion for the report only provided information for the the sensitivity and specificity in one fibrosis stage (≥F2) (Sensitivity 71%, Specificity 67%). Another SR[15] found that there was a significantly better diagnostic accuracy for the diagnosis of severe fibrosis in studies who excluded HBV-infected patients than those who included HBV-infected patients. The reason for this difference may be because liver fibrosis is distributed more heterogeneously in patients with HBV. A third SR[16] also found a trend towards lower diagnostic accuracy in patients with HBV than patients with hepatitis C. The SR reported good diagnostic accuracy using ARFI for the diagnosis of significant liver fibrosis (F≥3) (83% Area under the receiver operator characteristic curve [AUROC]) and excellent diagnostic accuracy for diagnosing liver cirrhosis (F=4) (90% AUROC) in patients with hepatitis B.

One NRS[17] concluded that ARFI exhibited an acceptable diagnostic accuracy in predicting liver fibrosis in patients with HBV. Another NRS[18] reported a sensitivity of 52.8% and a specificity of 70.5% in patients with hepatitis B using ARFI; it also concluded that the ARFI AUROC was lower in patients with HBV than patients with hepatitis C. Two NRSs[19,21] found that ARFI and TE were proven to be reliable methods to assess fibrosis in patients with HBV. Another NRS[20] concluded that ARFI may be a reliable method for diagnosing the stage of liver fibrosis in patients with HBV; ARFI also had similar predictive value to TE. One NRS[22] found that ARFI was a reliable modality for the assessment of advanced stages of liver fibrosis in chronic hepatitis B patients. Another NRS[23] reported that ARFI and TE could not differentiate between earlier stages of liver fibrosis in chronic hepatitis B patients. One NRS[5] concluded that ARFI was a reliable predictor of liver fibrosis in patients with hepatitis B (Sensitivity 93.9 to 95.7%, Specificity 91.8 to 95.0%).

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 B?

The economic evaluation[24] assessed the cost and effect tradeoff of various non-invasive methods relate to liver biopsy among chronic hepatitis B patients. Whether patients were e antigen-positive or negative, liver biopsy was dominated by less costly and more effect non-

invasive options. Mean costs for liver biopsy varied from £70,274 to £75,957 and mean effects varied from 9.64 QALYs to 11.41 QALYs. Among hepatitis B e antigen-positive patients (high levels of the virus and greater infectiousness) costs for ARFI and TE were £83,487 and £79,004 and mean effects were 11.71 QALYs and 11.61 QALYs, respectively. Among hepatitis B e antigen-negative patients (low to zero level of the virus and less infectious) costs for ARFI and TE were £77,512 and £73.007 and mean effects were 10.10 QALYs and 9.93 QALYs, respectively. ARFI was dominated by less costly and more effective options among hepatitis B e antigen-positive patients; ARFI was extendedly dominated in hepatitis B e antigen-negative patients.

Limitations

Two of the SRs[2,15] included a systematic review and an abstract that were of higher quality. One of the SRs[16] did not provide any evidence of quality assessment. This affects the reliability of the study findings, as it is not known if the included studies were of low, moderate, or high quality.

There were inconsistencies among all the studies in the various fibrosis stages when measuring ARFI and/or TE. The three SRs[2,15,16] used the same fibrosis staging (F≥2, F≥3, F=4). Some of the NRSs[21,22,25] used METAVIR to grade fibrosis, while other studies[5,19,20] used an alternate method to grade fibrosis. Two of the NRSs[18,23] did not report their findings according to fibrosis staging. These inconsistencies between studies make it more difficult to interpret findings. Additionally, two of the NRSs[5,21] only reported findings for select stages of liver fibrosis, rather than all the fibrosis stages.

Most of the NRSs[5,18-23] did not provide any indication that a pre-determined threshold was used to determined fibrosis stage. This represents a potential limitation in the interpretation of their results because cut-offs were created once the index test and reference standard had been conducted. This creates potential bias, as optimal cut-off values may have been chosen to maximize the sum of sensitivity and specificity.

Most of the studies identified in this report[5,15-24] 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 in this report.

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING

This report serves as a complementary review to a previous 2016 report[11] that examined the clinical and cost-effectiveness of ARFI among patients with hepatitis C; that report suggested that ARFI was a comparable method to liver biopsy to evaluate liver fibrosis and cirrhosis in patients with hepatitis C. An economic evaluation 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 treatment options and may not reflect current practice. This report includes: three SRs[2,15,16] and eight NRSs[5,17-23] on the diagnostic accuracy of acoustic radiation force impulse imaging compared with liver biopsy for detecting and grading liver fibrosis in patients with hepatitis B. One economic evaluation[24] was identified regarding the cost-effectiveness of ARFI compared with liver biopsy. No literature was identified regarding evidence-based guidelines and clinical effectiveness of ARFI compared with liver biopsy in patients with hepatitis B.

The studies in this report[2,5,15-23] had inconsistent results on the use of ARFI for detecting and grading liver fibrosis in patients with hepatitis B. Five studies[5,17,19,21,22] were favourable to the use of ARFI in patients with hepatitis B. These studies[5,17,19,21,22] all concluded that ARFI was a reliable method for diagnosing the stage of liver fibrosis in patients with hepatitis B. Four studies[15,16,18,23] found that ARFI was an unreliable method for detecting and grading liver fibrosis in patients with hepatitis B. Two SRs[15,16] and one NRS[18] found there was a lower diagnostic accuracy in patients with hepatitis B than with patients with other liver disease (i.e., hepatitis C). Another NRS[23] found that ARFI and TE could not differentiate between earlier stages of liver fibrosis in patients with hepatitis B. Based on the findings from this report, it remains unclear if ARFI is a reliable modality to measure liver fibrosis in patients with hepatitis B. One economic evaluation[24] compared the cost-effectiveness of different options for the detection of liver fibrosis and treatment of hepatitis B. In this analysis, ARFI was dominated by less costly and more effective options among hepatitis B e antigen-positive patients; however, ARFI was extendedly dominated in hepatitis B e antigen-negative patients.

PREPARED BY:

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

REFERENCES

1. Hepatitis B - get the facts [Internet]. Ottawa: Public Health Agency of Canada; 2014 Sep 22. [cited 2016 May 3]. Available from: http://www.phac-aspc.gc.ca/hcai-iamss/bbp-pts/hepatitis/hep_b-eng.php

2. 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

3. 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

4. Cirrhosis [Internet]. Markham (ON): Canadian Liver Foundation; 2015. [cited 2016 May 3]. Available from: http://www.liver.ca/liver-disease/types/cirrhosis/

5. Ye XP, Ran HT, Cheng J, Zhu YF, Zhang DZ, Zhang P, et al. Liver and spleen stiffness measured by acoustic radiation force impulse elastography for noninvasive assessment of liver fibrosis and esophageal varices in patients with chronic hepatitis B. J Ultrasound Med [Internet]. 2012 Aug [cited 2016 Apr 18];31(8):1245-53. Available from: http://www.jultrasoundmed.org/content/31/8/1245.full.pdf+html

6. 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

7. 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 Free registration required.

8. 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

9. 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.

10. 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.

11. Accoustic 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 [Internet]. Ottawa: CADTH; 2016 Apr 18. [cited 2016 Apr 21]. (Rapid response report: summary with critical appraisal). Available from: https://www.cadth.ca/sites/default/files/pdf/htis/apr-2016/RC0771%20ARFI%20Final.pdf

12. 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

13. 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.

14. 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

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. Park MS, Kim SW, Yoon KT, Kim SU, Park SY, Tak WY, et al. Factors influencing the diagnostic accuracy of acoustic radiation force impulse elastography in patients with chronic hepatitis B. Gut Liver [Internet]. 2016 Mar 23 [cited 2016 Mar 23];10(2):275-82. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4780458

18. 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.

19. Dong DR, Hao MN, Li C, Peng Z, Liu X, Wang GP, et al. Acoustic radiation force impulse elastography, FibroScan(R), Forns' index and their combination in the assessment of liver fibrosis in patients with chronic hepatitis B, and the impact of inflammatory activity and steatosis on these diagnostic methods. Mol Med Rep [Internet]. 2015 Jun [cited 2016 Mar 23];11(6):4174-82. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4394958

20. Zhang D, Chen M, Wang R, Liu Y, Zhang D, Liu L, et al. Comparison of acoustic radiation force impulse imaging and transient elastography for non-invasive assessment of liver fibrosis in patients with chronic hepatitis B. Ultrasound Med Biol. 2015 Jan;41(1):7-14.

21. Liu Y, Dong CF, Yang G, Liu J, Yao S, Li HY, et al. Optimal linear combination of ARFI, transient elastography and APRI for the assessment of fibrosis in chronic hepatitis B. Liver Int. 2015 Mar;35(3):816-25.

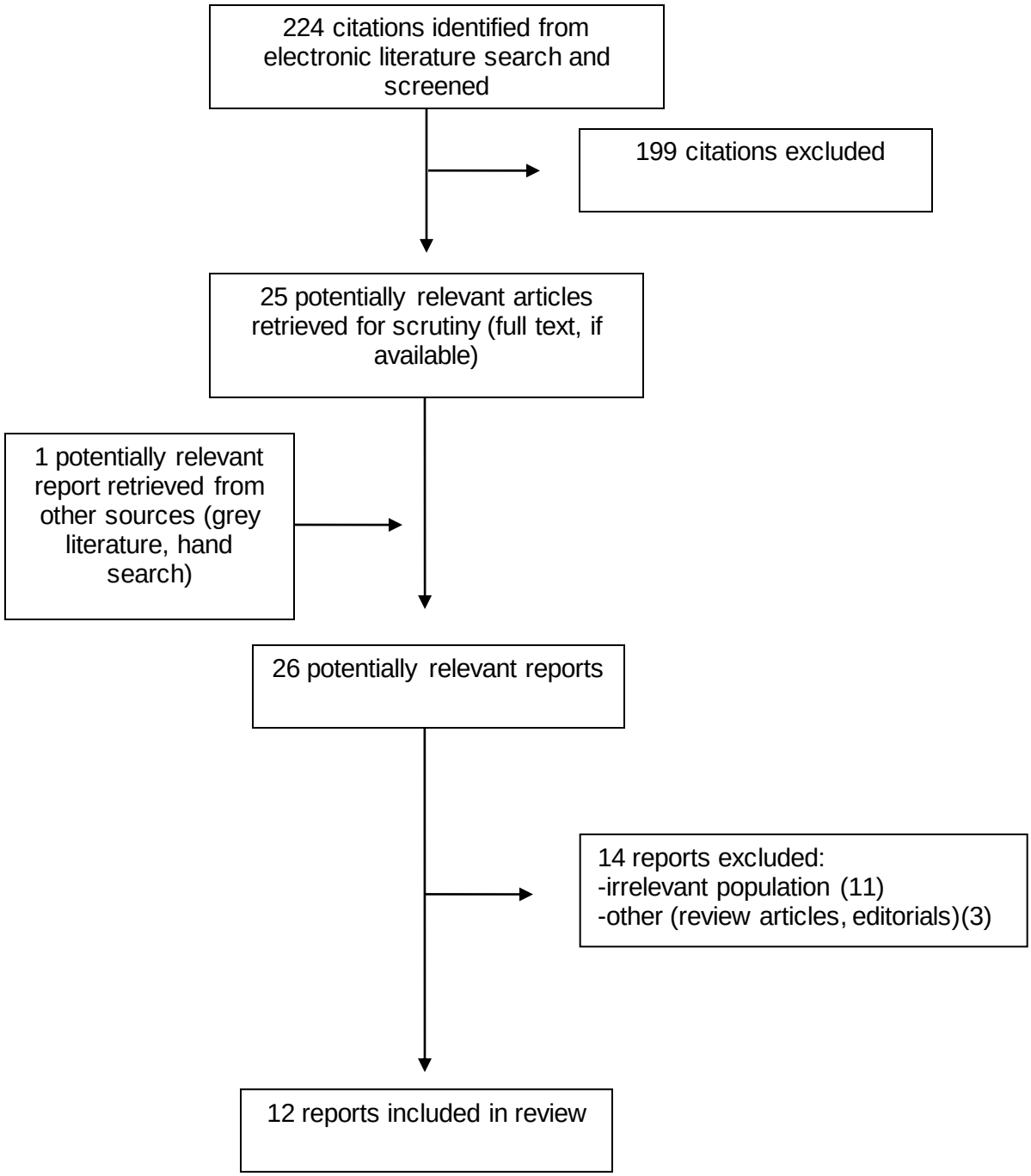
22. Friedrich-Rust M, Buggisch P, de Knegt RJ, Dries V, Shi Y, Matschenz K, et al. Acoustic radiation force impulse imaging for non-invasive assessment of liver fibrosis in chronic hepatitis B. J Viral Hepat. 2013 Apr;20(4):240-7.

23. 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.

24. 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/

25. Brouwers M, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in healthcare. CMAJ [Internet]. 2010 Dec;182(18):E839-E842. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3001530/pdf/182e839.pdf

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	Intervention	Comparator(s)	Clinical Outcomes
Brener, 2015, Canada[2]	Fourteen SRs published between 2007 and 2014 Five of these SRs evaluated the overall performance of ARFI and TE for the diagnosis of LF	All causes of liver disease, n = NR One of these reviews in this SR was reported on because it analyzed the DA of ARFI in patients with HBV	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 All of these studies evaluated the overall performance of ARFI for the diagnosis of LF	All causes of liver disease, n = 3591 One abstract in this SR was reported on because they analyzed the DA of ARFI in patients with HBV	ARFI	LB	Diagnostic accuracy of ARFI (measured by AUROC, SE, SP, PPV, NPV)
Friedrich-Rust, 2012, Germany[16]	Eight studies published until October 2010 All of these studies evaluated the overall performance of ARFI for the diagnosis of LF	All causes of liver disease, n = 518, HBV patients, n = 51	ARFI	LB	Diagnostic accuracy of ARFI (measured by measured by AUROC, SE, SP, PPV, NPV)

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; HBV = Hepatitis B virus; LB = Liver biopsy; LF = Liver fibrosis; MA = Meta-analysis; NPV = Negative predictive value; PPV = Positive predictive value; SE = Sensitivity; SP = Specificity; SR = Systematic Review ; TE = Transient elastography

Table A2: Characteristics of Included Non-Randomized Studies

First Author, Publication Year, Country	Study Design	Patient Characteristics	Intervention(s)	Comparator(s)	Clinical Outcomes
Park, 2016, Korea[17]	Cross-sectional	105 CHB patients	ARFI	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)
Tai, 2015, Taiwan[18]	Cross-sectional	204 CHC and CHB patients	ARFI	LB	Diagnostic accuracy (measured by ROC curves, AUROC, SE, SP)
Dong, 2015, China[19]	Cross-sectional	81 CHB patients	ARFI/TE	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)
Zhang, 2015, China[20]	Cross-sectional	197 CHB patients	ARFI/TE	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)
Liu, 2015, China[21]	Case-control	95 patients with CHB and 16 healthy controls	ARFI/TE	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)
Friedrich-Rust, 2013, Germany[22]	Cross-sectional	131 patients with CHB	ARFI/TE	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)
Sporea, 2012, Romania[23]	Cross-sectional	160 patients with CHB and CHC	ARFI/TE	LB	Diagnostic accuracy (measured by AUROC, SE, SP, PPV, NPV)
Ye, 2012, China[5]	Case-control	204 patients with CHB and 60 healthy controls	ARFI	LB	Diagnostic accuracy (measured by AUROC, sensitivity, specificity, PPV, NPV)

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; CHB = Chronic hepatitis B; DA = Diagnostic accuracy; LB = Liver biopsy; LS = Liver stiffness; NPV = Negative predictive value; PPV = Positive predictive value; TE = Transient elastography

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[24]	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 For chronic hepatitis C there were 57 interventions considered: 22 indirect methods, 22 direct methods and 13 imaging methods 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 Patients were assumed to enter the model at between 30 and 40 years of age Cost data were derived from the literature based on resource use information collected on inpatient and outpatient care, investigations, procedures, drug use and other services Costs were inflated to 2012 levels using NHS inflation indices Utility data were derived from the literature based on the EQ-5D preference measure within a UK population Death was assumed to have a utility value of 0 Health outcomes were measured in quality adjusted life years Costs and utilities were discounted at a rate of 3.5%

APPENDIX 3: Critical Appraisal of Included Publications

Table A4: Strengths and Limitations of Systematic Reviews and Meta-Analyses using AMSTAR[12]

Strengths	Limitations
Brener[2]	
• An "a priori" design was provided. • A comprehensive literature search was performed, including grey literature. A detailed search strategy and a flow diagram for the search results were provided. • A list of the included studies was provided. • The characteristics of the included studies were provided. • 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. • The scientific quality of the included studies was used appropriately in formulating conclusions. • 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. • A list of excluded studies was not provided. • The likelihood of publication bias was not assessed. • No conflicts of interest were mentioned by the authors.
Nierhoff[15]	
• An "a priori" design was provided. • A comprehensive literature search was performed. A detailed search strategy and a flow diagram for the search results were provided. • Study selection was done by two independent reviewers. • A list of the included studies was provided. • The characteristics of the included studies were provided. • The likelihood of publication bias was assessed. • 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. • The scientific quality of the included studies was used appropriately in formulating conclusions. • The methods used to combine the findings of studies were appropriate.	• A list of excluded studies was not provided. • No conflicts of interest were mentioned by the authors. • Abstracts were included in the meta-analysis, however, the authors of the abstracts were contacted if the data in the abstract were insufficient.
Friedrich-Rust[16]	
• A comprehensive literature search was performed. • A list of the included studies was provided. • The characteristics of the included studies were provided. • The likelihood of publication bias was assessed. • The methods used to combine the findings of studies were appropriate. • No conflicts of interest were declared by the authors. • The authors of the studies were contacted for original patient data.	• It is unclear if an "a priori" design was provided. • A flow diagram for the search results was not provided. • A list of excluded studies was not provided. • The scientific quality of the included studies was not assessed and documented, and the included studies were not rated on their quality. • The scientific quality of the included studies was not used appropriately in formulating conclusions. • It is unclear if study selection was done by two independent reviewers.

Table A5: Strengths and Limitations of Non-Randomized Studies using QUADAS-2[13]

Strengths	Limitations
Park[17]	
Patient Selection • Patient selection was consecutive • The study avoided inappropriate exclusions • A case-control design was avoided *Index Test* • A threshold was pre-specified • The index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • The reference standard was likely to classify patients appropriately • The reference standard results were interpreted without knowledge of the results of the index test *Flow and Timing* • All patients received the reference standard • There was an appropriate time interval between the index test and reference standard	*Patient Selection* • No limitations identified *Index Test* • No limitations identified *Reference Standard* • No limitations identified *Flow and Timing* • Not all patients were included in the analysis • It is unclear if all patients received the reference standard
Tai[18]	
Patient Selection • Patient selection was consecutive • The study avoided inappropriate exclusions • A case-control design was avoided *Index Test* • The index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • The reference standard was likely to classify patients appropriately • The reference standard results were interpreted without knowledge of the results of the index test *Flow and Timing* • All patients received the reference standard • All patients were included in the analysis • There was an appropriate time interval between the index test and reference standard	*Patient Selection* • No limitations identified *Index Test* • It is unclear if a threshold was pre-specified *Reference Standard* • No limitations identified *Flow and Timing* • No limitations identified
Dong[19]	
Patient Selection • Patient selection was consecutive • The study avoided inappropriate exclusions • A case-control design was avoided *Index Test* • No strengths identified *Reference Standard* • The reference standard was likely to classify patients appropriately *Flow and Timing* • All patients received the reference standard • All patients were included in the analysis	*Patient Selection* • No limitations identified *Index Test* • It is unclear if the index test results were interpreted without knowledge of the results of the reference standard • A threshold was not pre-specified *Reference Standard* • It is unclear if the reference standard results were interpreted without knowledge of the results of the index test

Table A5: Strengths and Limitations of Non-Randomized Studies using QUADAS-2[13]	
Strengths	**Limitations**
• There was an appropriate time interval between the index test and reference standard	*Flow and Timing* • No limitations identified
Zhang[20]	
Patient Selection • Patient selection was consecutive • The study avoided inappropriate exclusions • A case-control design was avoided *Index Test* • The index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • The reference standard results were interpreted without knowledge of the results of the index test • The reference standard was likely to classify patients appropriately *Flow and Timing* • All patients received the reference standard	*Patient Selection* • No limitations identified *Index Test* • A threshold was not pre-specified *Reference Standard* • No limitations identified *Flow and Timing* • Not all patients were included in the analysis • It is unclear if there was an appropriate time interval between the index test and reference standard
Liu[21]	
Patient Selection • The study avoided inappropriate exclusions *Index Test* • No strengths were identified *Reference Standard* • The reference standard was likely to classify patients appropriately *Flow and Timing* • All patients received the reference standard • There was an appropriate time interval between the index test and reference standard (same day or 1 day after)	*Patient Selection* • It is unclear if patient selection was consecutive • A case-control design was used *Index Test* • A threshold was not pre-specified • It is unclear if the index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • It is unclear if the reference standard results were interpreted without knowledge of the results of the index test *Flow and Timing* • Not all patients were included in the analysis
Friedrich-Rust[22]	
Patient Selection • Patient selection was consecutive • The study avoided inappropriate exclusions • A case-control design was avoided *Index Test* • The index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • The reference standard results were interpreted without knowledge of the results of the index test • The reference standard was likely to classify patients appropriately *Flow and Timing* • All patients received the reference standard	*Patient Selection* • No limitations identified *Index Test* • A threshold was not pre-specified *Reference Standard* • No limitations identified *Flow and Timing* • Not all patients were included in the analysis • There was a long time interval between the index test and reference standard (up to 4 weeks)

Table A5: Strengths and Limitations of Non-Randomized Studies using QUADAS-2[13]

Strengths	Limitations
Sporea, 2012[23]	
Patient Selection • A case-control design was avoided *Index Test* • The index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • The reference standard was likely to classify patients appropriately • The reference standard results were interpreted without knowledge of the results of the index test *Flow and Timing* • All patients received the reference standard • There was an appropriate interval between the index test and the reference standard	*Patient Selection* • Unclear if patient selection was consecutive • Exclusion criteria was unclear *Index Test* • A threshold was not pre-specified *Reference Standard* • No limitations identified *Flow and Timing* • Not all patients were included in the analysis
Ye[5]	
Patient Selection • The study avoided inappropriate exclusions *Index Test* • No strengths identified *Reference Standard* • The reference standard results were interpreted without knowledge of the results of the index test • The reference standard was likely to classify patients appropriately *Flow and Timing* • There was an appropriate time interval between the index test and reference standard (1 to 3 days) • All patients were included in the analysis	*Patient Selection* • A case-control design was used • It was unclear if the patient selection was consecutive *Index Test* • A threshold was not pre-specified • It is unclear if the index test results were interpreted without knowledge of the results of the reference standard *Reference Standard* • No limitations identified *Flow and Timing* • Not all patients received the reference standard

Table A6: Strengths and Limitations of Economic Studies using Drummond[14]

Strengths	Limitations
Crossan[24]	
• The perspective of the analysis was explicit and the alternatives being compared were clearly described • Details of the sources of evidence and methods of synthesis used to inform model parameters were given • The models were described • The time horizon for costs and effects and the associated discount rates were stated • Incremental analysis was reported • The conclusions followed from the reported data	• The form of the economic evaluation was not explicitly stated • Quantities of resource use were not reported separately from their unit costs • 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>Brener, 2015[2]</td><td></td></tr>
<tr><td>

One systematic review met the inclusion criteria for this report (Tsochatzis et al. 2015)
ARFI had similar value to TE for significant fibrosis and cirrhosis
ARFI had good diagnostic accuracy for staging liver fibrosis
Noninvasive tests (ie, ARFI and TE) were good at excluding advanced cirrhosis
TE was an accurate diagnostic method for moderate fibrosis or cirrhosis
The DA of LS measurement using TE for the assessment of LF according to METAVIR stages in the study that examined <u>HBV only</u>:

<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>1 (NR)</td><td>NR</td><td>0.71 (0.59-0.80)</td><td>0.67 (0.30-0.90)</td></tr>
<tr><td>≥F3</td><td>NR</td><td>NR</td><td>NR</td><td>NR</td></tr>
<tr><td>F=4</td><td>NR</td><td>NR</td><td>NR</td><td>NR</td></tr>
<tr><td>TE</td><td></td><td></td><td></td><td></td></tr>
<tr><td>≥F2</td><td>4 (NR)</td><td>NR</td><td>0.78 (0.71-0.84)</td><td>0.78 (0.68-0.85)</td></tr>
<tr><td>≥F3</td><td>NR</td><td>NR</td><td>NR</td><td>NR</td></tr>
<tr><td>F=4</td><td>6 (NR)</td><td>NR</td><td>0.80 (0.61-0.91)</td><td>0.86 (0.82-0.94)</td></tr>
</table>
<u>Quality of Included Studies</u>
The study analyzing ARFI/TE with biopsy as a diagnostic technology was of higher quality.
</td><td>

"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)

</td></tr>
<tr><td>Nierhoff[15]</td><td></td></tr>
<tr><td>

The overall AUROC for studies for HBV patients was 0.87 (95% CI, 0.85-0.90) –compared to an AUROC of 0.92 (95% CI, 0.89-0.95) in patients without HBV
One abstract met the inclusion criteria for this report (Friedrich-Rust et al. 2011)
The DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages in the abstract that examined <u>HBV only</u>:

<table>
<tr><th>Stage</th><th>AUROC</th><th>Cutoff</th><th>SE</th><th>SP</th><th>DOR</th></tr>
<tr><td>≥F2</td><td>0.73</td><td>1.39</td><td>50</td><td>90</td><td>0.94</td></tr>
<tr><td>≥F3</td><td>0.94</td><td>NR</td><td>NR</td><td>NR</td><td>0.97</td></tr>
<tr><td>F=4</td><td>0.97</td><td>NR</td><td>NR</td><td>NR</td><td>NR</td></tr>
</table>
</td><td>

"Further subgroup analyses showed a significantly better DA for the diagnosis of severe fibrosis for the studies without HBV-infected patients than for those including HBV-infected patients" (page 3050)
"A reason for this difference might be that liver fibrosis is distributed more heterogeneously in patients with HBV and liver cirrhosis is predominantly macronodular rather than micronodular in HCV" (page 3050)

</td></tr>
</table>

Table A7: Summary of Findings of Included SRs and MAs

Main Study Findings	Author's Conclusions
Quality of Included Studies The abstract in this review was of higher quality.	
Friedrich-Rust[16]	
• There was a trend towards lower DA in CHB patients as compared to CHC patients • A reason for this trend is that liver fibrosis is distributed more heterogeneously in patients with HBV • The DA of ARFI in studies that examine patients with **HBV only**: **Stage** / **AUROC**: F≥2: 0.79 (0.63, 0.96); F≥3: 0.83 (0.70, 0.96); F=4: 0.90 (0.79, 1.00)	• "A trend towards lower DA in patients infected with CHB as compared to CHC was observed" (page e218) • "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; CHB = Chronic hepatitis B; CHC = Chronic hepatitis C; CI = Confidence interval; DA = Diagnostic accuracy; HBV = Hepatitis B; LF = Liver fibrosis; LS = Liver stiffness; NR = Not reported; SE = Sensitivity; SP = Sensitivity; TE = Transient elastography

Table A8: Summary of Findings of Included Non-Randomized Studies

Main Study Findings	Author's Conclusions
Park, 2016[17]	
• ARFI exhibited acceptable DA in predicting liver fibrosis • The DA of LS measurement using ARFI elastography for the assessment of LF according to METAVIR stages: **Stage** / **AUROC** / **Cutoff** / **SE** / **SP** / **PPV** / **NPV**: ≥F2: 0.814 / 1.31 / 89.7 / 63.0 / 94.9 / 22.2; ≥F3: 0.848 / 1.81 / 78.4 / 78.8 / 90.2 / 53.7; F4: 0.752 / 1.98 / 66.7 / 73.3 / 66.7 / 73.3	• "…advanced fibrosis stage (F3-4) showed a positive correlation with nondiscordance between LB and ARFI elastography" (page 281)
Tai, 2015[18]	
• When the low ARFI cutoff value was set at 1.35 m/s, SE was 0.628 and SP was 0.705 for CHB • Through logistic regression, ARFI demonstrated to be an acceptable modality for predicting liver cirrhosis in CHB patients (OR = 2.810, 95% CI 0.871-9.063, p<0.084) • Conventional sonography was the best modality for predicting liver fibrosis in patients with HBV	• "The ARFI AUROC for predicting liver cirrhosis was lower in patients with CHB than patients with CHC" (page 819) • "It would be better to avoid ARFI studies at the time of acute exacerbation in patients with CHB" (page 819)
Dong, 2015[19]	
• ARFI, TE, and FI scores were significantly correlated with the stage of fibrosis in patients with HBV; these noninvasive methods were adequate to evaluate the stage of liver fibrosis • The combination of FI, or other serum tests, with ARFI or TE may increase the diagnostic accuracy of these tests • The DA of LS measurement using ARFI elastography for the assessment of LF according to diagnosis of different histological stages:	• "…ARFI, TE, and FI were proven to be reliable methods to assess fibrosis in patients with HBV" (page 4181)

Table A8: Summary of Findings of Included Non-Randomized Studies

Main Study Findings	Author's Conclusions

Stage	AUROC	95%CI	Cutoff	SE	SP
≥S1	0.720	0.524-0.916	1.295	68.3	80.0
≥S2	0.762	0.627-0.896	1.295	82.9	65.0
≥S3	0.884	0.798-0.970	1.54	76.2	90.0
S4	0.723	0.501-0.944	1.835	66.7	85.5

- The DA of LS measurement using TE for the assessment of LF according to diagnosis of different histological stages:

Stage	AUROC	95%CI	Cutoff	SE	SP
≥S1	0	NR	NR	NR	NR
≥S2	0.753	0.631-0.875	10.3	51.2	90.0
≥S3	0.888	0.805-0.970	11.85	71.4	90.0
S4	0.873	0.740-1.006	9.4	100	63.6

Zhang, 2015[20]

- ARFI and TE are significantly correlated with liver fibrosis
- There were no differences between ARFI and TE
- The DA of LS measurement using ARFI for the assessment of LF according to diagnosis of different histological stages:

Stage	AUROC	Cutoff (m/s)	SE	SP	PPV	NPV
≥S2	0.815	1.24	68	72	79	60
≥S3	0.832	1.32	79	76	61	88
S4	0.754	1.41	88	72	28	98

- The DA of LS measurement using TE for the assessment of LF according to diagnosis of different histological stages:

Stage	AUROC	Cutoff (kPa)	SE	SP	PPV	NPV
≥S2	0.835	6.85	71	93	94	68
≥S3	0.824	7.75	75	84	69	87
S4	0.816	10.25	88	85	41	98

- "For the non-invasive assessment of liver fibrosis in patients with CHB, ARFI elastography may be a reliable method for diagnosing the stage of liver fibrosis and is similar in predictive value to TE" (page 13)

Table A8: Summary of Findings of Included Non-Randomized Studies

<table>
<tr><th>Main Study Findings</th><th>Author's Conclusions</th></tr>
<tr><td>Liu, 2015[21]</td><td></td></tr>
<tr><td>

The stage of fibrosis and ARFI (r = 0.85, p<0.001) and TE (r = 0.81, p<0.001) were significantly correlated
Diagnostic performance of ARFI and TE for the diagnosis of ≥F2 and cirrhosis:

<table>
<tr><th>Stage</th><th>AUROC Accuracy (%)</th><th>Cutoff</th><th>SE</th><th>SP</th><th>PPV</th><th>NPV</th></tr>
<tr><td colspan="7">ARFI</td></tr>
<tr><td>≥F2</td><td>83.50</td><td>1.27 m/s</td><td>83.95</td><td>83.05</td><td>79.06</td><td>89.41</td></tr>
<tr><td>F4</td><td>88.76</td><td>1.65 m/s</td><td>93.10</td><td>76.83</td><td>92.36</td><td>83.93</td></tr>
<tr><td colspan="7">TE</td></tr>
<tr><td>≥F2</td><td>75.27</td><td>6.60 kPa</td><td>81.80</td><td>71.24</td><td>66.43</td><td>87.57</td></tr>
<tr><td>F4</td><td>87.61</td><td>9.47 kPa</td><td>96.23</td><td>41.67</td><td>82.53</td><td>83.73</td></tr>
</table>
</td><td>

"ARFI, TE and APRI are reliable methods for assessing fibrosis in patients with CHB" (page 824)

</td></tr>
<tr><td>Friedrich-Rust, 2013[22]</td><td></td></tr>
<tr><td>

No significant difference was found between ARFI and TE when measuring LF
ARFI demonstrated excellent results for the diagnosis of advanced fibrosis; the evaluation of LF on mild fibrosis stages needs improvement
AUROC (95% CI) for ARFI imaging and TE according to the METAVIR fibrosis stage:

<table>
<tr><th>Method</th><th>F≥1</th><th>F≥2</th><th>F≥3</th><th>F=4</th></tr>
<tr><td>ARFI (n=114)</td><td>0.66</td><td>0.73</td><td>0.94</td><td>0.97</td></tr>
<tr><td>TE (n=88)</td><td>0.78</td><td>0.82</td><td>0.94</td><td>0.94</td></tr>
</table>

Cut-off values of ARFI imaging for the diagnosis of significant liver fibrosis (≥F2):

<table>
<tr><th>Stage</th><th>Cutoff (m/s)</th><th>SE</th><th>SP</th><th>PPV</th><th>NPV</th></tr>
<tr><td>Exclusion of ≥F2</td><td>1.03</td><td>91</td><td>26</td><td>32</td><td>88</td></tr>
<tr><td>Diagnosis of ≥F2</td><td>1.39</td><td>50</td><td>90</td><td>67</td><td>82</td></tr>
</table>
</td><td>

"…this study shows that ARFI imaging is a reliable ultrasound-based method for the assessment of advanced stages of liver fibrosis in CHB with the advantage that it is incorporated into a regular ultrasound machine" (page 246)

</td></tr>
</table>

Table A8: Summary of Findings of Included Non-Randomized Studies

Main Study Findings	Author's Conclusions
Sporea, 2012[23]	
• LS measurements assessed by ARFI was statistically significant in patients with CHB (r=0.356, p=0.01) • LS measurements assessed by means of ARFI elastography with histologic fibrosis was similar patients with CHB vs. those with CHC (r=0.356 vs. r=0.490, p=0.36) • LS measurements assessed by TE was statistically significant in patients with CHB (r=0.403, p=0.004) • LS values with fibrosis stages was significantly better in patients with CHC vs. those with CHB (r=0.660 vs. r=0.403, p=0.001) • The correlation of LS measurements assessed by ARFI with LF was better in patients with CHC than those with CHB, but not statistically significantly so • The correlation of LS measurements assessed by TE with LF was significantly higher in patients with CHC than those with CHB	• "In patients with CHB, ARFI elastography and TE could not differentiate between patients with no fibrosis (F0, METAVIR) and mild fibrosis (F1, METAVIR) and those at least moderate fibrosis (F≥2, METAVIR) but had good value for differentiation between patients with or without severe fibrosis (F≥3, METAVIR) and especially between noncirrhotic and cirrhotic patients (F=4, METAVIR)" (page 1314) • "The mean LS values assessed by ARFI elastography are similar in patients in CHB and CHC for the same stage of fibrosis" (page 1314) • "…for the noninvasive assessment of liver fibrosis, ARFI had similar predictive value with TE in both CHB and CHC" (page 1314)
Ye, 2012[5]	
• Liver stiffness correlated well with fibrosis • Liver stiffness cutoff values according to fibrosis stage:	• "…liver and spleen stiffness values measured by ARFI elastography are reliable predictors of liver fibrosis, especially for patients who are at high risk or not willing to undergo liver biopsy" (page 1252)

Stage	Cutoff (m/s)	AUROC	SE	SP
S≥3	1.69	0.99	93.9	95.0
S4	1.88	0.97	95.7	91.8

ARFI = Acoustic radiation force impulse imaging; AUROC = Areas under the respective receiver operator characteristics curves; CHB = Chronic hepatitis B; CHC = Chronic hepatitis C; CI = Confidence Interval; DA = Diagnostic accuracy; HBV = Hepatitis B; LF = Liver fibrosis; LS = Liver stiffness; SE= Sensitivity; SP = Sensitivity; TE = Transient elastography

Table A9: Summary of Findings of Included Economic Studies	
Main Study Findings	**Author's Conclusions**
Crossan, 2015[24]	
Hepatitis B e antigen-positive base-case analysis Liver biopsy: Cost was £75,957 Mean effect was 11.41 QALYs ARFI: Cost was £83,487 Mean effect was 11.71 QALYs FibroScan (Transient Elastography): Cost was £79,004 Mean effect was 11.61 QALYs No treatment: Cost was £37,831 Mean effect was 9.64 QALYs Treat all: Cost was £101,484 Mean effect was 12.18 **Hepatitis B e antigen-negative base-case analysis** Liver biopsy: Cost was £70,274 Mean effect was 9.64 QALYs ARFI: Cost was £77,512 Mean effect was 10.10 QALYs FibroScan (Transient Elastography): Cost was £73,007 Mean effect was 9.93 QALYs No treatment: Cost was £37,579 Mean effect was 8.83 QALYs Treat all: Cost was £96,525 Mean effect was 10.92 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)