H.2 Diagnostic tests

Study	Arena 2008[8]
Study type	Prospective cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=161 consecutive patients, 11 excluded due to liver biopsy length, final analysis n=150). Recruitment between 1 September 2006 and 1 July 2007.
Countries and Settings	Italy, University Hospital.
Funding	Academic or government (grants from the Italian Ministry of Education, Universities and Research, the University of Florence and the Italian Liver Foundation and Instituto de Salud Carlos III, Spain).
Age, gender, ethnicity	Age, mean (SD): 50.6 (12.5), range 21–70 years; male/female: 92/58; ethnicity: not reported; ALT (U/l): not reported
Patient characteristics	Population: HCV-related chronic liver disease referred for the histopathological assessment of disease progression. Inclusion: levels of ALT >1.5-fold the upper normal limit either persistently or intermittently, and detectable HCV RNA. Exclusion: BMI ≥30; presence of ascites at clinical or ultrasound examination; presence of HCC or previous/current decompensation of the disease; co-infection with HIV or HBV; use of IV drugs, previous or current alcohol abuse or the use of hepatotoxic drugs, genetic liver disease, autoimmune hepatitis, vascular diseases of the liver, biliary tract disorders, ongoing or recent (within 1 year) therapy with antiviral agents, cardiac failure, age <18 or >70 years and pregnancy.
Index test (including threshold and	Transient elastography (Fibroscan, Echosens, France), optimal cut-off threshold calculated (14.8 kPa): operator was a staff physician (AU) who had previously performed determinations in patients with chronic liver disease. Considered

Study	Arena 2008[8]
whether threshold pre-specified)	representative measurements of the median value of 10 successful acquisitions with a success rate of at least 60%, and with an IQR over median ratio lower than 30%.
Reference standard	Liver biopsy (METAVIR F4): performed on the right lobe of the liver with a 16 G semiautomatic modified Menghini needle system (BIOMOL; Hospital Service, Aprilia, Italy) under local anaesthesia and ultrasound guidance. Only samples with a length >25 mm and including at least 11 complete portal tracts were considered adequate (average 33(0.7) mm and 15(3) portal tracts). Sections of liver tissue, 5 mm thick, were stained with haematoxylin & eosin and Masson trichrome, and were examined by an experienced pathologist.
Time between index test and reference standard	Same day
Target condition	Cirrhosis
Prevalence of cirrhosis according to reference standard	29/150 (19.33%)
Results: Fibroscan AUC (90% CI): 0.98 (0.950.99) Optimal cut-off threshold (if calculated): 14.8 kPa Threshold: 14.8 kPa (optimal) Sensitivity: 94 Specificity: 92 Positive predictive value (PPV): 73 Negative predictive value (NPV): 98 +ve/-ve likelihood ratios: 11.27/0.07 True positives (TP): Not reported False positives (FP): Not reported False negatives (FN): Not reported True negatives (TN): Not reported Other measures reported and conclusions: Also reported multilevel likelihood ratios (LRs) and concluded that thresholds of <12 kPa and >18 kPa were adequate to rule-out or rule-in cirrhosis respectively (LRs above 10 and below 0.1 and considered strong evidence to rule in and rule out respectively). Values between 12 and 18 kPa could not reliably predict the presence or absence of cirrhosis at multilevel LRs analysis.	

Study	Arena 2008[8]
<12 kPa: LR 0 (0–0.139); ≥12 and <15: LR 1.34 (0.472–3.831); ≥15 and <18: LR 2.318 (0.986–5.449); ≥18 LR 87.621 (16.760–458.074). Any complications associated with tests reported: No major complications were associated with percutaneous liver biopsy. Fifteen patients (10%) experienced a self-limiting abdominal and/or right shoulder pain, and 6 patients (4%) required a single dose of intravenous analgesic drug (tramadol). There were no complications associated with transient elastography (TE).	
General limitations according to QUADAS II: Unclear if reference standard interpreted without knowledge of the index test result.	

Study	Aykut 2014[10]
Study type	Prospective cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=88 NAFLD patients). Recruitment period not reported.
Countries and Settings	Department of Gastroenterology, University School of Medicine, Turkey
Funding	Academic or Government funding (Marmara University Scientific Research Fund).
Age, gender, ethnicity	Age, mean (SD): 46 (9); male/female: 50/38; ethnicity: not reported; ALT (U/l): 84 (56); BMI: 30.3 (4.6)
Patient characteristics	Population: NAFLD Inclusion: Persistent (>6 months) elevation of transaminases and steatosis on ultrasound; subjects with normal transaminases in presence of hepatomegaly and/or splenomegaly; subjects with normal transaminases but persistently increased gamma-glutamyl transferase. Absent to low alcohol consumption (<30 g/day men and <20 g/day women). Exclusion: Viral hepatitis B or C, Wilson's disease, alpha1-antitrypsin deficiency, autoimmune hepatitis, genetic haemochromatosis and use of steatogenic drugs. Other conditions known to cause liver dysfunction.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan, Echosens, France), optimal cut-off threshold not reported. A single operator performed all examinations according to the manufacturer's protocol. With the patient lying in the dorsal secubitus position, the tip of the transducer was placed on the skin between the ribs over the right lobe of the liver. Assessment performed using the M or XL probe as appropriate. Measurement depth between 25 and 65 mm for the M probe and 35 and 75 mm for the XL probe. Subjects with failures or unreliable measurements were excluded. Failure defined as zero valid shots and unreliable examinations were defined as fewer than 10 valid shots, a success rate <60% or an IQR >30%.

Study	Aykut 2014[10]
Reference standard	Liver biopsy (NAFLD activity score F4 [reference McPherson 2010 paper which used Kleiner score]): all liver biopsies were at least 20 mm long and/or contained more than 11 complete portal tracts.
Time between index test and reference standard	Not reported
Target condition	Cirrhosis
Prevalence of cirrhosis according to reference standard	9/88 (10.2%)
Results: Fibroscan AUC (95% CI): 0.907 (SE 0.034) Optimal cut-off threshold (if calculated): Not reported Threshold: Sensitivity and specificity values only given from ROC curve and threshold not reported Sensitivity: 100 (threshold not reported) Specificity: 76.3 (threshold not reported) PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: The accuracy of the Fibrometer NAFLD score and the NAFLD fibrosis score developed by Angulo. Any complications associated with tests reported: Not reported.	
General limitations according to QUADAS II: Consecutive or random recruitment not reported. Unclear if results of reference standard were interpreted without knowledge of the index test results or clinical data. Subjects with unreliable transient elastography measurements not included in the analysis.	

Study	Aykut 2014[10]
Liver biopsies could be <25 mm.	

Study	BORRONI 2006[13]
Study type	Retrospective analysis of chart and liver biopsy
Number of studies (number of participants). Recruitment period.	1 study (n=232 consecutive patients, 4 excluded due to liver biopsy <6 portal fields, final analysis n=228). Recruitment between 1999 and 2002.
Countries and Settings	Italy, General Hospital
Funding	No external funding
Age, gender, ethnicity	Age, mean (SEM): 42.4(0.9); male/female: 166/62; ethnicity: not reported; ALT (U/l): 117(7); duration of infection, mean (SEM): 5.6(0.4); genotype 1: 53.4%
Patient characteristics	Population: Chronic hepatitis C infection but no clinical evidence of cirrhosis. Inclusion: The diagnosis of chronic HCV infection was based on persistently high serum aminotransferase levels for at least 6 months and a positive polymerase chain reaction assay of HCV-RNA. Active IVDU were included in the study only after a period of at least 6 months of abstinence. Exclusion: (i) a previous biopsy-based diagnosis of cirrhosis; (ii) the presence of clinical (ascites, gastroesophageal varices, hepatic encephalopathy, prominent abdominal venous collaterals, spider angiomata) or ultrasonographic signs of cirrhosis (splenomegaly, liver surface nodularity); (iii) concomitant causes of liver disease diagnosed by means of standard clinical, serological and biochemical criteria; (iv) HIV-Ab positivity; (v) alcohol intake of >20 g/day during the previous 6 months; (vi) previous anti-viral treatment; (vii) any other conditions that may affect AST or platelet count.
Index test (including threshold and whether threshold pre-specified)	APRI: AST to Platelet Ratio Index (APRI)=AST (UNL)/Platelet count (109=L) x 100 (optimal cut-off ≥2, not pre-specified, so sensitivity and specificity maximal) AST/ALT ratio: AST (U/L)/ALT(U/L) (optimal cut-off ≥1, not pre-specified, so sensitivity and specificity maximal)
Reference standard	Liver biopsy (Knodell F4): The biopsies were performed under ultrasound guidance using 16-gauge needles and the lateral transcostal approach. Only samples with a length >20 mm analysed (average not reported) and 4 patients excluded as biopsy <6 portal fields. The histological sections were assessed by a single experienced pathologist (M. R.) blinded to the patients' clinical and laboratory characteristics; several sections of each specimen were evaluated in order to minimize variability.
Time between index test and	Undergone serum markers during the 3 months preceding liver biopsy.

Study	BORRONI 2006[13]
reference standard	
Prevalence of cirrhosis according to reference standard	30/228 (13.2%)
Target condition	Cirrhosis
Results: APRI AUC (95% CI): 0.86 (0.79–0.93) Optimal cut-off threshold (if calculated): ≥2 Threshold: ≥2 (optimal) Sensitivity: 43.0 Specificity: 94.0 PPV: 54.0 NPV: 92.0 +ve/-ve likelihood ratios: 7.2/0.6 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: AST/ALT ratio AUC (95% CI): 0.76 (0.68–0.84) Optimal cut-off threshold (if calculated): ≥1 Threshold: ≥1 (optimal) Sensitivity: 30.0 Specificity: 97.0 PPV: 57.0 NPV: 90.0 +ve/-ve likelihood ratios: 10/0.7 TP: Not reported FP: Not reported	

Study	BORRONI 2006[13]
FN: Not reported TN: Not reported Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Up to 3 months between index test and reference standard. Retrospective chart analysis. Liver biopsy sample <25 mm and 10 portal tracts.	

Study	BOTA 2011A[14]
Study type	Retrospective cohort
Number of studies (number of participants). Recruitment period.	1 study (n=212 patients). Recruitment between January 2008 and March 2010.
Countries and Settings	Romania, University Hospital
Funding	None declared
Age, gender, ethnicity	Age, mean (SD): not reported; male/female: not reported; ethnicity: not reported; ALT (U/l): not reported.
Patient characteristics	Population: Chronic hepatitis C infection Inclusion: Anti-HCV positive for at least 6 months and had detectable levels of HCV-RNA by RT-PCR Exclusion: Not reported
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan, Echosens, France), (cut-off 13.3 kPa, not-prespecified, from previous studies): 10 valid TE measurements, included only liver stiffness (LS) measurements with a success rate (the ratio of the number of successful acquisitions over the total number of acquisitions) of at least 60% and an interquartile range (IQR) lower than 30%. APRI: APRI score=[(AST/upper limit NV AST) ×100]/number of platelets (10^9/l). Cut-off ≥1, not-prespecified, from previous

Study	BOTA 2011A[14]
	studies. FIB-4: FIB-4 score=[age (years)] × AST (U/L)]/[number of platelets (10^9/L)] × ALT (U/L)½].
Reference standard	Liver biopsy (METAVIR F4): Echo-assisted LB was performed in all patients by using modified Menghini needles (1.4 and 1.6 mm in diameter). Only LB fragments including at least 8 portal tracts were included (average 3.35(0.9) cm). The LBs were assessed by a senior pathologist blinded to the results of the LS measurements.
Time between index test and reference standard	Single hospital visit
Prevalence of cirrhosis according to reference standard	30/212 (14.2%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.977 (CI not reported) Optimal cut-off threshold (if calculated): Not reported Threshold: 13.3 kPa (not pre-specified, from previous studies) Sensitivity: 93.3 Specificity: 97.2 PPV: 84.8 NPV: 98.8 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.879 (CI not reported) Optimal cut-off threshold (if calculated): Not reported Threshold: ≥1 (not pre-specified, from previous studies)	

Study	BOTA 2011A[14]
Sensitivity: 80.0 Specificity: 74.1 PPV: 33.8 NPV: 95.7 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: FIB-4 Not reported Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random selection not reported. Exclusions not reported. Liver biopsy sample <10 portal tracts.	

Study	BOTA 2015[15]
Study type	Retrospective cohort
Number of studies (number of participants). Recruitment period.	1 study (n=132 patients, 117 included in final analysis due to unreliable ARFI measurements). Recruitment between October 2009 to April 2013.
Countries and Settings	University Hospital, Romania
Funding	University Young Researchers Grant
Age, gender, ethnicity	Age, mean (range): 53 (21–65); male/female: 45/87; ethnicity: not reported; ALT (U/l): 1.5 (0.5–8)

Study	BOTA 2015[15]
Patient characteristics	Population: Chronic hepatitis C infection Inclusion: diagnosis of chronic infection with hepatitis C virus with positive serum anti-HCV antibodies for at least 6 months and detectable hepatitis C virus RNA in serum, by real-time polymerase chain reaction (PCR ARN-HCV). Exclusion: co-infection with hepatitis B or HIV; liver focal liver lesions or ascites on abdominal ultrasound examination.
Index test (including threshold and whether threshold pre-specified)	ARFI (pre-published cut-off 1.87 m/s): performed in all patients, in fasting condition, with a Siemens Acuson S2000TM ultrasound system using Virtual Touch Tissue Quantification application (Siemens AG, Erlangen, Germany) with a 4CI transducer. Scanning was performed between the ribs with the patient in supine position, in the right liver lobe (segment V/VIII). 10 valid LS measurements performed in the same place in the right liver lobe and a median value was calculated, the result being measured in m/s. If the measurement was not valid, "x.xx" was displayed on the screen. Reliable LS measurements were defined as median value of 10 valid measurements with an interquartile range interval (IQR) <30% and a success rate ≥60%. Transient elastography (pre-published cut-off 15.3 kPa): Transient Elastography was performed using a Fibro-Scan® device (EchoSens, Paris, France) (standard Mprobe) and was available in 123/132 patients (93.1%). In each patient aimed for 10 valid TE measurements using the standard M-probe. The LS measurements were performed under fasting conditions, in supine position, by intercostal approach, with the right arm in maximum abduction; then a median value was calculated and the results were expressed in kiloPascals (kPa). Reliable measurements were defined as: median value of 10 valid LS measurements with IQR <30% and SR ≥ 60%.
Reference standard	Liver biopsy (METAVIR F4): all liver specimens were at least 2 cm long. The biopsy fragment's length was evaluated by the physician who performed the procedure. Assessed by a senior pathologist, blinded to the results of ARFI measurements. Length of LB specimen 3.5 (2–6) cm, number of portal tracts 26.9 ± 10.1.
Time between index test and reference standard	Same session
Prevalence of cirrhosis according to reference standard	14/117 (12.0%)
Target condition	Cirrhosis
Results: ARFI AUC (95% CI): not reported Optimal cut-off threshold (if calculated): N/A Threshold: 1.87 m/s (pre-published)	

Study	BOTA 2015[15]
Sensitivity: not reported Specificity: not reported PPV: not reported NPV: 97.8% +ve/-ve likelihood ratios: Not reported TP: 12 FP: 17 FN: 2 TN: 86 Transient elastography results only reported for the FPs by ARFI. Any complications associated with tests reported: Not reported.	
General limitations according to QUADAS II: Consecutive or random selection not reported. Some liver biopsies <25 mm. Reliable LS measurements by means of ARFI elastography were obtained in 117/132 patients (87.9%), patients included in the final analysis.	

Study	CARDOSO2012 [16]
Study type	Prospective cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (hepatitis C population: n=392 consecutively recruited, n=26 excluded due to unreliable results, n=3 excluded due to unsuccessful tests; final analysis chronic hepatitis C [CHC] n=363). Recruitment between 2006 and 2008. Also recruited a hepatitis B population (n=221).
Countries and Settings	France, hospital hepatology service
Funding	Author funding or speaker for Roche, Schering Plough, Gilead, Novartis, Pharmasset, Tibotec, Boehringer, Biolex, Intermune,

Study	CARDOSO2012 [16]
	Abbott.
Age, gender, ethnicity	Age, mean (SD): 49.0(10.2); male/female: 218/145; ethnicity: 87% Caucasian, 12% Asian, 1% other; ALT (U/l): 2.5(1.2–3.1)
Patient characteristics	Population: Treatment-naïve chronic hepatitis B or chronic hepatitis C (only CHC population data extracted) Inclusion: Presence of anti-HCV antibodies and detectable serum HCV-RNA by PCR (>50IU/ml) Exclusion: Excessive alcohol consumption (>30 g/day for men, >20 g/day for women); co-infection with HIV and/or hepatitis delta virus; decompensated liver disease; HCC; previous liver surgery or transplant.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan; cut-off 12.5 kPa, according to previous studies): performed by a single experienced operator. Only patients with at least 10 valid measurements were included (IQR less than 30% median stiffness and at least 60% success rate).
Reference standard	Liver biopsy (METAVIR F4): percutaneous liver biopsy performed under ultrasound guidance using the Menghini technique with disposable 16-gauge diameter needle. A single experienced pathologist who was unaware of the clinical data evaluated all slides. Only patients with a liver biopsy length of ≥15 mm and/or at least 6 portal tracts were included.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	31/363 (8.5%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.947 (SEM 0.027) Optimal cut-off threshold (if calculated): Not reported Threshold: 12.5 (pre-specified from literature) Sensitivity: 83.9 Specificity: 94.3 PPV: 57.8 NPV: 98.4 +ve/-ve likelihood ratios: 14.65/0.17 TP: 26 FP: 19 FN: 5	

Study	CARDOSO2012 [16]
TN: 313 Other measures reported and conclusions: TE is an accurate tool for the non-invasive diagnosis of liver fibrosis in patients with chronic viral hepatitis, either related to HBV or HCV. Any complications associated with tests reported: Not reported.	
General limitations according to QUADAS II: Excluded patients with unreliable TE measurements from analysis. Liver biopsy sample <15 mm or 10 portal tracts.	

Study	CASTERA 2010A[17]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=314 CHC patients, 12 patients that had a biopsy length of less than 10 mm and/or less than 6 portal tracts were excluded, final analysis N=302, TE could not be performed in 8 patients). Recruitment period from June 2003 to February 2007.
Countries and Settings	France
Funding	Nothing to declare regarding funding from industry or conflicts of interest.
Age, gender, ethnicity, ALT (U/l):	Age: mean (SD): 52 (12) years; male/female: 176/126; ethnicity: not reported; ALT (IU/L): 106 (76)
Patient characteristics	Population: chronic hepatitis C (CHC) Inclusion: CHC was defined by detectable serum anti-HCV antibodies and HCV RNA with chronically elevated serum alanine aminotransferase (ALT) levels. Elevated ALT were defined as values above the upper limit of normal (ULN) range (50 IU/L) on at least 2 consecutive measurements over a period of 6 months. Exclusion: co-infection with hepatitis B virus (HBV) or human immunodeficiency virus (HIV), other causes of liver disease, decompensated liver disease, and liver transplantation.
Index test (including threshold and whether threshold pre-specified)	Algorithms:

Study	CASTERA 2010A[17]
	SAFE: Based on sequential use of APRI, FibroTest and liver biopsy. APRI as the initial screening test with a low and high cut-off and FibroTest as a second step. If APRI lower than low cut-off (1.0) then cirrhosis absent, if higher than 1.0 then FibroTest performed. FibroTest ≤0.48 (cirrhosis absent), FibroTest 0.49-0.74 (liver biopsy needed) and ≥0.75 (cirrhosis present). Castera: combination of TE and FibroTest. When TE and FibroTest agree no biopsy is performed whereas when they disagree, liver biopsy is needed. TE ≥12.5 and FT <0.75 (disagree), TE <12.5 and FT ≥0.75 (disagree), TE failure (disagree), TE <12.5 and FT <0.75 (agree cirrhosis absent), TE ≥12.5 and FT ≥0.75 (agree cirrhosis present). Transient elastography (Fibroscan): 10 successful measurements were performed on each patient. The success rate was calculated as the number of validated measurements divided by the total number of measurements. The median value of successful measurements was considered representative of the liver stiffness in a given patient, according to the manufacturer's recommendations IQR <30% of the median value and success rate >60%. FibroTest: Score was purchased from Biopredicitve website (www.biopredictive.com). APRI (cut-off from original publication): Formula taken from the original publication. Parameters (aspartate aminotransferase, alanine aminotransferase, c-glutamyl-transpeptidase, total bilirubin, a2-macroglobulin, apolipoprotein A1, haptoglobin and platelet count) allowing the calculation of FT and APRI were determined in the same laboratory on blood sampled the day of liver biopsy.
Reference standard	Liver biopsy (METAVIR F4): performed by senior operators using the Menghini technique with a 1.6 mm-diameter needle (Hepafix, Braun, Melsungen, Germany). All biopsy specimens were analysed by the same trained pathologist blinded to the results of non-invasive markers. Specimens with a length of less than 10 mm and/or less than 6 portal tracts were excluded (note: all biopsies would be ≥6 portal tracts even if shorter than 15 mm). The mean liver biopsy length was 20 ± 8 mm and the mean number of portal tracts was 15 ± 8. Biopsy length was greater than 15 mm in 70% of patients and greater than 25 mm in 25%.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	25%
Target condition	Cirrhosis

Study	CASTERA 2010A[17]
	Results: SAFE algorithm AUC (95% CI): 0.87 (0.84–0.90) Optimal cut-off threshold (if calculated): Not reported Threshold: as above Sensitivity: 86.4 Specificity: 89.7 PPV: 77.6 NPV: 94.1 +ve/-ve likelihood ratios: 8.4/0.15 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Castera algorithm AUC (95% CI): 0.93 (0.90–0.96) Optimal cut-off threshold (if calculated): Not reported Threshold: As above Sensitivity: 89.4 Specificity: 98.2 PPV: 95.0 NPV: 95.9 +ve/-ve likelihood ratios: 49.6/0.1 TP: Not reported FP: Not reported FN: Not reported TN: Not reported

Study	CASTERA 2010A[17]
Other measures reported and conclusions: Liver biopsy saved in 226/302 patients using SAFE algorithm and 238/302 patients using Castera algorithm. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Liver biopsy could be <25 mm or <10 portal tracts	

Study	CATANZARO 2013[18]
Study type	Prospective cohort study
Number of studies (number of participants). Recruitment period.	1 study (n=162 with chronic hepatitis C, consecutively recruited). Recruitment between January 2011 and March 2013. Also recruited 67 healthy controls to assess the diagnostic accuracy of ELF and APRI to distinguish F0 from F≥1 (note: presumed healthy control group not included in the analysis for diagnostic accuracy for F4).
Countries and Settings	Italy. Admitted to Complex Unit for liver biopsy.
Funding	None
Age, gender, ethnicity	Age, mean (SD): 55.19(9.53); male/female: 57/105; ethnicity: not reported; ALT (U/l): not reported
Patient characteristics	Population: Chronic hepatitis C Inclusion: Diagnosis of chronic hepatitis C was determined according to the positivity of anti-HCV and HCV-RNA for at least 6 months. The levels of HCV-RNA were determined by RNA extracted from serum, with reverse transcription and amplification of cDNA in real time PCR with TaqMan probes, with a sensitivity of 10 IU/ml. Exclusion: Previous history of antiviral therapy, the presence of ascites, chronic kidney failure or chronic co-infection HBV/HCV or HIV/HCV, chronic liver disease of other aetiology (HBV, non-alcoholic steatohepatitis, hemochromatosis, Wilson's disease, autoimmune hepatitis and α-1 anti-trypsin deficiency), liver failure, patients with alcohol abuse (taking more than 30 g/day of ethanol), heart failure or pregnancy, and patients with BMI >30 kg/m^2.
Index test (including threshold and whether threshold pre-specified)	ELF test (best cut-off values were determined by optimization of the Younden index). Laboratory analysis of 0.3 ml of blood taken at MedLab of Catania. Abstinence from alcohol prior to sampling was respected. Serum sample was processed through the ELF test ADVIA Centaur® (Siemens Healthcare Diagnostics Inc.), which generates a single score (ELF score) combined with doses of HA, PIIINP and TIMP-1. ELF score per ADVIA Centaur XP=2.278+0.851 ln[CHA]+0.751 ln[CPIIINP]+0.394 ln[CTIMP-1]

Study	CATANZARO 2013[18]
	APRI: details not reported
Reference standard	Liver biopsy (METAVIR F4): Percutaneous liver biopsies were performed under ultrasound guidance by a specialist, using an 18-G disposable needle. All of the liver biopsies were evaluated by expert pathologists, who were blinded to the patients' clinical histories. Only biopsies longer than 15 mm with at least 6 portal tracts were accepted.
Time between index test and reference standard	ELF test 2 weeks after liver biopsy
Prevalence of cirrhosis according to reference standard	43/162 (26.5%)
Target condition	Cirrhosis
Results: ELF AUC (95% CI): 0.94 (0.88–0.96). Adjusted AUC (DANA method): 0.90 Optimal cut-off threshold (if calculated): 9.3 Threshold: ≥9.3 (optimal) Sensitivity: 79.1 Specificity: 90.8 PPV: 75.6 NPV: 92.3 +ve/-ve likelihood ratios: LH+ 9.55 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.89 (0.83–0.93). Adjusted AUC (DANA method): 0.85 Optimal cut-off threshold (if calculated): 1.19 Threshold: ≥1.19 (optimal) Sensitivity: 74.4	

Study	CATANZARO 2013[18]
	Specificity: 87.4 PPV: 68.1 NPV: 90.4 +ve/-ve likelihood ratios: LH+ 5.9 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: ELF test more reliable than APRI score in the diagnosis of significant fibrosis and cirrhosis. It was not effective in discriminating healthy volunteers from patients with liver fibrosis. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Liver biopsy sample <25 mm and <10 portal tracts.

Study	CAVIGLIA 2013[19]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=57 with chronic hepatitis C, consecutively recruited). Recruitment period not reported.
Countries and Settings	Italy, University hospital
Funding	None to declare
Age, gender, ethnicity	Age, mean (SD): 52.5(11.9); male/female: 32/25; ethnicity: not reported; ALT (IU/l): 85(47)
Patient characteristics	Population: chronic hepatitis C Inclusion: CHC patients tested positive for anti-HVC (Ortho HCV SAVe 3.0, Raritan, USA) and HCV RNA (TaqMan, Roche,

Study	CAVIGLIA 2013[19]
	detection limit 15IU/ml). Exclusion: Patients with other aetiologies of chronic hepatitis, such as chronic hepatitis B, NASH, autoimmune hepatitis, primary biliary cirrhosis, alcoholic liver disease and haemochromatosis.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan, Echosens, Paris): (cut-off 13.8 kPa, optimal chosen to maximise sensitivity and specificity) performed on the right lobe of the liver through the intercostal spaces. Measurement depth between 25 and 65 mm below the skin surface. Liver stiffness expressed as the median value of the successful measurements. Only data with at least 10 successful measurements, success rate higher than 60% and IQR inferior to 30% considered reliable.
Reference standard	Liver biopsy (METAVIR F4): underwent liver biopsy the year preceding non-invasive assessment (from 6 to 12 months). All biopsy specimens were analysed by an experiences pathologist blinded to the clinical results of the patients. Liver specimens shorter than 20 mm were excluded from the analysis.
Time between index test and reference standard	Liver biopsy in the year preceding non-invasive liver assessment (from 6–12 months)
Prevalence of cirrhosis according to reference standard	18/57 (31.6%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.95 (0.86–0.99) Optimal cut-off threshold (if calculated): 13.8 kPa Threshold: 13.8 kPa (optimal) Sensitivity: 88.9 Specificity: 97.4 PPV: 94.1 NPV: 95.0 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	CAVIGLIA 2013[19]
Other measures reported and conclusions: Also assessed the accuracy of serum markers (hyaluronic acid, C-aminopyrine, cytokeratin). Transient elastography performed significantly better than the other tested methods. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Up to 12 months between index test and reference standard Liver biopsy sample < 25 mm	

Study	CHEN 2012[25]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=142 consecutive patients, 5 refused or were contraindicated for liver biopsy, 2 patients excluded with HCC, 2 with ALD, 1 with end stage renal disease, 2 with unreliable LSM [liver stiffness measurement] results, and 3 with inadequate specimen quality, final analysis n=127). Recruitment between November 2010 and October 2011.
Countries and Settings	Taiwan, University Hospital
Funding	Academic or Government (Department of Medical Research, China Medical University Hospital grant)
Age, gender, ethnicity	Age, mean (SD): F0-3: 51.6(1.2); F4: 62.7(1.5); male/female: 59/68; ethnicity: Taiwanese; ALT (IU/l): F0-3: 97.94(8.24); F4: 64.28(8.07).
Patient characteristics	Population: Chronic hepatitis C (referred to liver centre for liver biopsy prior to the initiation of standard care for CHC). Inclusion: Positive serum anti-HCV antibody (Abbott Laboratories, Abbott Park, Illinois, USA) for more than 6 months with the presence of serum HCV RNA (Cobas Amplicor HCV Monitor 2.0; Roche Diagnostics, New Jersey, USA). Exclusion: Interferon or nucleos(t)ide analogue treatment, exposure to hepatotoxic drugs or chemicals, primary biliary cirrhosis, primary sclerosing cholangitis, Wilson's disease, autoimmune hepatitis, alcoholic liver disease (ALD), hepatitis B virus (HBV) co-infection, human immunodeficiency virus (HIV) co-infection, liver abscess, acute hepatitis, extrahepatic cholestasis, severe haemolysis, Gilbert's syndrome with high unconjugated hyperbilirubinemia, autoimmune disorders, myeloproliferative disorders, thalassemias, schistosomiasis, major abdominal surgery, cardiac congestion, blood product transfusion within the previous 30 days, pregnancy, liver cancer, serum creatinine higher than 221 umol/L (2.5 mg/dL), hepatic encephalopathy, refractory ascites, and variceal bleeding.

Study	CHEN 2012[25]
Index test (including threshold and whether threshold pre-specified)	FibroTest (optimal cut-off value from the ROC): Serum markers including α2-macroglobulin, alanine aminotransferase (ALT), apolipoprotein A1, total bilirubin, γ-glutamyl transpeptidase (GGT) and haptoglobin were tested in the same laboratory, and results were then sent to www.biopredictive.com to determine a measure of liver fibrosis (FibroTest F score) using patented artificial intelligence algorithms. ARFI (optimal cut-off value from the ROC): ARFI technology was integrated into a conventional ultrasound system (Acuson S2000 with a Siemens 4C1 curved array, 4.00 MHz for B-mode, 2.67 MHz for push pulses and 3.08 MHz for detection pulses; Siemens Medical Solutions, Mountain View, California, USA). All ARFI stiffness measurements were performed by the same hepatologist, who was experienced in digestive system ultrasonography and blinded to the patient data. The right lobe of the liver was approached intercostally, with the patient lying in a dorsal decubitus position with both arms above the head and holding their breath during VTQ measurements. Each patient received 10 successful LSMs (failed measurements were defined as SWV= "x.xx m/s"). Reliable cases were defined as those with an IQR of less than 30% of the median of 10 successful LSMs, and a successful rate of LSMs greater than 60%. Other cases were deemed unreliable and excluded.
Reference standard	Liver biopsy (METAVIR F4): Senior hepatologists performed the percutaneous right lobe liver biopsy. All biopsy specimens were interpreted by an expert pathologist blinded to the results of LSMs and patient data. Biopsy specimens at least 15 mm in length containing at least 5 portal tracts were defined adequate (mean 21.7 [3.3] mm, range 15–32 mm).
Time between index test and reference standard	Liver biopsy within 1 hour of receiving blood tests (including those for FibroTest) and stiffness measurements
Prevalence of cirrhosis according to reference standard	18/127 (14.2%)
Target condition	Cirrhosis
Results: FibroTest AUC (95% CI): 0.757 (0.648–0.865) Optimal cut-off threshold (if calculated): Not reported Threshold: Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported	

Study	CHEN 2012[25]
	TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: ARFI AUC (95% CI): 0.831 (0.723–0.939) Optimal cut-off threshold (if calculated): 1.98 m/s Threshold: 1.98 m/s (optimal) Sensitivity: 88.9 Specificity: 79.8 PPV: 42.1 NPV: 97.8 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: 32 FN: Not reported TN: Not reported Other measures reported and conclusions: A comparison of the AUCs using ARFI and FibroTest results showed insignificant differences: p=0.341. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Liver biopsy sample < 25 mm and <10 portal tracts

Study	CHRYSANTHOS 2006[26]

Study	CHRYSANTHOS 2006[26]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (hepatitis C population: n=284 consecutively recruited). Recruitment between January 1998 and May 2004. Also recruited a hepatitis B population (n=205).
Countries and Settings	Greece, University Hospital
Funding	None reported
Age, gender, ethnicity	Age, mean (SD): 49 (15); male/female: 145/139; ethnicity: not reported; ALT (IU/l): 81 (10-647). Alcohol abuse reported in n=16 patients but had no evidence of alcohol-induced liver disease.
Patient characteristics	Population: Chronic hepatitis C Inclusion: Detectable antibodies against HCV (anti-HCV), detectable HCV RNA in serum and increased ALT activity (ALT >upper limit of normal) on at least 2 separate monthly determinations within the last 6 months. Exclusion: Patients with chronic hepatitis B virus or chronic hepatitis C virus co-infection, detectable antibodies against hepatitis delta virus (anti-HDV) or against HIV (anti-HIV), other causes of liver injury (alcohol abuse, use of known hepatotoxic drugs, autoimmune hepatitis, metabolic or cholestatic liver diseases), malignancy, or any type of antiviral or immunosuppressive therapy within the past 6 months. No patient had decompensated liver disease (history or evidence of ascites, variceal bleeding, hepatic encephalopathy or jaundice). Excluded patients with an inadequate liver biopsy length.
Index test (including threshold and whether threshold pre-specified)	APRI (2.0 and 1.0 cut-off value pre-specified from the literature): liver function tests evaluated by commercially available assays in all patients on the liver biopsy day. APRI = [(AST/ULN) / PLT (109/l)] x 100
Reference standard	Liver biopsy (Ishak F5/F6): adequate biopsy specimen with length of at least 1.5cm. All liver biopsies were evaluated blindly.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	58/284 (20.4%)
Target condition	Cirrhosis
Results: APRI AUC (95% CI): Not reported for CHC population separately Optimal cut-off threshold (if calculated): Not reported Threshold: 1.0 (pre-specified from literature) Sensitivity: 72	

Study	CHRYSANTHOS 2006[26]
	Specificity: 60 PPV: 35 NPV: 88 +ve/-ve likelihood ratios: Not reported TP: 35 FP: 64 FN: 23 TN: 162 Threshold: 2.0 (pre-specified from literature) Sensitivity: 38 Specificity: 91 PPV: 52 NPV: 85 +ve/-ve likelihood ratios: Not reported TP: 22 FP: 20 FN: 36 TN: 206 Other measures reported and conclusions: data provided for hepatitis B populations and overall viral hepatitis. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Unclear if all the liver biopsy specimens were evaluated by the same pathologist Liver biopsy sample <25 mm

Study	DE 2006[30]
Study type	Multicentre cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (HIV HCV co-infection: n=77 consecutively recruited, 5 excluded due to unsuccessful liver biopsy <7 mm, final analysis n=72). Recruitment between January 2003 and January 2005.
Countries and Settings	France
Funding	Equipment made available by Echosens (Paris, France)
Age, gender, ethnicity, ALT (U/l):	Age: Mean 42.4 (SD 5.9), gender M/F: 52/20, ethnicity: not reported, ALT: 74.4 (SD 54.7)IU/L
Patient characteristics	Population: HIV infected patients with chronic HCV Inclusion: Presence of HCV RNA and HIV antibodies in serum Exclusion: Not reported
Index test (including threshold and whether threshold pre-specified)	TE (Fibroscan, Echosens, Paris, France; optimal calculated for highest sensitivity with specificity forced no less than 90%, cut-off 11.8 kPa, and for the highest sensitivity with specificity forced no less than 95%, cut=of 14.5 kPa): tip of probe transducer placed on the skin between the ribs at the level of the right lobe of the liver. Measurement depth 25–65 mm below the skin surface. At least 5 successful measurements were performed on each patient, with the ratio of the number of successful measurements over the total number of acquisitions not lower than 30%. Platelet count (cut-off <140G/L, published cut-off) APRI index (published cut-off >2): AST X ULN x 100/platelet count (109/L) AST/ALT ratio (published cut-off >1): AST X ULN x 100/platelet count (109/L) FIB-4 (published cut-off >3.25): age x AST /(platelet count x square root ALT)
Reference standard	Liver Biopsy (METAVIR F4): Liver biopsies less than 10 portal tracts (except for cirrhosis) were excluded from histological analysis. Median length 22 mm (range 7–48 mm) All biopsy specimens were analysed by 2 experienced pathologists blinded to the clinical data and results of TE.
Time between index test and reference standard	Not reported
Prevalence of cirrhosis according to	17/72 (23.6%)

Study	DE 2006[30]
reference standard	
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.97 (0.94–1) Optimal cut-off threshold: 11.8 kPa (highest sensitivity with specificity no less than 90%), 14.5 kPa (highest sensitivity with specificity no less than 95%) Threshold: 11.8 kPa (optimal) Sensitivity: 100 (80.5–100) Specificity: 92.7 (82.4–98) PPV: 81 (58.1–94.6) NPV: 100 (93–100) +ve/-ve likelihood ratios: 13.8 (5.35–35.3)/0 TP: 17 FP: 4 FN: 0 TN: 51 Threshold: 14.5 kPa (optimal) Sensitivity: 88.2 (63.6–98.5) Specificity: 96.4 (87.5–99.6) PPV: 88.2 (63.6–98.5) NPV: 96.4 (87.5–99.6) +ve/-ve likelihood ratios: 24.3 (6.2–95.6)/0.12 (0.03–0.45) TP: 15 FP: 2 FN: 2 TN: 53 Results: Platelet count (n=64) AUC (95% CI): 0.80 (0.64–0.95)	

Study	DE 2006[30]
	Results: AST/ALT ratio (n=46) AUC (95% CI): 0.45 (0.20–0.70) Results: APRI (n=47) AUC (95% CI): 0.76 (0.59–0.92) Results: FIB-4 (n=46) AUC (95% CI): 0.73 (0.57–0.89) Other measures reported and conclusions: Area under the receiver operating characteristic curve (AUROC) of TE significantly higher than those for platelet count, AST/ALT ratio, APRI and FIB-4 Any complications associated with tests reported: Not reported
General limitations according to QUADAS II: Unclear time between index test and reference standard	

Study	Esmat 2013[34]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=164 patients). Recruitment period not reported. (Study also included 67 patients with concurrent schistosomiasis but results from these patients were not extracted).
Countries and Settings	Egypt
Funding	None reported
Age, gender, ethnicity	Age, mean (SD not reported): 40 (10.5); male/female: 111/53; ethnicity: Egyptian; ALT (U/l): not reported (but multivariate logistic regression found ALT not to be associated with agreement between biopsy and TE)
Patient characteristics	Population: Hepatitis C

Study	Esmat 2013[34]
	Inclusion: 18 to 60 years; naivety to antiviral therapy; all patients were referred for assessment prior to interferon therapy as part of the national programme for combating viral hepatitis. HCV diagnosed by seropositivity for HCV antibodies and HCV RNA by polymerase chain reaction. Exclusion: Other liver disease, decompensated liver cirrhosis, HCC, liver biopsy contraindication, those not fit for combined IFN and ribavirin treatment due to persistent haematological abnormalities and those with BMI >30
Index test (including threshold and whether threshold pre-specified)	TE (cut-off 12.5 kPa; from published literature: Castera et al): using the ultrasound TE fibroscan device (Echosens, Paris, France) with a 5-MHz ultrasound transducer probe mounted on the axis of a vibrator. Measurements were made in liver segment from 25 and 65 mm below the skin surface in a cylindrical shape 1 cm wide and 4 cm long.
Reference standard	Liver biopsy (METAVIR F4): performed on the same day as TE; performed using a semi-automatic true-cut needle (16G); specimens were analysed by an experienced pathologist blinded to the TE result. Only samples at least 15 mm and with 6 portal tracts were considered for assessment (mean of actual size of samples included was not reported).
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	18/164 (11%)
Target condition	Liver fibrosis and cirrhosis
Results: Fibroscan AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 12.5 kPa (published) Sensitivity: 72.2 Specificity: 92.5 PPV: 54.2 NPV: 96.4 +ve/-ve likelihood ratios: Not reported TP: 13 FP: 11	

Study	Esmat 2013[34]
FN: 5 TN: 135 Other measures reported and conclusions: Multivariate logistic regression, using fibrosis level as the independent variables found OR 7.12 (95%CI 2.38, 21.39, p value 0.00) for the agreement between TE and biopsy in those with liver biopsy F4. Any complications associated with tests reported: None (ARFI was feasible in all patients)	
General limitations according to QUADAS II: Consecutive or random selection not reported. Liver biopsy sample <25 mm and <10 portal tracts	

Study	Fahmy 2011[35]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (hepatitis C population: n=110). Recruitment between March 2010 to February 2011.
Countries and Settings	Italian hospital and a fibroscan centre in Cairo
Funding	Not reported
Age, gender, ethnicity, ALT (U/l):	Age, mean (SD): 41 (9); male/female: 84/26; ethnicity: not reported; ALT (IU/l): 73.61 (4.24).
Patient characteristics	Population: Newly diagnosed CHC patients Inclusion: Positive for HCVAb and HCV-RNA by polymerase chain reaction and who did not start interferon treatment Exclusion: Patients with other causes of chronic liver disease, bleeding tendency, cardiac disease, and decompensated liver disease
Index test (including threshold and whether threshold pre-specified)	TE (Fibroscan, Echosens, Paris, France; cut-off 16.5 kPa; unclear if published or optimal): the measurements were made on patients lying in dorsal decubitus with the right arm in maximal abduction. The operator, assisted by ultrasound time-motion and A-mode images, located a portion of the liver free of large vascular structures that was at least 6 cm thick. Ten validated measurements were made on each patient. Only procedures with 10 validated measurements and a success rate of at least 60% were considered reliable.

Study	Fahmy 2011[35]
Reference standard	Liver biopsy (METAVIR F4): specimens composed of core >15 mm were assessed
Time between index test and reference standard	Within 1 week
Prevalence of cirrhosis according to reference standard	22/110 (20%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.95 (CI not reported) Optimal cut-off threshold (if calculated): Not reported Threshold: 16.5 kPa; unclear if published or optimal Sensitivity: 87 Specificity: 91 PPV: 71 NPV: 96 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Also reported the diagnostic accuracy of Doppler indices (splenic artery pulsatile index, SAPI, and hepatic vein dampening index, DI). TE had a significantly higher AUROC in predicting significant fibrosis and cirrhosis than the Doppler indices ($p < 0.001$), with no significant difference found between DI and SAPI ($p > 0.05$). Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random selection not reported. Unclear whether reference standard tests results were interpreted with knowledge of other results. Liver biopsy sample <25 mm	

Study	Fernandes 2015[37]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=120, transient elastography failed in 2 patients) consecutive patients, January 2011 to July 2012
Countries and Settings	Two liver units in Brazil
Funding	Not reported
Age, gender, ethnicity, ALT (U/l):	Age, mean (SD): 53 (11.3); male/female: 41/79; ethnicity: not reported; ALT (IU/l): 84.0 (75.4)
Patient characteristics	Population: Patients with chronic hepatitis C submitted for liver biopsy to assess the indication for treatment. Inclusion: No other inclusion criteria reported. Exclusion: HIV and HBV co-infection; alcohol daily intake >20 g for women and 40 g for men; cholestasis; chronic kidney failure; right-sided heart failure; fibrogenic drug use; biopsies with < 6 portal tracts.
Index test (including threshold and whether threshold pre-specified)	ELF (cut-off 10.44 optimal): 15 ml blood sample taken and serum frozen at minus 70°C within 3 hours). PIIINP, HA and TIMP-1 measured in a random access automated clinical immunochemistry analyser that performs magnetic separation enzyme immunoassay tests (ADIVA Centaur, Siemens). ELF=2.278+0.851 ln[CHA]+0.751 ln[CPIIINP]+0.394 ln[CTIMP-1] Transient elastography (cut-off 12.5 kPa, published): performed using Fibroscan (EchoSens) using the M probe and an experienced operator blinded to the biopsy and ELF results. The median value of 10 acquisitions was considered for analysis. Only examinations with a success rate of at least 60% and an IQR/M ratio of 30% were considered for a valid measurement. If no valid measurements were achieved the examination was considered a failure.
Reference standard	Liver biopsy (METAVIR F4): ultrasound guided percutaneous liver biopsies performed under local anaesthesia. Biopsies classified by the same experienced pathologist, blinded to patient data. People with biopsies <6 portal tracts were excluded. Mean (SD) length 22 mm (1.02) and the mean number of portal tracts was 11 (4).
Time between index test and reference standard	Maximum time 3 months
Prevalence of cirrhosis according to reference standard	7%
Target condition	Cirrhosis
Results: ELF	

Study	Fernandes 2015[37]
AUC (95% CI): 0.78 (0.70–0.85) Optimal cut-off threshold (if calculated): 10.44 Threshold: 10.44 (optimal) Sensitivity: 87.5 (47.2–99.7) Specificity: 77.6 (68.8–85) PPV: 21.9 (9.1–40.3) NPV: 98.9 (93.88–100) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Transient elastography (AUC, sensitivity/specificity or 2x2 table values not reported) Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Liver biopsy sample <25 mm or <10 portal tracts	

Study	FERRAIOLI 2014[40]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=134 total population with viral hepatitis, n=102 with hepatitis C analysed separately and reported here). Consecutive patients with chronic viral hepatitis.
Countries and Settings	Infectious Diseases Department of Policlinico San Matteo, Italy
Funding	The FibroScan device was made available for this study by Echosens (Paris, France), and the iU22 ultrasound equipment was provided by Philips Medical Systems (Bothell, WA, United States)
Age, gender, ethnicity	Age, mean (SD): 45.2 (11); male/female: 82/20; ethnicity: not reported; ALT (U/l): 70 (IQR 43–127)

Study	FERRAIOLI 2014[40]
Patient characteristics	Population: Chronic viral hepatitis Inclusion: Chronic viral hepatitis Exclusion: None reported
Index test (including threshold and whether threshold pre-specified)	Transient elastography (pre-published cut-off 9.3 kPa): measurements were performed using the M probe of the FibroScan® device by two physicians with experience performing at least 50 TE procedures. During the acquisition, the patients lay in the dorsal decubitus position with the right arm in maximum abduction. The results were expressed in kilopascals (kPa). Only examinations with 10 valid measurements and an interquartile range/mean (IQR/M) <30% for values greater than 7.1 kPa were considered reliable Point shear wave elastography (pSWE; optimal cut-off): the examinations were performed using the iU22 ultrasound system (Philips Healthcare, Bothell, WA, United States) with a convex broadband probe and the ElastPQ® technique. If the amount of non-shear wave motion exceeds a threshold, the system does not display a calculation. The two raters performing the PSWE measurements had 7 years and 2 years, respectively, of experience in real-time elastography studies. They received training in PSWE measurements for two days before the study began. The examinations were performed in the right lobe of the liver through intercostal spaces, with the subject lying supine with the right arm in maximal abduction. Each rater performed 10 valid measurements, which were expressed in kPa. Measurements <1 kPa were rejected by the raters.
Reference standard	Liver biopsy (METAVIR F4): performed by three experienced physicians using a 17-gauge modified Menghini needle (Hepafix; Braun, Melsungen, Germany). The same intercostal space used for the TE and PSWE measurements was chosen for LB. The specimens were assessed on site by a single expert liver pathologist who was blind to both the TE and PSWE results. Out of the total 134 patients, specimen length described as adequate for histology in all but one patient and the mean was 2.5 (0.78) cm.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	10/102 (9.9%) (for transient elastography n=98, for pSWE n=101)
Target condition	Cirrhosis
Results: Transient elastography AUC (95% CI): 0.92 (0.85-0.97) Optimal cut-off threshold (if calculated): N/A Threshold: 9.3 kPa (pre-published)	

Study	FERRAIOLI 2014[40]
	Sensitivity: 90.0 (55.5–99.7) Specificity: 87.8 (79.2–93.7) PPV: 45.0 (23.1–78.5) NPV: 98.7 (93.2–100) +ve/-ve likelihood ratios: 7.4 (4.1-13.3)/0.1 (0.02–0.7) TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Point shear wave elastography AUC (95% CI): 0.95 (0.89–0.99) Optimal cut-off threshold (if calculated): 7.2 kPa Threshold: 7.2 kPa (optimal) Sensitivity: 90.0 (55.5–99.7) Specificity: 88.6 (80.1–94.4) PPV: 47.4 (24.4–71.1) NPV: 98.7 (93.1–100) +ve/-ve likelihood ratios: 7.9 (4.3–14.7)/0.1 (0.02–0.7) TP: Not reported FP: Not reported FN: Not reported TN: Not reported Any complications associated with tests reported: Not reported
General limitations according to QUADAS II: Liver biopsy sample <10 portal tracts and <25 mm.	

Study	FIERBINTEANU BRATICEVICI 2013[43]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=64 patients; of 93 patients with histologically proven NAFLD, 15 excluded because biopsy sample lengths were <20 mm, 14 because they were considered to have borderline NASH). Recruitment between 2007 and 2010. Note: also includes a healthy control group – presumed not to be included in calculations of diagnostic accuracy for F4).
Countries and Settings	Romania, University Hospital Bucharest
Funding	None reported
Age, gender, ethnicity	Age, mean (SD not reported): 51 (NASH) and 47 (steatosis); male/female: 28/36; ethnicity: not reported; ALT (U/l): 92 (NASH) and 67 (steatosis) (SD not reported)
Patient characteristics	Population: NAFLD Inclusion: Histologically proven NAFLD Exclusion: History of significant alcohol abuse (>20 g daily), evidence of hepatitis B and C, drug-induced liver disease or other specific liver diseases, haemochromatosis, alpha 1-antitrypsin deficiency, Wilson's disease, autoimmune diseases, congestive heart failure, biopsy <20 mm including those with biopsies less than 6 (none included had hepatic decompensation such as with ascites, variceal bleeding, or encephalopathy).
Index test (including threshold and whether threshold pre-specified)	ARFI (cut-off 1.636 m/s; determined using ROC curves with sensitivity of 91% and specificity of 92%): using the Virtual Touch Tissue Quantification mode on the Siemens Acuson S2000 ultrasound system (Siemens AG, Erlangen, Germany) with a 4-MHz transducer. Measurements were made in liver segment VIII at 1 cm depth below the liver capsule through intercostal spaces with the patient lying in decubitus dorsal position with the right hand under the head (patients were evaluated at least 8 hours after their last meal). Patients were asked to momentarily stop normal breathing while minimal scanning pressure was applied by the operator. Ten successful acquisitions were performed in each patient with results expressed at mean value of the total measurements in m/s (with values between 0.72 to 2.53 m/s). If measurements were not reliable, "X-X-X" was displayed on the screen. Liver stiffness assessed by the same physician who was blinded to the clinical and biological data.
Reference standard	Liver biopsy (Kleiner, stage 4): performed up to 6 months before ARFI; percutaneous liver biopsy was performed by senior physicians using the Menghini technique with a 1.4 mm diameter needle. All biopsy specimens were analysed by an expert pathologist with 25 years of experience who was blinded to the patient's clinical results. Only samples at least 20 mm and with 8 portal tracts were considered for assessment (average 22 mm, range 20 to 24 mm).
Time between index test and reference standard	<6 months
Prevalence of cirrhosis according to	12/64 (18.75%)

Study	FIERBINTEANU BRATICEVICI 2013[43]
reference standard	
Target condition	Liver fibrosis and cirrhosis
Results: ARFI AUC (95% CI): 0.984 (0.958–1.000) Optimal cut-off threshold (if calculated): 1.636 m/s Threshold: 1.636 m/s Sensitivity: 91.7 Specificity: 92.3 PPV: 73.33 NPV: 97.96 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Spearman's correlation coefficient between ARFI measurements and histologically determined fibrosis Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random selection not reported. Up to 6 months between index test and reference standard. Liver biopsy sample <10 portal tracts and <25 mm.	

Study	FLOREANI 2011[45]
Study type	Cross-sectional study

Study	FLOREANI 2011[45]
Number of studies (number of participants). Recruitment period.	1 study (primary biliary cirrhosis: n=120 consecutively recruited, 6 excluded because TE measurement was judged unreliable (due to an unsuccessful acquisition in 4 patients and a success rate below 60% in 2, all obese females with BMI > 34), final analysis n=114). Recruitment between January and December 2009.
Countries and Settings	Italy
Funding	Partially supported by a University grant (ex 60% fund), no conflicts declared
Age, gender, ethnicity, ALT (U/l):	Age: mean 58 (12), gender male/female: 8/96 (as reported, does not equal n=114), ethnicity: not reported, ALT: 1.1(0.9)xULN
Patient characteristics	Population: Primary biliary cirrhosis (PBC) Inclusion: PBC was defined according to the EASL 2009 guidelines; 112 patients (93.3%) had anti-mitochondrial antibody positivity of at least 1:40, whilst 8 had an antinuclear antibody positivity of at least 1:160, fulfilling the criteria for a diagnosis of AMA-negative PBC. Exclusion: Ascites, hepatocellular carcinoma, severe obesity (BMI > 40), hepatitis B or C virus infection, overlap syndrome with autoimmune hepatitis or primary sclerosing cholangitis, a history of alcohol abuse, and any other causes of liver injuries other than PBC.
Index test (including threshold and whether threshold pre-specified)	TE (Fibroscan, Echosens, Paris, France; optimal cut-off obtained analysing the AUROC at the maximum of total sensitivity and specificity): the same dedicated operator took all the measurements, obtained in the right lobe of the liver through the intercostal spaces and the median depth of measurement was 55 mm. Ten validated measurements were obtained for each patient and the minimum success rate (the ratio of successful acquisition to total acquisitions) was calculated to be 60%. The final LS result was the median of the 10 valid measurements. APRI (optimal cut-off obtained analysing the AUROC at the maximum of total sensitivity and specificity): aspartate transaminase (×upper limit of normal)/platelet count (109/L) FIB-4 (optimal cut-off obtained analysing the AUROC at the maximum of total sensitivity and specificity): age (years) × aspartate transaminase (IU/L)/(platelet count (109/L) × alanine transaminase (IU/L)) AST/ALT ratio (optimal cut-off obtained analysing the AUROC at the maximum of total sensitivity and specificity): combination of TE with each marker.

Study	FLOREANI 2011[45]
Reference standard	Liver biopsy (METAVIR F4): All specimens were analysed independently by 2 experienced pathologists blinded to patients' FibroScan results and clinical details. The length of each LB specimen and the number of fragments were recorded and only ones with a minimum length of 14 mm and including at least 10–15 portal space were considered.
Time between index test and reference standard	Within 6 months (80% within the same month)
Prevalence of cirrhosis according to reference standard	17/114 (14.9%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.99 (0.94–1) Optimal cut-off threshold: 11.4 Threshold (11.4 optimal): Sensitivity: 99 Specificity: 94 PPV: 77 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.84 (0.74–0.97) Results: FIB-4 AUC (95% CI): 0.74 (0.58–0.88)	

Study	FLOREANI 2011[45]
Results: AST/ALT ratio AUC (95% CI): 0.58 (0.42–0.74) Results: Fibroscan + APRI AUC (95% CI): 0.99 (0.94–1) Results: Fibroscan + FIB-4 AUC (95% CI): 0.99 (0.94–1) Results: Fibroscan + AST/ALT ratio AUC (95% CI): 0.99 (0.94–1) Other measures reported and conclusions: Correlation between liver stiffness and Mayo score prognostic index Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Time between index test and reference standard up to 6 months	

Study	FRIEDRICH-RUST 2010[47]
Study type	Retrospective cohort study
Number of studies (number of participants). Recruitment period.	1 study (n=74 patients with serum available dated around the time of the FibroTest of patients with chronic liver disease, who received a liver biopsy, transient elastography and FibroTest). September 2005 to June 2008. Only n=36 included here (HCV population)
Countries and Settings	University Hospital, Germany
Funding	None
Age, gender, ethnicity, ALT (U/l):	Not reported for HCV population alone

Study	FRIEDRICH-RUST 2010[47]
Patient characteristics	Population: Chronic liver disease (HCV, HVB, PBC) Inclusion: Serum available dated around the time of the FibroTest of patients with chronic liver disease, who received a liver biopsy, transient elastography and FibroTest Exclusion: Not reported
Index test (including threshold and whether threshold pre-specified)	FibroTest (pre-published cut-off): Computed on the Biopredictive website http://www.biopredictive.com. ELF test (pre-published cut-off): Serum samples were analysed for levels of tissue inhibitor of matrix metalloproteinase 1 (TIMP-1), hyaluronic acid (HA), and amino-terminal propeptide of type III collagen (P3NP) using the proprietary assays developed for ELF test by Siemens Healthcare Diagnostics Inc. (Tarrytown, New York USA). TE (Fibroscan, Echosens, Paris, France; pre-published cut-off): The examination was performed on the right lobe of the liver through the intercostal space. After the area of measurement was located, the examiner pressed the button of the probe to start the acquisition. The measurement depth was between 25 and 65 mm. As suggested by the manufacturer, 10 successful acquisitions were performed on each patient. Only TE results obtained with 10 valid measurements with a success rate of at least 60% and an IQR range ≤30% were considered reliable. Blood parameters were determined after overnight fasting in the same laboratory on the same day as transient elastography in all patients.
Reference standard	Liver biopsy (METAVIR): All biopsy specimens were analysed by an experienced pathologist blinded to the clinical results of the patients. The biopsies were judged as adequate if the number of portal tracts was at least 6 and the length of liver biopsy at least 1 cm. The mean length of the included liver biopsies was 22.3 ± 9.3 mm (median 20 mm, range 10–54 mm).
Time between index test and reference standard	Up to 12 months
Prevalence of cirrhosis according to reference standard	11/74 (not reported for HCV population alone)
Target condition	Cirrhosis
Results: FibroTest AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 0.73 (pre-published)	

Study	FRIEDRICH-RUST 2010[47]
Sensitivity: 67 Specificity: 81 PPV: 54 NPV: 88 +ve/-ve likelihood ratios: 3.6/0.41 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: ELF AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 10.31 (pre-published) Sensitivity: 89 Specificity: 63 PPV: 44 NPV: 94 +ve/-ve likelihood ratios: 2.4/0.18 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Fibroscan AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 12.5 (pre-published) Sensitivity: 78	

Study	FRIEDRICH-RUST 2010[47]
	Specificity: 84 PPV: 64 NPV: 91 +ve/-ve likelihood ratios: 4.86/0.27 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: AUROC for mixed aetiologies and for HBV and PBC separately (for the latter, measured against the Ludwig scoring system)
	General limitations according to QUADAS II: Retrospective analysis of samples Time period between index test and reference standard up to 12 months Size of liver biopsy <6 portal tracts

Study	FRIEDRICH-RUST 2010A[46]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=50 consecutive patients with NAFLD or NASH. Recruitment period August 2008 to November 2009.
Countries and Settings	Germany
Funding	XL probe provided by Echosens. No financial support.
Age, gender, ethnicity	Age, mean (SD): 44 (15), range 21–71 years; male/female: 27/23; ethnicity: not reported; ALT (IU/l): 73 (45); BMI: 29 (5.5), range 20–43 kg/m^2

Study	FRIEDRICH-RUST 2010A[46]
Patient characteristics	Population: NAFLD or NASH Inclusion: Diagnosis of NAFLD or NASH made histologically by liver biopsy. Exclusion: Men with alcohol consumption more than 30 g/week and women with alcohol consumption more than 20 g/week. Other causes of liver disease (positive hepatitis B surface antigen or anti-hepatitis C virus antibody, positive auto-antibodies) or histological evidence of other chronic liver diseases.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (FibroScan using standard M probe and using the XL probe): distance between the skin and the liver capsule at the site of TE was measured using conventional ultrasound. Performed on the right lobe of the liver through intercostal spaces. Ten successful acquisitions performed on each patient using each probe. Only results with 10 valid measurements, with a success rate of at least 60% and an IQR≤30% of the median were considered reliable. Study aims to compare the M and XL probe in the same patients. Note: The Fibroscan XL probe has been designed specifically for use in obese patients by utilisation of a lower frequency and more sensitive ultrasonic transducer, a deeper focal length, a larger vibration amplitude and a greater depth of measurement below the skin surface.
Reference standard	Liver biopsy (Kleiner F4): All specimens analysed by an experienced pathologist who was blinded to the clinical results. The biopsies were judged to be accurate if the number of portal tracts was at least 6 and the length of the biopsy at least 1cm. Mean length 21.5 (8.0) mm, median 20 mm, range 10–40 mm.
Time between index test and reference standard	Up to 18 months (median 5.5 months, mean 7.9 (6.2) months, range 0–18)
Prevalence of cirrhosis according to reference standard	3/50 (6%)
Target condition	Cirrhosis
Results: Fibroscan M probe AUC (95% CI): 0.91 (0.75-1.00) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported	

Study	FRIEDRICH-RUST 2010A[46]
NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Fibroscan XL probe AUC (95% CI): 0.95 (0.85–1.00) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Number of valid measurements significantly higher for the XL probe than the M probe. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Time between reference standard and index test up to 18 months Size of liver biopsy <6 portal tracts	

Study	FUJII 2009[50]
Study type	Unclear
Number of studies (number of participants). Recruitment period.	n=50 patients with NASH (also 100 patients with HCV but liver biopsy fibrosis scoring system does not match reference standard for HCV, Desmet et al (Scheuer classification)). Recruitment period 1998–2007.
Countries and Settings	Osaka City University Hospital
Funding	Not reported
Age, gender, ethnicity	Age, mean (SD): 55.8 (15.2); male/female: 13/37; ethnicity: presumed Japanese; ALT (IU/l): 106 (24–368)
Patient characteristics	Population: NASH Inclusion: Diagnosis of NASH based on histological features of steatohepatitis Exclusion: Clinically significant alcohol consumption (20 g/day), and other identifiable causes of liver disease including drug-induced hepatotoxicity, infection with hepatitis B or C virus, autoimmune diseases, Wilson's disease, haemochromatosis, and α1-antitrypsin deficiency.
Index test (including threshold and whether threshold pre-specified)	AAR: AST/ALT APRI: [(AST/ULN) / platelet count (x109/l] x 100 AST, ALT, alkaline phosphatase, total bilirubin, total cholesterol, triglycerides, plasma glucose, prothrombin time and platelet count were routinely determined by standard procedures within 4 weeks of biopsy.
Reference standard	Liver biopsy (Brunt F4 for NASH patients): obtained by ultrasound guided biopsy using a 15-guage Tru-cut needle (Hakko, Nagona, Japan). All specimens fulfilled the criteria for size as suggested by Janiec et al. (>1 cm with >10 portal tracts). Histological diagnosis was performed.
Time between index test and reference standard	Within 4 weeks
Prevalence of cirrhosis according to reference standard	9/50 (18%)
Target condition	Cirrhosis
Results: AAR AUC (95% CI): 0.813 (0.674–0.952) Optimal cut-off threshold (if calculated): Not reported Results: APRI	

Study	FUJII 2009[50]
AUC (95% CI): 0.786 (0.625–0.947) Optimal cut-off threshold (if calculated): Not reported Other measures reported and conclusions: AP index, CDS, HALT-C score. Sensitivity and specificity values only reported for CDS and HALT-C score. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random recruitment not reported Unclear if reference standard results interpreted without knowledge of the index test results	

Study	GAIA 2011[51]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	290 initially enrolled 21 excluded due to unsuccessful liver stiffness measurements 10 excluded due to inadequate liver biopsy specimens 259 included (77 HCV, 70 HCB, 72 NAFLD, 40 controls) January 2007–March 2009
Countries and Settings	San Giovanni Battista Hospital, Gastroenterology, Italy
Funding	Not reported
Age, gender, ethnicity, ALT (U/l):	HCV: age: 46 (29–69); male/female: 42/35; ethnicity: not reported; ALT: 76 (22–324) UI/L NAFLD: age: 48 (24–65), male/female: 52/20, ethnicity: not reported, ALT: 58 (12–264)
Patient characteristics	Population: All patients with viral or metabolic chronic liver disease who underwent liver biopsy at the Hepatology Unit. Inclusion: Chronic hepatitis C was defined by detectable anti-hepatitis C virus antibodies and serum HCV RNA. Diagnosis of NAFLD was confirmed by liver biopsy in patients with abnormal liver function tests or fatty liver at ultrasound and no other

Study	GAIA 2011[51]
	known cause of liver disease. Exclusion: Patients with alcoholic liver disease (>40 g/day alcohol consumption) and patients with acute viral hepatitis were excluded. TE and biopsy performed before any therapeutic approach, including diet and antiviral therapy.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan; optimal cut-off values to maximize sensitivity, specificity, and diagnostic accuracy): Performed on the right lobe of the liver through intercostal spaces on patients lying in the dorsal decubitus position with the right arm in maximal abduction. Measurement depth was between 25 mm and 65 mm below the skin surface. TE acquisitions with abnormal vibration shape or propagation were automatically rejected by the software. The success rate was calculated as the ratio of the number of successful measurements over the total number of acquisitions. Liver stiffness was expressed as the median value of the successful measurements. Only liver stiffness data with at least 10 successful measurements, success rate higher than 60%, and inter quartile ratio inferior to 30%, were considered reliable. TE was performed by officially trained operators who were blinded to liver histology but had access to medical records of the patients. Presumed to have used appropriate probe for patient's BMI according to manufacturer's instructions (not reported).
Reference standard	Liver biopsy (METAVIR F4 for HCV; Brunt F4 for NAFLD): All specimens were analysed by an expert pathologist blinded to the results of TE but not to the clinical and biochemical data. Liver specimens shorter than 20 mm were excluded, median length of the available specimens was 25.2 mm (range 20–30.2 mm).
Time between index test and reference standard	Within 6 months
Prevalence of cirrhosis according to reference standard	HCV 13/77 (16.8%) NAFLD 9/72 (12.5%)
Target condition	Cirrhosis
Results: HCV group AUC (95% CI): 0.922 (0.86–0.985) Optimal cut-off threshold (if calculated): 11.5 kPa Threshold: 11.5 kPa (optimal) Sensitivity: 69 Specificity: 93 PPV: (given as positive predictive accuracy, PPA): 64 NPV: (given as negative predictive accuracy, PPA): 94 +ve/-ve likelihood ratios: Not reported TP: Not reported	

Study	GAIA 2011[51]
FP: Not reported FN: Not reported TN: Not reported Results: NAFLD group AUC (95% CI): 0.942 (0.881–1.003) Optimal cut-off threshold (if calculated): 10.5 kPa Threshold: 10.5 kPa (optimal) Sensitivity: 78 Specificity: 96 PPV: (given as positive predictive accuracy, PPA) 70 NPV: (given as negative predictive accuracy, PPA) 97 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Independent predictors of severe fibrosis and cirrhosis, steatosis. TE can be considered a valid support to detect fibrosis in chronic liver disease related to HCV but it should be interpreted with caution in NAFLD patients, where host or disease-related factors may modify its accuracy.	
General limitations according to QUADAS II: Time between index and reference tests up to 6 months. Excluded patients with unsuccessful liver stiffness measurements from the analysis. Length of biopsy <25 mm.	

Study	GUECHOT 2012[58]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=590 enrolled, consecutive recruitment reported previously Zarski 2012[164] (512 included in analysis, 42 had insufficient liver biopsy, 5 had previous interferon, 9 had co-infection with HBV, 5 had excessive alcohol consumption, 1 had immunosuppressant therapy, 13 incomplete data, 3 non-confirmed HCV positive status). November 2007 to July 2008.
Countries and Settings	19 academic centres in France, Fibrostar study cohort (previously reported the ELFG score and other fibrosis tests, Zarski 2012)
Funding	The French National Agency for Research on AIDS and Viral Hepatitis (ANRS).
Age, gender, ethnicity, ALT (U/l):	Age: median 50 (18–79), gender: 60% male, ethnicity: not reported, ALT: median 69 (12–594 IU/L)
Patient characteristics	Population: Untreated hepatitis C patients Inclusion: Anti-HCV antibodies positive and RNA-HCV positive Exclusion: Associated co-infection (hepatitis B or HIV), other causes of liver disease (drug hepatitis, Wilson's disease, hemochromatosis, autoimmune hepatitis, alcohol consumption >30 g/day for men and >20 g/day for women, primary biliary cirrhosis, α-1 antitrypsine deficiency), severe systemic diseases. Individuals receiving antiviral drug therapy, immunosuppressive therapy.
Index test (including threshold and whether threshold pre-specified)	ELF score (optimal cut-off calculated by maximising the sum of sensitivity plus specificity): Fasting blood samples were collected be venepuncture. The same kinds of tubes from the same lots were used for all patients (BD Vacutainer, type Z, Becton-Dickinson, Plymouth, UK). Each of the biological parameters included in the ELF score were measured in a single laboratory using serum samples immediately separated and fractioned in fractions of 0.5 ml in 1.5 ml screw cap microtubes (Sarstedt, Numbrecht, Germany). All fractions were immediately frozen and stored at -80°C until the assays were undertaken. The transport of samples from the hepatology centres to the laboratory was achieved in carbonic ice by a specialised transporter (Area Time Logisitics, Cergy Pontoise, France). All biological tests were processed blindly without knowledge of the clinical and histological data. Serum HA was assayed using a latex agglutination method that can be applied to general clinical chemistry analysers using an AU640 analyser. Serum PIIINP was assayed using a radio immunoassay and the serum TIMP-1 was assayed using an ELISA kit. ELF score was computed from the results using the simplified algorithm published by Parkes. ELF score= -7.412+[ln HA(ng/ml)x0.681]+[ln PIIINP(ng/ml)x0.775]+[ln TIMP1(ng/ml)x0.494]+10
Reference standard	Liver biopsy (METAVIR F4): Performed by 2 senior pathologists, academic experts in liver pathology, without knowledge of any clinical and biological data except that patients had chronic hepatitis C. To be considered as adequate for scoring, the

Study	GUECHOT 2012[58]
	liver biopsies had to measure at least 15 mm and/or contain at least 11 portal tracts except for cirrhosis for which no limitation was required. Mean 25.1 (8.8) mm and longer than 25 mm in 40.2%. In case of discrepancies, slides were simultaneously reviewed by 2 pathologists using a multi-pipe microscope in order to reach a consensus.
Time between index test and reference standard	Within 2 months
Prevalence of cirrhosis according to reference standard	76/512 (14.8%)
Target condition	Cirrhosis
Results: ELF score AUC (95% CI): 0.85 (0.81–0.90) Optimal cut-off threshold (if calculated): 9.35 Threshold: 9.35 (optimal) Sensitivity: 0.83 (0.79–0.66) Specificity: 0.75 (0.64–0.84) PPV: 0.44 NPV: 0.95 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Youden index 0.59 Other measures reported and conclusions: Obuchowski measures for ELF versus ELFG and FibroTest. This study confirms the ELF score performance as an index to predict liver fibrosis or cirrhosis in chronic HCV. The ELF test, using validated reagents, could be added to the health authorities approved non-invasive tests in assessing fibrosis as surrogate to liver biopsy. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Not all patients included in the analysis and length of time between reference standard and index test up to 2 months. Liver biopsy size <25 mm.	

Study	Halfon 2007[60]
Study type	Retrospective cohort
Number of studies (number of participants). Recruitment period.	n=356. Recruitment from October 1994 to March 2004 in Tours centre and from September 2002 to January 2004 in Provence area.
Countries and Settings	University Hospital in Tours, and 5 units (2 University Hospital, 2 public hospitals, 1 private clinic) from Provence-Cote d'Azur area, France
Funding	Not reported
Age, gender, ethnicity, ALT (U/l):	Age: 44.9±12.9; male: 189 (53%); ethnicity: not reported; ALT (IU/L): 76.5±66.2
Patient characteristics	Population: Chronic viral hepatitis C Inclusion: Positive HCV-RNA in the serum and a liver biopsy and an alcohol consumption <30 g/day for the past 5 years Exclusion: Liver specimen <15 mm or other cause of liver disease or complicated cirrhosis or were given putative anti-fibrotic treatment (for example interferon or sartan) in the past 6 months
Index test (including threshold and whether threshold pre-specified)	FibroTest: Cut-off of regression score was determined according to the highest Youden index (Se + Spe 1) APRI: Cut-off of regression score was determined according to the highest Youden index (Se + Spe 1) Blood markers were measured either on fresh blood or frozen sample of serum stored at -20C. Sampling was performed for routine diagnostic aim within 1 week of liver biopsy.
Reference standard	Liver biopsy (METAVIR F4): Patients were not included if they had liver specimen <15 mm (average 22.0 ± 7.1). Fibrosis was staged by 2 independent expert pathologists. Observers were blinded for patient characteristics. When the pathologists did not agree, the specimens were re-examined under a double-headed microscope to analyse discrepancies and reach a consensus.
Time between index test and reference standard	Within 1 week
Prevalence of cirrhosis according to reference standard	13/356 (4%)
Target condition	Cirrhosis
Results: FibroTest AUC (95% CI): 0.86 (0.82; 0.89)	

Study	Halfon 2007[60]
Optimal cut-off threshold (if calculated): 0.56 Threshold: 0.56 (optimal) Sensitivity: 85 Specificity: 74 PPV: 11 NPV: 99 +ve/-ve likelihood ratios: 3.19/0.21 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.92 (0.88; 0.94) Optimal cut-off threshold (if calculated): 0.83 Threshold: 0.83 (optimal) Sensitivity: 100 Specificity: 83 PPV: 18 NPV: 100 +ve/-ve likelihood ratios: 5.81/0.00 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Fibrometer and hepascore reported. Subgroup analysis by centre and by biopsy size (≥21 mm and <21 mm). Any complications associated with tests reported: Not reported	

Study	Halfon 2007[60]
General limitations according to QUADAS II: Consecutive or random recruitment not reported. Retrospective recruitment. Liver biopsy size <25 mm.	

Study	Janssens 2010[63]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=255 patients admitted, 16 excluded due to unsuccessful TE due to obesity or ascites, 167 patients excluded as were F0-2 according to TE value, 72 patients had severe fibrosis according to TE but 21 refused biopsy and biopsy not possible in 2 patients. Final analysis n=49) Recruitment between January 1, 2006 and February 29, 2008.
Countries and Settings	University hospital, Brussels, Belgium
Funding	No conflict of interest or financial support to be declared
Age, gender, ethnicity	Age, median (range): 53 (29–73) years; male/female: 34/15; ethnicity: ; ALT (U/l): 62 (36.6). Six patients had diabetes mellitus, 1 patient was hepatitis B surface antigen positive, and 1 patient was hepatitis C antibody and HCV-RNA positive but liver biopsies did not show signs of chronic viral hepatitis and therefore it was decided to keep them in the study.
Patient characteristics	Population: Actively drinking alcoholic patients admitted for detoxification and rehabilitation during a 2-week hospitalisation period, separated by 1 outpatient week. Lab tests and TE performed during the first week. Those with a suspicion of severe fibrosis (TE ≥9.5 kPa) underwent liver biopsy during the second hospitalisation week. Inclusion: All patients drank actively until the day of their first admission. Self-reported minimum daily alcohol intake was 7 standard drinks (70 g of alcohol). Exclusion: Patients who desired not to be rehospitalised for a second week. Patients who declined TE or had unsuccessful TE (as it was a prerequisite for liver biopsy). Patients who refused liver biopsy.
Index test (including threshold and whether threshold pre-specified)	APRI (pre-published cut off value of 2.0): Calculated from routine lab blood tests collected at admission. APRI calculated as follows: AST/ULN x 100/platelet count (109/L). Transient elastography (Fibroscan, optimal cut-offs for population reported, also used validated cut-off in HCV population but results not reported): Performed by an experienced examiner who was unaware of the biological, radiological and clinical

Study	Janssens 2010[63]
	data. Final result reported as the median value of at least 10 validated measurements with a minimum success rate of 60% and an IQR <30%.
Reference standard	Liver biopsy METAVIR (F4): Performed through the right jugular vein approach using a Ross-modified Colapinto catheter needed with a diameter of 1.5 mm (Cook, Denmark). All specimens analysed by an experienced liver pathologist blinded to the biological, radiological and clinical data. Liver biopsy specimen of at least 15 mm containing a minimum of 6 portal tracts were considered suitable for fibrosis staging, or when obvious regenerating nodules were present allowing the unequivocal diagnosis of cirrhosis.
Time between index test and reference standard	Within 3 weeks
Prevalence of cirrhosis according to reference standard	20/49 (40.8%) for TE. 11/28 (39.3%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.864 (CI not reported) Optimal cut-off threshold (if calculated): ranged between 19.6 and 23.5 kPa Threshold: 19.6 kPa Sensitivity: 80 Specificity: 76 PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 21.1 kPa Sensitivity: 75 Specificity: 80 PPV: Not reported	

Study	Janssens 2010[63]
NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 23.5 kPa Sensitivity: 65 Specificity: 83 PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI (n=48) AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 2.0 Sensitivity: 40 Specificity: 61 PPV: 42 NPV: 59 +ve/-ve likelihood ratios: Not reported TP: 8 FP: 11 FN: 12	

Study	Janssens 2010[63]
TN: 17 Other measures reported and conclusions: Forns score. Evaluation of factors that influence the liver stiffness measurement. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Random or consecutive recruitment not reported. Liver biopsy samples <25 mm Indirectness: Only patients with severe fibrosis (transient elastography ≥9.5 kPa) underwent liver biopsy.	

Study	KAYADIBI 2014[66]
Study type	Retrospective cohort study
Number of studies (number of participants). Recruitment period.	1 study (n=214; 202 with sufficient data to complete) Recruitment between 2008–2010
Countries and Settings	Department of Gastroenterohepatology of Haydarpasa Numune Training Hospital, Istanbul
Funding	Not reported
Age, gender, ethnicity	Age, mean (range): 52 (42–59); male/female: 61% male; ethnicity: presumed from Istanbul; ALT (U/l): not reported for whole group, only grouped by presence or absence of cirrhosis.
Patient characteristics	Population: Hepatitis C patients who underwent liver biopsy Inclusion: Anti-HCV and HCV RNA positivity Exclusion: Co-infection with HIV, hepatitis B, hepatitis D, use of steroids, NSAIDs, antiviral therapy, other liver disorders
Index test (including threshold and whether threshold pre-specified)	FIB-4=Age (years) x AST (U/L) / [platelet count (109L) x ALT1/2 (U/L)] APRI=([AST/ULN]/platelet count [109L]) x100 AST/ALT ratio (AAR) AST ALT

Study	KAYADIBI 2014[66]
	Platelet count: Performed by the blood count analyser All measured by commercial assays using the fasting serum sample results.
Reference standard	Liver biopsy METAVIR (F4) obtained with an 18-gauge needle and assessed by a single senior pathologist blinded to the clinical history and lab results. Samples ≥25 mm, ≥8 portal tracts were used.
Time between index test and reference standard	1 week
Prevalence of cirrhosis according to reference standard	47/202 (23%)
Target condition	Cirrhosis
Results: ALT: AUC (95% CI): 0.626 (0.534–0.717) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported AST: AUC (95% CI): 0.752 (0.671–0.832) Optimal cut-off threshold (if calculated): Not reported Threshold:	

Study	KAYADIBI 2014[66]
Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Platelet count: AUC (95% CI): 0.827 (0.745–0.908) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported FIB-4: AUC (95% CI): 0.853 (0.784–0.921) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported	

Study	KAYADIBI 2014[66]
Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported APRI: AUC (95% CI): 0.847 (0.776–0.919) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported AST/ALT ratio: AUC (95% CI): 0.610 (0.510–0.709) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported	

Study	KAYADIBI 2014[66]
	PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Multivariate regression analysis revealed that fibrosis index was the best predictor of cirrhosis, potentially decreasing the need for biopsy in 83% of patients, and Forns index, platelet count and APRI were statistically significant predictors of cirrhosis. Sensitivity and specificity values at a given cut-off threshold only provided for the created fibrosis index. Any complications associated with tests reported: Not reported
General limitations according to QUADAS II: Random or consecutive recruitment not reported.	

Study	KETTANEH 2007[67]
Study type	Prospective multicentre
Number of studies (number of participants). Recruitment period.	935 consecutive HCV patients enrolled 79 inadequate FibroScan measurements 292 biopsy length <15 mm 54 biopsy length unknown 560 patients included in analysis November 2002–April 2005

Study	KETTANEH 2007[67]
Countries and Settings	Multiple centres in France: Hopital Saint-Antoine, Paris; Hopital Beaujon, Paris; Hopital Henri Mondor, Paris; Hopital Jean Verdier, Paris; Hopital Haut-Leveque, Bordeaux
Funding	No funding received from any source
Age, gender, ethnicity, ALT (U/l):	Mean age: 24.5±4.0; gender: 62.3% male; ethnicity: not reported; ALT: 93±80 IU/l
Patient characteristics	Population: Chronic HCV patients Inclusion: HCV defined by detectable serum anti-HCV antibodies and HCV RNA in subjects with chronically elevated serum alanine aminotransferase levels. Exclusion: Co-infection with HIV or HBV. Hepatocellular carcinoma.
Index test (including threshold and whether threshold pre-specified)	TE via FibroScan The tip of the probe transducer was placed on the skin, between the rib bones at the level of the right lobe of the liver where liver biopsy would be done. Once the measurement area had been located, the operator pressed the probe button to start an acquisition. The measurement depth was between 25 mm and 65 mm below the skin surface.
Reference standard	Liver biopsy was fixed in formalin and paraffin-embedded. All biopsy specimens were analysed by 1 experienced pathologist blinded to the clinical data and the results of the FibroScan. Fibrosis and necro-inflammatory activity were staged according to METAVIR. Only those with a minimal length of 15 mm were eligible as the gold standard for the prediction of cirrhosis by elastography.
Time between index test and reference standard	Not reported
Prevalence of cirrhosis according to reference standard	58/560 (10.4%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 90.7 (87.1–94.3) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported	

Study	KETTANEH 2007[67]
	Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Patient and operator characteristics associated with the success rate of liver stiffness measurements. Effect of number of valid Fibroscan shots (at least 3 versus at least 10) on outcome. Fibroscan provides a reasonable performance for the diagnosis of cirrhosis that is not influenced substantially by any other feature. More patients will benefit from this procedure with no significant loss in performance if only 5 valid shots are requested. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Time between reference standard and index test not reported. Patients with unsuccessful TE excluded from the analysis. Liver biopsies <25 mm.

Study	LACKNER 2005[74]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=211 consecutive patients with chronic hepatitis C (17 excluded due to inadequate biopsy, final analysis n=194). Between 1994 and 2004.
Countries and Settings	Medical University Graz or at the Landeskrankenhaus Hoergas, Austria
Funding	Not reported. No conflicts of interest.

Study	LACKNER 2005[74]
Age, gender, ethnicity, ALT (U/l):	Age: mean 48 (12) years; male/female: 111/83; ethnicity: not reported; ALT: 2.8 (2.0) ULN
Patient characteristics	Population: Treatment- naïve patients with chronic HCV Inclusion: Tested positive for the presence of HCV RNA using a polymerase chain reaction assay and did not suffer from additional causes of chronic liver disease as confirmed by standard clinical, serological, biochemical, and radiological criteria. Exclusion: Antiviral treatment before liver biopsy, alcohol consumption in excess of 20 g/d, and previous liver transplantation.
Index test (including threshold and whether threshold pre-specified)	AST/ALT ratio: Pre-published cut-off threshold APRI: Pre-published cut-off threshold Platelet count: Optimal cut-off from ROC Because of the introduction of the International Federation of Clinical Chemistry reference method for the determination of aminotransferase activities at 37°C, the upper limits of normal (ULN) for AST and ALT changed in the course of the study (ULN before March 2003: AST, 18 U/L; ALT, 22 U/L; after March 2003: AST, 35 U/L male or 30 U/L female, ALT, 45 U/L male or 35 U/L female). Therefore, both AST and ALT were transformed into multiples of the ULN for further analysis except for the calculation of AAR. The reference range for platelet count was 140x109/L.
Reference standard	Liver biopsy (Ishak F5-6): Biopsy specimens with at least 6 portal fields were considered representative. Histological grading performed independently by 2 pathologists. Mean biopsy length 19 (8) mm, median number of portal tracts 11 (IQR 9–16).
Time between index test and reference standard	Same day (n=96); within 1 month (n=98)
Prevalence of cirrhosis according to reference standard	32/194 (16.4%) (reported in the paper for 2 pathologists' opinions separately as 16% and 17%, however, the results in the table show that both pathologists rated 32/194 as F5-6. Results also reported as similar for the 2 pathologists, so results for all tests below were taken for pathologist 1).
Target condition	Cirrhosis
Results: AST/ALT ratio AUC (95% CI): 0.73 (0.63–0.83) Optimal cut-off threshold (if calculated): Not reported Threshold: 1.0 (pre-published) Sensitivity: 36 Specificity: 90 PPV: 41 NPV: 87	

Study	LACKNER 2005[74]
+ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.90 (0.85–0.95) Optimal cut-off threshold (if calculated): Not reported Threshold: 1.0 (pre-published) Sensitivity: 93 Specificity: 70 PPV: 38 NPV: 98 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 2.0 (pre-published) Sensitivity: 55 Specificity: 93 PPV: 59 NPV: 91 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	LACKNER 2005[74]
Results: Platelet count AUC (95% CI): 0.89 (0.83–0.94) Optimal cut-off threshold (if calculated): 150x109L Threshold: 130x109L (published) Sensitivity: 53 Specificity: 93 PPV: 59 NPV: 91 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 150x109L (optimal) Sensitivity: 77 Specificity: 88 PPV: 56 NPV: 95 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: APRI accuracy in good agreement with previous studies but AST/ALT and platelet count accuracies considerably lower than previous reports. Any complications associated with tests reported: Not reported
General limitations according to QUADAS II:

Study	LACKNER 2005[74]
Unclear if reference standard result interpreted without knowledge of clinical data or the index test results. Liver biopsy <10 portal tracts	

Study	LEROY 2014[77]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	510 patients (CHC n=255, CHB n=255)
Countries and Settings	Clinique Universitaire d'Hepato-Gastroenterologie, CHU de Grenoble, France
Funding	'Direction de la Recherche Clinique' Grenoble University Hospital
Age, gender, ethnicity, ALT (U/l):	Age: 46.5±12.1, gender: 56.9% male, ethnicity: not reported, ALT: 59.5±56.5 IU/L
Patient characteristics	Population: Consecutive naïve patients with chronic HCV addressed to the centre were considered for inclusion if they had interpretable liver biopsy and a fasting serum sample collected the same day. Inclusion: Presence of HCV RNA for at least 6 months. During the inclusion period a liver biopsy was systematically recommended and performed as part of clinical care for staging and grading liver disease. Exclusion: <18 years, HBV or HIV co-infection, hepatitis delta virus, other causes of liver disease alcohol consumption over 30 g/day, hepatocellular carcinoma, Gilbert's disease, chronic haemolysis, inflammatory syndrome, previous antiviral treatment, previous liver transplantation.
Index test (including threshold and whether threshold pre-specified)	FibroTest (optimal calculated according to Youden's Index which maximises the sum of sensitivity and specificity): Parameters were measured in fresh blood samples. Alpha-2 macroglobulin, haptoglobin and apolipoprotein A1 were measured by immunonephelometry using a BN ProsPec analyser. GGT and bilirubin were measured using a Roche modular analyser with reagents from the manufacturer and CFAS. Using laboratory values FibroTest was purchased from Biopredictive.
Reference standard	Percutaneous liver biopsy was performed by 2 senior operators using a 16G disposable needle. Tissue samples were fixed in formalin and embedded in paraffin. All specimens were analysed twice by a single senior pathologist who was unaware of biochemical markers. Liver fibrosis was evaluated according to the METAVIR system.
Time between index test and reference standard	Same day

Study	LEROY 2014[77]
Prevalence of cirrhosis according to reference standard	Not reported for HCV group 56/510 (11% in whole group)
Target condition	
Results: FibroTest AUC (95% CI): 0.87 (0.8–0.94) Optimal cut-off threshold (if calculated): 0.63 (calculated according to Youden method) Threshold 0.63 (optimal): Sensitivity: 74 Specificity: 82 PPV: 53 NPV: 96 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold 0.74 (published): Sensitivity: 59 Specificity: 91 PPV: 45 NPV: 95 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions:	

Study	LEROY 2014[77]
Steatosis, Fibrometer, Hepascore. Applicability of HCV cut-offs to HBV. Overall the diagnostic performance of blood tests is similar in hepatitis B and C. The risk of underestimating significant fibrosis and cirrhosis is greater in hepatitis B and cannot be entirely corrected by use of more stringent cut-offs. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Liver biopsy length <25mm	

Study	LUPSORPLANTON 2013[83]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=1202 consecutive CHC patients. Between May 2007 and December 2012.
Countries and Settings	Regional Institute of Gastroenterology and Hepatology, Cluj-Napoca, Romania
Funding	Part of a research project from the "Iuliu-Hatieganu" University of Medicine and Pharmacy, Cluj-Napoc.
Age, gender, ethnicity, ALT (U/l):	Age: mean 50.61 (10.84) years, range 21–85; male/female: 465/737; ethnicity: not reported; ALT: 86.16 (66.88) U/l
Patient characteristics	Population: Chronic hepatitis C (CHC) patients Inclusion: Positive serum HCV-RNA and underwent percutaneous LB for disease grading and staging Exclusion: Evidence of ascites on physical or ultrasound examination (ascites is a physical limitation of the technique because elastic waves do not propagate through fluids), co-infection with HBV and/or HIV, active infectious diseases other than HCV, severe cholestasis, right heart failure, history of alcohol consumption (>30 g/day in men and >20 g/day in women) and pregnancy.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan; optimal cut-off values were chosen to maximize the sum of sensitivity and specificity): After an overnight fast, each patient was examined in a dorsal decubitus position, with the right arm in maximum abduction. The Fibroscan transducer was placed perpendicularly to the intercostal space, in an area free of any large vascular structure. The median value of 10 successful acquisitions was recorded. We considered as representative 10 successful acquisitions, regardless of the success rate (SR) as long as 10 valid LSMs were obtained and with an IQR lower than 30% of the median value.

Study	LUPSORPLANTON 2013[83]
Reference standard	Liver biopsy (METAVIR F4): Performed using the TruCut technique with a 1.8 mm (14G) diameter automatic needle device – Biopsy Gun (Bard GMBH, Karlsruhe, Germany). Only LB specimens with more than 6 intact portal tracts were eligible for evaluation. Median size of the LB sample was 11 (8–27) mm, with a median of 11 (7–30) portal spaces.
Time between index test and reference standard	TE 1 day prior to biopsy
Prevalence of cirrhosis according to reference standard	374/1202 (31.1%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.970 (0.969–0.979) (also reports adjusted DANA AUC: 0.9774, no significant difference with AUC) Optimal cut-off threshold (if calculated): 13.2 kPa Threshold: 13.2 kPa (optimal) Sensitivity: 93.75 (90.8–96.0) Specificity: 93.31 (91.4–94.9) PPV: 86.5 NPV: 97.0 +ve/-ve likelihood ratios: 14.01/0.067 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Any complications associated with tests reported: In 27 patients (2.2%) no valid measurement was obtained. In 11.2% of cases, the SR was <60%, although 10 valid LSMs were recorded.	
General limitations according to QUADAS II: Unclear who performed fibrosis staging of biopsy and whether it was performed without knowledge of the index test result or clinical data Liver biopsy less than 10 portal tracts	

Study	MACIAS 2006[85]
Study type	Retrospective cross sectional
Number of studies (number of participants). Recruitment period.	1 study (n=357; only n=263 with adequate liver biopsy included in the analysis reported here). Liver biopsy between January 1991 and January 2005.
Countries and Settings	Southern Spain, 5 hospitals
Funding	Fondo de Investigaciones Sanitarias, Fundacio Barcelona SIDA, Fundacion para la Investigacion y la Prevencion del SIDA en Espana
Age, gender, ethnicity	Age, mean (range): 37 (34–41); male/female: 84% male; ethnicity: not reported; ALT (U/l): 80 (UI/L) (54–133)
Patient characteristics	Population: Hepatitis C and HIV co-infected Inclusion: Admitted for liver biopsy to establish prognosis and indicate therapy for chronic hepatitis C. Exclusion: Hepatitis B, other causes of liver disease (autoimmune, tumoural, biliary, vascular-associated), prior anti-HCV therapy.
Index test (including threshold and whether threshold pre-specified)	AST:ALT ratio (cut-off value 1, pre-specified from published threshold) Platelet count (cut-off value 150x109/l, pre-specified from published threshold) APRI (cut-off value 1 and 2, pre-specified from published thresholds): Calculated by assigning arbitrary scores to 3 laboratory parameters and summing them with a possible value of 0 to 11.
Reference standard	Liver biopsy (Knodell F4). A minimum liver biopsy length of 10 mm was required but only biopsies above 15 mm were included in the analysis. Specimens were immediately placed in buffer formalin. After 24 hours of fixation they were embedded in paraffin using routine methods. Histological evaluation was made on sections stained with haematoxylin-eosin and Masson's trichrome by a single pathologist who was blinded to clinical data.
Time between index test and reference standard	Within 1 month
Prevalence of cirrhosis according to reference standard	40/263 (15%)
Target condition	Cirrhosis
Results: APRI AUC (95% CI): 0.79 (0.71–0.87) Optimal cut-off threshold (if calculated): Not reported Threshold: 1 (published cut-off)	

Study	MACIAS 2006[85]
Sensitivity: 78 Specificity: 57 PPV: 24 NPV: 93 +ve/-ve likelihood ratios: Not reported TP: 31 FP: 97 FN: 9 TN: 126 Threshold: 2 (published cut-off) Sensitivity: 53 Specificity: 89 PPV: 46 NPV: 91 +ve/-ve likelihood ratios: Not reported TP: 21 FP: 25 FN: 19 TN: 198 Results: AST/ALT AUC (95% CI): 0.6 (0.5–0.69) Optimal cut-off threshold (if calculated): Not reported Threshold: 1 (published cut-off) Sensitivity: 38 Specificity: 77 PPV: 23 NPV: 87	

Study	MACIAS 2006[85]
	+ve/-ve likelihood ratios: Not reported TP: 15 FP: 51 FN:25 TN:172 Results: Platelet count AUC (95% CI): 0.79 (0.72–0.86) Optimal cut-off threshold (if calculated): Not reported Threshold: 150x109/l (published cut-off) Sensitivity: 63 Specificity: 77 (incorrectly reported in paper, calculated from 2x2 table) PPV: 33 NPV: 92 +ve/-ve likelihood ratios: Not reported TP: 25 FP: 51 FN: 15 TN: 172 Other measures reported and conclusions: Forns and Bonacini models, Saadeh model. The diagnostic accuracy of these models was lower in HIV/HCV co-infected patients than in the validation studies performed in HCV mono-infected patients, however simple fibrosis tests may render liver biopsy unnecessary in deciding anti-HCV treatment in over one-third of patients with HIV infection and chronic hepatitis C. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Not all patients included in the analysis Liver biopsy sample <25 mm

Study	MARTINEZ 2011A[86]
Study type	Cohort study
Number of studies (number of participants). Recruitment period.	n=340 August 2001–November 2007
Countries and Settings	Liver Unit, Hospital Clinic, IDIBAPS and Ciberehd, Barcelona, Spain
Funding	Not reported
Age, gender, ethnicity	Mean age=47.7 years, male/female: 217/123. Ethnicity: not reported; ALT (presented as ALT/upper limit of normal): 2.94± 2.5
Patient characteristics	Population: Chronic hepatitis C patients (established by the presence of HCV RNA using polymerase chain reaction assays) tested prior to antiviral therapy. Inclusion: Consecutive patients who underwent antiviral treatment and underwent a pretreatment liver biopsy within 6 months prior to the initiation of therapy. Exclusion: Patients with HIV, hepatitis B or other causes of chronic liver disease were not included.
Index test (including threshold and whether threshold pre-specified)	APRI, FIB-4, ELF (cut-off values as pre-published): Measured in blood samples collected on the day of antiviral treatment initiation, all according to standard cut-offs (also taken following antiviral treatment). Patient values were entered into the ELF algorithm, where the original score was simplified by removing age (J. Parkes, unpublished observation).
Reference standard	Liver biopsy (METAVIR F4): Percutaneous liver biopsies were performed under local anaesthesia and ultrasound guidance with a Tru-Cut 14 gauge needle (Angiomed, Bard, Karlsruhe, Germany) by expert radiologists. A minimum length of 10 mm and the presence of 6 portal tracts were required for diagnosis. Histological grade and stage were determined by the same pathologist, who was blinded to patient data. Liver fibrosis was considered significant (stages 2, 3 or 4) when it spread out of the portal tract. Mean biopsy length was 15 mm (range 10–30 mm) with 55% of specimens >15 mm, 16% >20 mm and 1% >25 mm. Mean number of portal tracts was 9.
Time between index test and reference standard	Within 6 months
Prevalence of cirrhosis according to reference standard	124/340 (36.4%)
Target condition	Cirrhosis
Results: APRI	

Study	MARTINEZ 2011A[86]
AUC (95% CI): 0.86 (0.82–0.90) standard threshold Optimal cut-off threshold (if calculated): Not reported Threshold: 1 Sensitivity: 82 Specificity: 74 PPV: 64 NPV: 88 +ve/-ve likelihood ratios: 3.2/0.2 Diagnostic odds ratio: 16 TP: 102 FP: 57 FN: 22 TN: 159 Threshold: 2 Sensitivity: 49 Specificity: 91 PPV: 75 NPV: 76 +ve/-ve likelihood ratios: 5.4/0.6 Diagnostic odds ratio: 9 TP: 61 FP: 20 FN: 63 TN: 196 Results: ELF AUC (95% CI) 0.82 (0.78–0.87) standard threshold Optimal cut-off threshold (if calculated): Not reported Threshold: 0.06	

Study	MARTINEZ 2011A[86]
Sensitivity: 90 Specificity: 53 PPV: 52 NPV: 90 +ve/-ve likelihood ratios: 1.9/0.2 Diagnostic odds ratio: 9.5 TP: 111 FP: 102 FN: 13 TN: 114 Threshold: 1.73 Sensitivity: 52 Specificity: 90 PPV: 76 NPV: 77 +ve/-ve likelihood ratios: 5.2/0.5 Diagnostic odds ratio: 10.4 TP: 65 FP: 21 FN: 59 TN: 195 Results: FIB-4 AUC (95% CI) 0.89 (0.85–0.92) standard threshold Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported	

Study	MARTINEZ 2011A[86]
NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Extracellular matrix tests and virological response to treatment. Simple panel markers and ELF score are accurate at identifying significant fibrosis and cirrhosis in chronic hepatitis C. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Time between reference and index tests up to 6 months. Liver biopsy <10 portal tracts	

Study	MUELLER 2010[89]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=106 patients with histologically staged ALD, 5 excluded because of invalid TE, final analysis 101 (second validation part of study – includes diagnostic accuracy of overall population, in addition to internal validation of accuracy for proposed algorithm depending on glutamic oxaloacetic transaminase [GOT] level)
Countries and Settings	Germany
Funding	The Dietmar Hopp Foundation and the Manfred Lautenschlager Foundation
Age, gender, ethnicity	Age, mean (SD): 53.6 (10.6) years; male/female: 73/28; ethnicity: not reported; ALT (IU/l): not reported
Patient characteristics	Population: Alcohol-related liver disease (ALD)

Study	MUELLER 2010[89]
	Inclusion: Patients with histologically staged ALD, a full set of blood tests and FS examination at the time of liver biopsy Exclusion: Ultrasound examination was routinely performed in addition to FS measurements to exclude extrahepatic cholestasis, liver congestion or liver tumours.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (FibroScan, using the M probe; cut-off of 12.5 kPa based on previous studies and cut-off 11.5 to give optimal sensitivity): The tip of the probe transducer was placed on the skin between the rib bones and the level of the right lobe of the liver. The measurement depth was between 25 and 65 mm below the skin surface. Ten measurements were performed with success rates of at least 60%. FS measurements with an IQR higher than 40% were excluded.
Reference standard	Liver biopsy (Kleiner F4): All biopsy specimens were analysed independently by 2 experienced pathologists blinded to the results of FS and other clinical data. Only biopsies >15 mm were included.
Time between index test and reference standard	Same time
Prevalence of cirrhosis according to reference standard	26/101 (25.7%)
Target condition	Cirrhosis
Results: AUC (95% CI): 0.921 (0.87–0.97) Optimal cut-off threshold (if calculated): 11.5 kPa (to give 100% sensitivity) Threshold: 11.5 kPa (optimal sensitivity) Sensitivity: 100 Specificity: 77 PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	MUELLER 2010[89]
	Threshold: 12.5 kPa (pre-published) Sensitivity: 96 Specificity: 80 PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Development and internal validation of an algorithm for TE in people with ALD based on subgrouping into degree of alcoholic steatohepatitis and GOT level (exclusion of patients with GOT >100U/L, but not with GOT >50U/L, increased the accuracy of TE). Any complications associated with tests reported: Not reported
General limitations according to QUADAS II: Consecutive or random recruitment not reported Liver biopsy sample <25 mm	

Study	MYERS 2012B[90]
Study type	Multicentre cross-sectional study
Number of studies (number of participants). Recruitment period.	n=276 total. 'Viral' group comprised hepatitis C and B therefore did not extract. NAFLD group=127 Recruitment period July 2009–July 2010
Countries and Settings	Four academic hospitals in Canada
Funding	Echosens, Paris

Study	MYERS 2012B[90]
Age, gender, ethnicity	Whole group data (n=276): age, mean (range): 50 (43–57); male/female: 63% male; ethnicity: not reported; ALT (IU/l): 55 (36–87)
Patient characteristics	Population: NAFLD, BMI≥28 Inclusion: Patients who had undergone percutaneous liver biopsy within 6 months or were scheduled to undergo one in the next month were eligible. Exclusion: Pregnancy, ascites, implantable cardiac devices, previous liver transplant, terminal disease.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (FibroScan M probe and Fibroscan XL probe [optimal liver stiffness cut-offs that maximized the sum of sensitivity and specificity: M probe 22.3 kPa, XL probe 16 kPa]): Performed by 9 experienced operators as per manufacturer's instructions. Both M (standard) and XL (specifically designed for obese patients) were used on all subjects. No successful measurements after 10 attempts was deemed a failure. Exams with fewer than 10 valid measurements, an IQR>30% or <60% were considered unreliable. Study aims to compare the M and XL probe in the same patients. Note: The Fibroscan XL probe has been designed specifically for use in obese patients by utilisation of a lower frequency and more sensitive ultrasonic transducer, a deeper focal length, a larger vibration amplitude and a greater depth of measurement below the skin surface.
Reference standard	Liver biopsy (METAVIR F4): Specimens analysed by 2 experienced hepatopathologists without knowledge of other clinical data. Biopsies less than 15 mm in length and/or with fewer than 6 portal triads were deemed uninterpretable (length range 15–53 mm, portal tracts range 7–39), obtained under ultrasound guidance. Tissue was fixed, paraffin-embedded and stained with at least hematoxylin, eosin and Masson's trichrome.
Time between index test and reference standard	Within 6 months
Prevalence of cirrhosis according to reference standard	32/276 (12%), not reported for NAFLD population separately
Target condition	Cirrhosis
Results: Fibroscan M probe AUC (95% CI): 0.88 (0.75–1.00) Optimal cut-off threshold (if calculated): 22.3 kPa Threshold: 22.3 kPa Sensitivity: 80 (28–99) Specificity: 91 (82–97) PPV: 40 (12–74)	

Study	MYERS 2012B[90]
NPV: 98 (92–100) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Fibroscan XL probe AUC (95% CI): 0.95 (0.89–1.00) Optimal cut-off threshold (if calculated): 16.0 kPa Threshold: 16.0 kPa Sensitivity: 100 (54–100) Specificity: 91 (84–96) PPV: 40 (16–68) NPV: 100 (96–100) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Invalid liver stiffness measurements in whole population: XL probe 1.1%, M probe 16%. Failure of the M probe increased as BMI increased. Also reported data for a mixed hepatitis B and C population (did not use). Comparable with the M probe, the FibroScan XL probe reduces TE failure and facilitates reliable LSM in obese patients. Although the probes have comparable accuracy, lower liver stiffness cut-offs will be necessary when the XL probe is used to non-invasively assess liver fibrosis. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II:	

Study	MYERS 2012B[90]
Random or consecutive recruitment not reported Up to 6 months between index test and reference standard Liver biopsy sample <25 mm and 10 portal tracts	

Study	RIZZO 2011[106]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=146 consecutive patients evaluated, 5 excluded for suboptimal liver biopsy, 2 excluded with alcohol abuse, enrolled n=139). Recruitment between November 2008 and October 2009.
Countries and Settings	Italy, 3 Hospitals (Infectious Diseases Units of the Garibaldi Nesima and Ferrarotto Hospitals in Catania and the Hepatology Unit of the University Hospital, Palermo)
Funding	None
Age, gender, ethnicity	Age, mean (SD): 55 (12); male/female: 83/56; ethnicity: not reported; ALT (U/l): 77.2 (33.0)
Patient characteristics	Population: Chronic hepatitis C (viral and histologic diagnosis) Inclusion: Presence of active HCV replication, and on a liver histology consistent with chronic hepatitis Exclusion: HBV/ HIV co-infection, alcohol abuse (>20 g/ day in the last year or more, evaluated by questionnaire), with Child B or C cirrhosis, and those under antiviral treatment
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan, Echosens, France [cut-off 11 kPa, determined as optimal cut-off by Kolmogorov – Smirnov index]: Performed by 2 expert physicians, 1 in Palermo and 1 in Catania, according to the manufacturer's instructions. Both examiners were blinded to clinical and pathological data. ARFI (cut-off 2 m/s, determined as optimal cut-off by Kolmogorov – Smirnov index): B-mode standard ultrasonography scanning and ARFI elastography were performed using a Siemens Acuson S2000 (Siemens AG, Erlangen, Germany) with a 4Cl transducer. Liver stiffness was measured with ARFI elastography by 2 independent investigators: 1 in Catania and 1 in Palermo. Both investigators were blinded to all patients ' clinical, serological, and histological data. ARFI elastography was performed on fasting patients, choosing as the target the right lobe of the liver, which was accessed through the intercostal spaces. The velocity of the shear wave (in m/s) in the liver tissue was collected and recorded from 20 different sites, 5 sites for each segment (V, VI, VII, and VIII) within the right lobe. A median of the 20 results has been calculated.

Study	RIZZO 2011[106]
Reference standard	Liver biopsy (METAVIR F4): Liver biopsy specimens were obtained using Menghini 16G disposable needles. All biopsy specimens contained at least 10 portal tracts and were minimum 1.5 cm in length. All biopsy specimens were coded and evaluated by a single experienced pathologist, who was blinded to the patients ' clinical and imaging results.
Time between index test and reference standard	Within 6 months, median 3 months (range 1–6 months)
Prevalence of cirrhosis according to reference standard	30/139 (21.6%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.80 (0.72–0.86) Optimal cut-off threshold (if calculated): 11 kPa Threshold: 11 kPa (optimal) Sensitivity: 70 Specificity: 82 PPV: 53 NPV: 90 +ve/-ve likelihood ratios: 3.9/0.4 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: ARFI AUC (95% CI): 0.89 (0.83–0.94) Optimal cut-off threshold (if calculated): 2 m/s Threshold: 2 m/s (optimal) Sensitivity: 83 Specificity: 86 PPV: 63	

Study	RIZZO 2011[106]
	NPV: 95 +ve/-ve likelihood ratios: 6.1/0.2 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: TE was unreliable in 9 patients (6.5 %). In an extra analysis to check interobserver agreement, there was no significant difference between the ARFI values of the 21 patients obtained from the 2 different sonographers. ARFI performance was not statistically significantly higher than TE performances for the diagnosis of cirrhosis (p= 0.09). Also analysed partial AUC. Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Up to 6 months between index test and reference standard Liver biopsy sample <25 mm

Study	SANCHEZ-CONDE 2010[117]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study, n=105 (3 excluded due to inadequate biopsies, 2 excluded due to uninterpretable TE). n=100 included in the analysis. January 2007–January 2008
Countries and Settings	HIV outpatient clinic of 2 teaching hospitals in Spain, Madrid
Funding	Spanish AIDS investigation group and Spanish Health Research Fund
Age, gender, ethnicity	Age, mean (range): 42 (39–46); male/female: 29% female; ethnicity: not reported; ALT (U/l): 67.6±41.8 IU/ml
Patient characteristics	Population: Hepatitis C and HIV co-infected, mostly potential candidates for HCV therapy Inclusion: Detectable HCV-RNA by polymerase chain reaction Exclusion: Hepatic decompensation, hepatitis B, anti-HCV therapy.

Study	SANCHEZ-CONDE 2010[117]
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan): Optimal cut-off values based on the highest NPV with an acceptable PPV higher than 50%. Performed according to standard procedure. Performed by the same trained personnel at each centre. IQR <30% and procedures with at least 10 validated measurements and a success rate of 60% accepted. APRI, FIB-4: Diagnostic accuracy for significant fibrosis only.
Reference standard	Liver biopsy (METAVIR F4): Ultrasound routinely performed to determine percutaneous biopsy site. Biopsies evaluated by an experienced pathologist who had no knowledge of clinical and laboratory data. Biopsies were '25 mm in length in most cases'. Formalin-fixed, paraffin-embedded liver tissue was stained by haematoxylin-eosin, Mason's trichrome and Perl's iron.
Time between index test and reference standard	No more than 6 months
Prevalence of cirrhosis according to reference standard	8/100 (8%)
Target condition	Cirrhosis
Results: Transient elastography AUC (95% CI): 0.99 (0.97–1.00) Optimal cut-off threshold (if calculated): (chosen threshold) 14 kPa Threshold: 14 kPa (optimal) Sensitivity: 100 (93.7–100.0) Specificity: 93.5 (87.9–99.1) PPV: 57.1 (27.6–86.6) NPV: 100 (99.4–100) +ve/-ve likelihood ratios: 15.33 (7.07–33.24)/not reported TP: 8 FP: 6 FN: 0 TN: 86 Other measures reported and conclusions: TE accurately predicted liver fibrosis and outperformed other simple non-invasive indexes in HIV/HCV co-infected patients. Any complications associated with tests reported: Not reported	

Study	SANCHEZ-CONDE 2010[117]
General limitations according to QUADAS II: Random or consecutive recruitment not reported Up to 6 months between index test and reference standard Some liver biopsies <25 mm (unclear how many)	

Study	Shehab 2014[129]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study. n=994 (split into training and validation cohorts for the development of a new fibrosis marker, PLASA. However, all patients used for diagnostic accuracy of index tests measures reported here, minus those without available data on all variables: final analysis n=842). Consecutive treatment naïve patients with chronic hepatitis C. January 2010–October 2013
Countries and Settings	Two hospitals in Egypt
Funding	Not reported
Age, gender, ethnicity	Age, mean (SD): 42.4 (9.7); male/female 875/119; ethnicity: not reported; ALT (U/l): 56.6 (14–350)
Patient characteristics	Population: Treatment-naïve patients with chronic hepatitis C (HCV) Inclusion: Positive HCV RNA, compensated liver disease and availability of serum biomarker results done within 1 month prior to liver biopsy. Exclusion: Co-infection with HBV or HIV; other causes of liver disease; alcohol consumption higher than 20 g/day, HCC, prior liver transplant; Gilbert disease; chronic haemolysis; previous antiviral treatment and use of medications that could alter the measured laboratory parameters.
Index test (including threshold and whether threshold pre-specified)	APRI; FIB-4: From routine lab parameters and basic clinical data, retrieved from medical records. Only lab tests performed within 1 month before the biopsy were included. APRI (pre-published cut-off values of 0.5 and 2.0): [(AST/ULN) x100] / platelet count 109/l FIB-4 (pre-published cut-off of 3.25): [age (years) x AST (IU/l)] / platelet count 109/l x ALT (IU/l)1/2

Study	Shehab 2014[129]
Reference standard	Liver biopsy (METAVIR F4): Patients with biopsy samples shorter than 1.5 cm or containing less than 7 portal tracts were excluded. A single experienced pathologist examined the biopsy specimens in each centre. This person was blind to the laboratory data of the patient.
Time between index test and reference standard	Within 1 month
Prevalence of cirrhosis according to reference standard	260/994 (26.2%). Not reported for the 842 included in the final analysis.
Target condition	Cirrhosis
Results: APRI AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 0.5 (published) Sensitivity: 100 Specificity: 12.8 PPV: 5.3 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 2.0 (published) Sensitivity: 15.4 Specificity: 96 PPV: 15.8 NPV: 95.9 +ve/-ve likelihood ratios: Not reported	

Study	Shehab 2014[129]
TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: FIB-4 AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 3.25 (published) Sensitivity: 28.2 Specificity: 93.5 PPV: 17.5 NPV: 96.4 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Liver biopsies <25 mm and <10 portal tracts Data were not available for all variables for a large proportion of patients and only 842 included in the final analysis.	

Study	SILVIA JUNIOR 2014[130]
Study type	Cross-sectional study

Study	SILVIA JUNIOR 2014[130]
Number of studies (number of participants). Recruitment period.	1 study (n=51 consecutive patients). Recruitment from January 2012-March 2013
Countries and Settings	Santa Casa de Sao Paulo Hospital, Brazil
Funding	Not stated
Age, gender, ethnicity	Age, mean (SD): 53.8±1.53; male/female: 18 male, 33 female; ethnicity: not reported; ALT (IU/l): 60.55±6.3
Patient characteristics	Population: Chronic untreated hepatitis C Inclusion: CHC diagnosis was established by the presence of hepatitis C virus RNA using qualitative polymerase chain reaction. Exclusion: HIV, hepatitis B, alcohol abuse, cholestatic chronic hepatitis, non-alcoholic steatohepatitis, autoimmune chronic hepatitis, hemochromatosis, Wilson's disease, hepatocellular carcinoma, prior liver transplantation, prior interferon therapy, immunosuppressive therapy.
Index test (including threshold and whether threshold pre-specified)	ARFI elastography (optimal cut-off value 1.95 m/s determined by a common optimisation step that maximised the sum of the sensitivities in predicting the single stages): Performed with Siemens Acuson S2000 ultrasound system (Siemens Medical Solutions, Brazil) using a standard ultrasonographic probe on the right lobe of the liver. All procedures performed in a single centre by a single physician, experienced in digestive system ultrasonography and blinded to the clinical, serological and histological data. A median was calculated based on 10 measurements. APRI (optimal cut-off value 1.71 determined by a common optimisation step that maximised the sum of the sensitivities in predicting the single stages): [(AST/ULN) x100] / platelet count 109/l FIB-4: [age (years) x AST (IU/l)] / platelet count 109/l x ALT (IU/l)1/2 Blood tests performed within the same week as liver biopsy (ARFI and FIB-4).
Reference standard	Liver biopsy METAVIR F4. Biopsy length median 20.6 mm (range 15–28 mm), median portal tracts 10.1 (range 8–14). Percutaneous liver biopsy was performed by senior operators using the TruCut technique with manual or semi-automatic instruments. Tissue was fixed in formalin paraffin-embedded and stained with hematoxylin-eosin and Masson's trichrome. Specimens were analysed by an expert pathologist blinded to biological and clinical data.
Time between index test and reference standard	Up to 6 months (median 2.8 months)
Prevalence of cirrhosis according to reference standard	9/51 (17.6%)

Study	SILVIA JUNIOR 2014[130]
Target condition	Cirrhosis
Results: ARFI AUC (95% CI): 0.98 (CI not reported) Optimal cut-off threshold (if calculated): 1.95 m/s Threshold: 1.95 m/s (optimal) Sensitivity: 100 Specificity: 95.2 PPV: 81.8 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.89 (CI not reported, value taken from table, incorrectly reported in text) Optimal cut-off threshold (if calculated): 1.71 Threshold 1.71 (optimal): Sensitivity: 66.7 Specificity: 92.9 PPV: 60 NPV: 90.5 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	SILVIA JUNIOR 2014[130]
Results: FIB-4 AUC (95% CI): 0.94 (CI not reported) Optimal cut-off threshold (if calculated): Not reported Threshold: Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Forns score, King score. ARFI elastography had very good accuracy for the assessment of fibrosis and was more effective for the prediction of cirrhosis than the blood tests. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Up to 6 months between index test and reference standard Liver biopsies <25 mm	

Study	SIRLI 2010[132]
Study type	Retrospective cohort study
Number of studies (number of participants). Recruitment period.	1 study (n=150; TE measurements only obtained for 144 patients) Recruited from January – December 2008

Study	SIRLI 2010[132]
Countries and Settings	Department of Gastroenterology and Hepatology, Timisoara, Romania
Funding	Not stated
Age, gender, ethnicity	Age, mean (SD): 50.1±10.3; male/female: 48/102; ethnicity: not stated; ALT (U/l): not stated
Patient characteristics	Population: Chronic hepatitis C Inclusion: Normal iron load and ceruloplasmin Exclusion: Ascites, hepatitis B, alcohol abuse, cholestasis, steatosis, autoimmune hepatitis, primary biliary cirrhosis, biliary obstruction
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan [optimal cut-off value of 13.3 kPa chosen to maximise the sum of the sensitivity and specificity]): Performed by 3 experienced physicians by standard method. Ten valid measurements. Only those with a success rate of at least 60% with IQR <30%. APRI (optimal cut-off value of 1.38 chosen to maximise the sum of the sensitivity and specificity): [(AST/ULN) x100] / platelet count 109/l FIB-4 (optimal cut-off value of 2.3122 chosen to maximise the sum of the sensitivity and specificity): [age (years) x AST (IU/l)] / platelet count 109/l x ALT (IU/l)1/2 Platelet count (optimal cut-off value of 155000/mm^3 chosen to maximise the sum of the sensitivity and specificity) Blood collected in the same session as TE and liver biopsy.
Reference standard	Liver biopsy (METAVIR F4). Echo-assisted using Menghini-type modified needles, 1.4 and 1.6 mm in diameter. Only biopsies of at least 20 mm and 8 portal tracts considered adequate and included in the study. Assessed by a senior pathologist.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	15/150 (10%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.979 (0.85–0.951) Optimal cut-off threshold (if calculated): 13.3 kPa Threshold 13.3 kPa (optimal):	

Study	SIRLI 2010[132]
Sensitivity: 93.3 Specificity: 96.1 PPV: 73.7 NPV: 99.2 +ve/-ve likelihood ratios: 24.08/0.07 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.909 (0.85–0.951) Optimal cut-off threshold (if calculated): 1.38 Threshold 1.38 (optimal): Sensitivity: 93.3 Specificity: 83 PPV: 37.8 NPV: 99 +ve/-ve likelihood ratios: 5.48/0.08 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: FIB-4 AUC (95% CI): 0.842 (0.772–0.898) Optimal cut-off threshold (if calculated): 2.3122 Threshold 2.3122 (optimal): Sensitivity: 80	

Study	SIRLI 2010[132]
Specificity: 77.8 PPV: 28.6 NPV: 97.2 +ve/-ve likelihood ratios: 3.6/0.26 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Platelet count AUC (95% CI): 0.899 (0.838–0.943) Optimal cut-off threshold (if calculated): 155000 mm^3 Threshold 155000 mm^3 (optimal): Sensitivity: 86.7 Specificity: 83.7 PPV: 37.1 NPV: 98.3 +ve/-ve likelihood ratios: 5.32/0.16 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Forns test, Lok test LSM better than blood fibrosis tests for predicting cirrhosis but all had excellent predictive value. Any complications associated with tests reported: Not reported
General limitations according to QUADAS II:

Study	SIRLI 2010[132]
Consecutive or random selection not reported. Unknown if the reference standard results were interpreted without knowledge of the index test results Liver biopsies <25 mm and <10 portal tracts	

Study	SPOREA2011A[138]
Study type	Prospective cross sectional
Number of studies (number of participants). Recruitment period.	1 study (n=197 patients). Recruitment period not reported.
Countries and Settings	Romania, 2 university hospitals
Funding	None reported
Age, gender, ethnicity	Age, mean (SD): 50(9.8); male/female: 78/119; ethnicity: not reported; ALT (U/l): not reported
Patient characteristics	Population: Chronic HCV hepatitis Inclusion: Anti-HCV antibodies positive, with or without cytolysis for at least 6 months, PCR HCV RNA positive. Exclusion: Patients with other causes of chronic hepatitis (HBV infection, chronic alcohol abuse, cholestatic chronic hepatitis, non-alcoholic steatohepatitis, autoimmune chronic hepatitis, haemochromatosis, Wilson's disease)
Index test (including threshold and whether threshold pre-specified)	Transient elastography (optimal cut-off value 12.2 kPa was chosen to maximize the sum of sensitivity and specificity): Fibroscan device (Echosens, Paris, France) by experienced physicians (more than 500 TE), blinded to the results of LB and ARFI measurements. In each patient, 10 valid measurements were performed, after which a median value of LS was obtained. Only patients in which LS measurements by means of TE had a success rate of at least 60%, with an IQR <30%, were included. ARFI (optimal cut-off value 1.8 m/s was chosen to maximize the sum of sensitivity and specificity): Ultrasound device ACUSON S2000 (Siemens). Scanning was performed between the ribs in the right liver lobe in order to avoid cardiac motion (approximately in the place where we usually perform LB), 1 cm under the capsule. Ten measurements in every patient, and a median value was calculated, the result being measured in m/s. Only patients in which LS measurements by means of ARFI had a success rate of at least 60%, with an IQR <30%, were included. Operators were blinded to the results of LB and TE measurements.

Study	SPOREA2011A[138]
	Combination of TE and ARFI (values both for TE and ARFI above the mentioned cut-offs) Combination of TE or ARFI (values both for TE and ARFI above the mentioned cut-offs)
Reference standard	Liver biopsy (METAVIR F4): Echo-guided TruCut technique, with a 1.8 mm (14 G) diameter automatic needle device-Biopty Gun (Bard GMBh), or echo-assisted, using Menghini type modified needles, 1.4 and 1.6 mm in diameter. Only LB fragments including at least 6 portal tracts were included. The LBs were assessed by a senior pathologist (1 in each centre) blinded to the results of TE and ARFI measurements.
Time between index test and reference standard	Same session
Prevalence of cirrhosis according to reference standard	53/197 (26.9%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.97 Optimal cut-off threshold (if calculated): 12.2 kPa Threshold: 12.2 kPa (optimal) Sensitivity: 96.2 Specificity: 89.6 PPV: 78.1 NPV: 98.3 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: ARFI AUC (95% CI): 0.91 Optimal cut-off threshold (if calculated): 1.8 m/s	

Study	SPOREA2011A[138]
Threshold: 1.8 m/s (optimal) Sensitivity: 90.4 Specificity: 85.6 PPV: 50.3 NPV: 95.8 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Combination of Fibroscan and ARFI AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: Values both for TE and ARFI above the cut-offs12.2 kPa and 1.8 m/s (optimal) Sensitivity: 84.9 Specificity: 94.4 PPV: 84.9 NPV: 94.4 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Combination of Fibroscan or ARFI AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: Values for TE or ARFI above the cut-offs12.2 kPa or 1.8 m/s (optimal)	

Study	SPOREA2011A[138]
Sensitivity: 96.2 Specificity: 83.3 PPV: 68.0 NPV: 98.3 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Obtained valid TE measurements in 187/197 patients (94.9%) and valid ARFI measurements in 191/197 patients (96.9%). Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random selection not reported. Liver biopsy sample <10 portal tracts.	

Study	SPOREA 2012A [137] Also included data from the following studies: 5 studies which were included and additional information extracted from the individual study[33,42,84,103,136] and 2 studies which were excluded from our review due to data only being available for mixed aetiologies[48,145] (presumed authors were contacted for further information).
Study type	Retrospective multi-centre
Number of studies (number of participants). Recruitment period.	914 (10 centres, 5 countries) ARFI obtained in 911 TE measured in 400
Countries and Settings	Romania, Japan, Germany, Italy, Austria
Funding	Not reported (however 4 authors are associated with Siemens and 1 is associated with Echosens)
Age, gender, ethnicity, ALT (U/l):	Mean age: 55.7±13.1, gender: 53.7% women, ethnicity: 49.6% European, 50.4% Asian, ALT: 1.6±1.7 x ULN

Study	**SPOREA 2012A [137] Also included data from the following studies: 5 studies which were included and additional information extracted from the individual study[33,42,84,103,136] and 2 studies which were excluded from our review due to data only being available for mixed aetiologies[48,145] (presumed authors were contacted for further information).**
Patient characteristics	Population: Chronic HCV Inclusion: Positive anti-HCV antibodies and positive PCR HCV RNA for more than 6 months. Homogenous liver structure (without liver masses). Exclusion: HIV or hepatitis B co-infection, ascites
Index test (including threshold and whether threshold pre-specified)	ARFI (optimal cut-off values were chosen so that the sum of sensitivity [Se] and specificity [Sp] would be the highest) – performed in all patients with a Siemens Acuson S2000TM ultrasound system with 4Cl transducers. Scanning was performed with a right intercostal approach, in the right liver lobe, segment V-VIII, 1–2 cm (Hyogo, Timisoara) or 2–3 cm (other centres) under the liver capsule, with minimal scanning pressure applied by the operator, while the patients were asked to stop normal breathing for a moment in order to minimize breathing motion. The operator selects the depth at which the liver elasticity is evaluated by placing a "measuring box" (10 mm long, 5 mm wide) in the desired area. The maximum depth at which ARFI measurements can be performed is 8 cm. A total of 5 (Saga), 6 (Bologna, Verona) or 10 (all other centres) valid measurements were performed in every patient and the median value was calculated. Operators who performed ARFI measurements were blinded to all patients' clinical, serological and histological data. TE (optimal cut-off values were chosen so that the sum of sensitivity [Se] and specificity [Sp] would be the highest) – measured using FibroScan. 10 measurements were performed in each patient and the median calculated. Only measurements with a success rate ≥60% and an interquartile range <30% were considered reliable. ARFI and TE were performed in the same session.
Reference standard	Liver biopsy (METAVIR F4): Percutaneous liver biopsy using Menghini needle in 5 centres (Timisoara – needle diameter 1.4 or 1.6) Bucharest 1.4 mm, Bologna and Verona – 1.4 or 1.6 mm and Frankfurt – 1.2 mm). Percutaneous biopsy using TruCut technique with automatic needle device in 2 centres (Cluj-Napoca – 14 G needle and Hyogo – 16 G needle) percutaneous biopsy using semi-automatic instruments in 2 centres (Saga – 16 G needle and Tokyo – 18 G needle) and transjugular biopsy in 1 centre (Vienna). Only fragments of at least 1.5 cm in length were included. Biopsies were performed in the right lobe and assessed by a senior pathologist, blinded to the results of liver stiffness measures.
Time between index test and reference standard	Up to 6 months
Prevalence of cirrhosis according to reference standard	223/911 (24.4% in whole group) 95/400 (23.8% in TE subgroup)
Target condition	Cirrhosis
Results: ARFI	

Study	SPOREA 2012A [137] Also included data from the following studies: 5 studies which were included and additional information extracted from the individual study[33,42,84,103,136] and 2 studies which were excluded from our review due to data only being available for mixed aetiologies[48,145] (presumed authors were contacted for further information).
AUC (95% CI): 0.842 Optimal cut-off threshold (if calculated): 1.55 m/s (or 1.69 m/s reported for n=400 subgroup who also had TE) Threshold: 1.55 m/s (optimal) Sensitivity: 84.3 Specificity: 76.3 PPV: 53.1 NPV: 93.7 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Accuracy: 77.9% Results: TE (n=400) AUC (95% CI): 0.932 Optimal cut-off threshold (if calculated): 11.9 kPa Note: Sporea 2012a did not report the sensitivities and specificities for TE at a cut-off threshold. This information was extracted separately for 5 of the studies used in the Sporea 2012 pooled data and is reported below (this did not include additional patients included in Sporea 2012a who weren't reported in previous papers, nor did it include Takahashi 2012 or Friedrust 2009A as these papers did not report data separately for HCV and/or for people with biopsy as the reference standard). ARFI data were not extracted from these papers separately, as this will be included in the above analysis. Lupsor 2009[84] (n=112); cirrhosis F4: 42/112 (37.5%): Threshold: >13.1 (optimal) Sensitivity: 95.12 Specificity: 89.17 PPV: 84.8	

Study	SPOREA 2012A [137] Also included data from the following studies: 5 studies which were included and additional information extracted from the individual study[33,42,84,103,136] and 2 studies which were excluded from our review due to data only being available for mixed aetiologies[48,145] (presumed authors were contacted for further information).
NPV: 96.8 +ve/-ve likelihood ratios: 9.24/0.05 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Fierbinteanu-braticeuici 2009[42] (n=74) TE not assessed by study, APRI assessed by study but accuracy values not reported Ebinuma 2011[33]; cirrhosis F4: Diagnostic accuracy of TE not reported separately for HCV aetiology (only splits into viral and non-viral aetiologies) Piscaglia 2011[103]; cirrhosis F4: Diagnostic accuracy of TE for cirrhosis not reported Sporea 2011D[136]; cirrhosis F4: Diagnostic accuracy of TE for cirrhosis not reported (only for diagnosis of significant fibrosis) Other measures reported and conclusions: Predictive ARFI values separated by ethnicity. Performance of ARFI according to ALT level. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Consecutive or random recruitment not reported Up to 6 months between reference standard and index test Liver biopsies <25 mm.	

Study	STIBBE 2011[139]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=89 (48 HBV patients, 41 HCV patients [only 40 included in FibroTest, 36 included in TE], 31 controls) February 2007–November 2007
Countries and Settings	Department of Gastroenterology and Hepatology, Erasmus MC, University Medical Centre, Rotterdam, The Netherlands
Funding	Not reported
Age, gender, ethnicity, ALT (U/l):	Mean age: 47 years; 66% men; ethnicity: not reported; ALT: not reported for HCV patients separately
Patient characteristics	Population: Chronic viral hepatitis C Inclusion: Mono-infected HCV patients referred for liver biopsy to the outpatient clinic. Exclusion: Alcohol intake >20 g/day, co-infection with HIV or hepatitis D, presence of hepatocellular carcinoma
Index test (including threshold and whether threshold pre-specified)	FibroTest (pre-published cut-off from Poynard et al.): blood samples were obtained from all patients on the day of biopsy. FibroTest was based on sex, age, α2M, haptoglobin, total bilirubin, γGT and ApoA1. Transient elastography (Fibroscan; pre-published cut-off Verveer, personal communication): preceded the biopsy in the same session. TE measured low-frequency elastic waves (50 Hz) through a medium and the speed of these waves was positively correlated with stiffness of the liver. A success rate of >60% was considered reliable in 10 validated measurements with an interquartile range (IQR) <30% of the median.
Reference standard	Liver biopsy (METAVIR F4): Two well-experienced hepatologists performed all biopsies. To reduce complications, during this procedure abdominal ultrasound was used to identify liver parenchymal and vascular structures. Biopsies were taken with a 14 G true-cut needle and required a length ≥20 mm. Two expert hepatopathologists scored all specimens (double read) for different fibrosis categories using Metavir scoring. No biopsies obtained from controls.
Time between index test and reference standard	Same day
Prevalence of cirrhosis according to reference standard	11/41
Target condition	Cirrhosis
Results: FibroTest (n=40) AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported	

Study	STIBBE 2011[139]
Threshold: 0.75 (published) Sensitivity: 100 Specificity: 24 PPV: 64 NPV: 100 +ve/-ve likelihood ratios: 1.31/0 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: TE (n=36) AUC (95% CI): Not reported Optimal cut-off threshold (if calculated): Not reported Threshold: 14 kPa (pre-published) Sensitivity: 88 Specificity: 73 PPV: 88 NPV: 73 +ve/-ve likelihood ratios: 3.23/0.16 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Breath tests, APRI, FIB-4. For APRI and FIB-4, and for a combination of TE and fibrosis tests, results were only given for all patients combined and not for HCV separately. Hyaluronic acid, APRI, FibroTest, Fib-4 and TE reliably distinguish non-cirrhotic and cirrhotic patients.	
General limitations according to QUADAS II:	

Study	STIBBE 2011[139]
Consecutive or random recruitment not reported. Blinding unclear during interpretation of reference standard test results. Liver biopsy size <25 mm.	

Study	Wong 2010B[160]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=309 consecutive patients, 35 excluded due to biopsy length, 28 excluded due to failure to obtain 10 valid LSM acquisitions, final analysis n=246). Recruitment between May 2003 and April 2009.
Countries and Settings	France and Hong Kong. Two University Hospitals.
Funding	Academic. Supported in part by the research fund of the Department of Medicine and Therapeutics, The Chinese University of Hong Kong.
Age, gender, ethnicity	Age, mean (SD): 51(11); male/female 135/111: ethnicity: Caucasian (n=128) and Chinese (n=118); ALT (IU/L): 75(54); BMI: 28.0(4.5); Diabetes: 36.2%.
Patient characteristics	Population: NAFLD Inclusion: Aged 18 years or older, with NAFLD undergoing liver biopsy. Exclusion: Men who consumed more than 30 g alcohol per day and women who consumed more than 20 g alcohol per day; secondary causes of hepatic steatosis (such as chronic use of systemic corticosteroids), positive hepatitis B surface antigen, anti-hepatitis C virus antibody, or histological evidence of other concomitant chronic liver diseases; patients with clinical and radiological evidence of cirrhosis were excluded (for example, bilirubin 30 ≥ mol/L, albumin <35 g/L, INR>1.3, platelet count <150x109/L, ascites, varices, splenomegaly).
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan), optimal cut-off threshold calculated (10.3 kPa) according to highest Youden's index. Accuracy also given at cut-off of 11.5 kPa (not pre-specified). Performed according to the instructions and training provided by the manufacturer. Ten successful acquisitions were performed on each patient. The median value represented the liver elastic modulus. Only cases with 10 successful acquisitions were evaluated. The operators were blinded to all clinical data and the diagnoses of the patients. Presumed to have used appropriate probe for patient's BMI according to manufacturer's instructions (not reported).

Study	Wong 2010B[160]
	APRI, AST/ALT and FIB-4
Reference standard	Liver biopsy (NAFLD specific scoring system, Kleiner et al 2005, F4): Percutaneous liver biopsy was performed using the 16 G Temno or Menghini needle. Liver histology was assessed by experienced histopathologists (B.L.B., P.C.C.) who were blinded to the clinical data. Liver specimens shorter than 15 mm were excluded (mean (SD) length 21(7)mm)
Time between index test and reference standard	Index test 1 week before
Prevalence of cirrhosis according to reference standard	25/246 (10.2%)
Target condition	Cirrhosis
Results: Fibroscan AUC (95% CI): 0.95 (0.91–0.99) Optimal cut-off threshold (if calculated): 10.3 kPa Threshold: 10.3 kPa (optimal) Sensitivity: 92.0 Specificity: 87.8 PPV: 46.0 NPV: 99.0 +ve/-ve likelihood ratios: 7.5/0.091 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 11.5 kPa (not pre-specified: cut-off giving specificity >90%) Sensitivity: 76.0 Specificity: 91.0 PPV: 48.7 NPV: 97.1	

Study	Wong 2010B[160]
+ve/-ve likelihood ratios: 8.4/0.26 TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: APRI AUC (95% CI): 0.75 (0.64–0.85) Results: FIB-4 AUC (95% CI): 0.81 (0.73–0.89) Results: AST/ALT AUC (95% CI): 0.66 (0.55–0.77) Other measures reported and conclusions: Transient elastography had high accuracy in detecting advanced fibrosis and cirrhosis. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS: Patients with unreliable TE excluded from the analysis Liver biopsy sample <25 mm and 10 portal tracts.	

Study	WONG 2012[159]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	n=205 consecutive NAFLD patients (12 patients were excluded because of liver biopsy length < 15 mm, final analysis 193). Recruitment period October 2009 to September 2011. Reliable results were obtained in 67% with M probe and 75% with XL probe (note: report intention to diagnose results here and cases with failed liver stiffness measurements were labelled as

Study	WONG 2012[159]
	incorrect classifications, study also reports accuracies not including those without valid TE measurements).
Countries and Settings	France and Hong Kong. Two University Hospitals.
Funding	Partially supported by the PROCORE-France/Hong Kong Joint Research Scheme (F-HK17 / 10T) and a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project no. CUHK477710).
Age, gender, ethnicity	Age, mean (SD): 52±11 years; male/female: 110/83; ethnicity: Caucasian 77, Chinese 116; ALT (IU/L): 73 (76); BMI: 28.9± 4.8. Sixty-eight (35 %) patients had BMI ≥ 30.
Patient characteristics	Population: NAFLD Inclusion: Indications of liver biopsy included persistently abnormal liver biochemistry and the presence of risk factors of advanced disease such as type 2 diabetes. Enrolled patients aged ≥ 18 years. Exclusion: Men who consumed more than 30 g alcohol per day and women who consumed more than 20 g alcohol per day; patients with secondary causes of hepatic steatosis (such as use of systemic corticosteroids and methotrexate), positive hepatitis B surface antigen, anti-hepatitis C virus antibody, or histological evidence of other concomitant liver diseases.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan) optimal cut-offs chosen at points with the highest Youden 's index based on cases with 10 valid measurements, cut-offs with sensitivity and specificity over 90% were also determined. Measurements were performed on the right lobe of the liver through intercostal spaces with the patient lying in dorsal decubitus with the right arm in maximal abduction. Ten successful acquisitions were performed on each patient. The success rate was calculated as the number of successful measurements divided by the total number of measurements. In each patient, measurements were performed by M probe followed by XL probe. The maximum number of measurements by each probe was limited at 20. The operators were blinded to all clinical data and the diagnoses of the patients, and had performed LSM on at least 50 patients before this study. An LSM was considered reliable only if 10 valid acquisitions were obtained, the success rate was over 60%, and the IQR-to-median ratio (IQR/M) of the measurements was below 0.3. Study aims to compare the M and XL probe in the same patients.
Reference standard	Liver biopsy (NAFLD specific scoring system, Kleiner et al 2005, F4): Percutaneous liver biopsy was performed using the 16 G Temno or Menghini needle. Liver histology was assessed by 2 experienced histopathologists who were blinded to the clinical data. Liver specimens shorter than 15 mm were excluded (mean 24±6).
Time between index test and reference standard	TE 24 hours before liver biopsy
Prevalence of cirrhosis according to reference standard	25/193 (13%)
Target condition	Cirrhosis

Study	WONG 2012[159]
Results: Fibroscan M probe AUC (95% CI): 0.53 (0.36–0.70) Optimal cut-off threshold (if calculated): 10.3 kPa (Youden's) Threshold: 10.3 (Youden's and highest sensitivity) Sensitivity: 52 (32–72) Specificity: 69 (62–76) PPV: 20 (10–30) NPV: 91 (86–96) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 11.5 (highest specificity) Sensitivity: 44 (25–64) Specificity: 71 (64–78) PPV: 18 (9–28) NPV: 90 (84–95) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Results: Fibroscan XL probe AUC (95% CI): 0.86 (0.79–0.94) Optimal cut-off threshold (if calculated): 7.9 kPa (Youden's) Threshold: 7.9 kPa (Youden's) Sensitivity: 84 (70–98)	

Study	WONG 2012[159]
Specificity: 72 (65–79) PPV: 31 (20–42) NPV: 97 (94–100) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 7.2 kPa (best sensitivity) Sensitivity: 88 (75–100) Specificity: 67 (60–74) PPV: 28 (18–38) NPV: 97 (95–100) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Threshold: 11.0 kPa (best specificity) Sensitivity: 68 (50–86) Specificity: 86 (81–92) PPV: 43 (27–58) NPV: 95 (91–98) +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	WONG 2012[159]
Other measures reported and conclusions: By intention-to-diagnose analysis, the performance of M probe was unsatisfactory due to the large number of patients with failed LSM. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Liver biopsy sample <25 mm	

Study	Yamanda 2006[161]
Study type	Pilot study
Number of studies (number of participants). Recruitment period.	n=74 HCV and HBV in total (including 44 with hepatitis C)
Countries and Settings	Chiba University Hospital, Japan
Funding	Not reported
Age, gender, ethnicity	In the whole group mean age=51±11 years (range 19–70 years); 55.4% males; ethnicity not stated (presumed Japanese).
Patient characteristics	Hepatitis C infected
Index test (including threshold and whether threshold pre-specified)	Ultrasound (SSA 770A, Toshiba Medical Systems, Tokyo, Japan). Transforming and receiving frequencies were 2.0 and 4.0 MHz respectively. The transducer was applied lengthways to the epigastric lesion of the patient's body surface, moving it in a linear fashion along the patient's skin manually about 3 cm for 100 consecutive ultrasound images. Patients held their breath during scanning (approximately 15 seconds).
Reference standard	Percutaneous liver biopsy by 18-gauge needle with 20 mm specimen notch. Only samples presenting at least 10 portal tracts were considered suitable for evaluation. Specimens were evaluated with regard to inflammatory activity and fibrosis in a blind fashion by 2 independent liver pathology specialists based on the New European Classification (same as METAVIR).
Time between index test and reference standard	A few days
Prevalence of cirrhosis according to reference standard	Not reported for HCV population

Study	Yamanda 2006[161]
Target condition	Cirrhosis
Results: Ultrasound AUC (95% CI): 0.79 (CI not reported) Optimal cut-off threshold (if calculated): Not reported Other measures reported and conclusions: The fibrosis extraction method has great potential for diagnosing liver fibrosis using ultrasound.	
General limitations according to QUADAS II: Random or consecutive recruitment not reported. Indirectness: Patient exclusion criteria unclear and 5 patients had partial liver resection because of malignancy.	

Study	Yoneda 2008[162]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	102 (5 excluded due to unreliable TE measurement [all BMI>30] leaving 97 included)
Countries and Settings	Yokohama City University Hospital and Dokkyo Medical University, Japan
Funding	Grant-in-Aid from Ministry of Health, Labour and Welfare of Japan Ministry of Education, Culture, Sports, Science and Technology of Japan National institute of Biomedical Innovation
Age, gender, ethnicity	Age, mean (SD): 51.8±13.7; male/female: 40, 57; ethnicity: presumed Japanese; ALT (U/l): 80.0±62.3
Patient characteristics	Population: NASH. No evidence of hepatic decompensation. Inclusion: Presence of macrovesicular fatty change in hepatocytes with displacement of the nucleus to the edge of the cell. Exclusion: Hepatitis C, hepatitis B, autoimmune hepatitis, primary biliary hepatitis, sclerosing cholangitis, hemochromatosis, α1-antitrypsin deficiency, Wilson's disease, hepatic injury caused by substance abuse, current or past history of more than 20 g alcohol daily.

Study	Yoneda 2008[162]
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan): Performed on right lobe of the liver through intercostal spaces with patients lying in the dorsal decubitus position. Success rate of at least 60% or IQR <30% considered reliable. Presumed to have used appropriate probe for patients' BMI according to manufacturer's instructions (not reported).
Reference standard	Liver biopsy (Brunt scoring system, 4=cirrhosis) obtained with an 18-gauge needle. Specimens were stained with haematoxylin-eosin, reticulin and Masson trichrome stains. Minimum length 20 mm. Minimum 7 portal tracts. Analysed independently by 2 experience pathologists blinded to the results of the clinical data.
Time between index test and reference standard	Within 3 months
Prevalence of cirrhosis according to reference standard	9/97 (9.3%)
Target condition	Cirrhosis
Results: [TE] AUC (95% CI): 0.991 (CI not reported) Optimal cut-off threshold (if calculated): 17.5 unclear if published or calculated Threshold: 17.5 kPa (unclear if published or calculated) Sensitivity: 100 Specificity: 96.6 PPV: 75 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Other measures reported and conclusions: Very highly significant correlations between liver stiffness measure and serum hyaluronic acid and type IV collagen 7s domain. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II:	

Study	Yoneda 2008[162]
Random or consecutive recruitment not reported. Length of time between index test and reference standard not reported. Liver biopsy samples <10 portal tracts	

Study	Yoneda 2010[163]
Study type	Cross-sectional study
Number of studies (number of participants). Recruitment period.	1 study (n=54 consecutive patients with NAFLD, also a healthy control group n=10 not included in calculations of diagnostic accuracy). Recruitment between January 2008 and December 2008.
Countries and Settings	Yokohama City University Hospital
Funding	Supported in part by a Collaborative Development of Innovative Seeds program grant from the Japan Science and Technology Agency. A.N. supported in part by a grant from the National Institute of Biomedical Innovation. M.Y. supported by a grant from the Yokohama Foundation for Advancement of Medical Science
Age, gender, ethnicity	Age, mean (SD): 50.6 (13.7); male/female: 25/29; ethnicity: presumed Japanese; ALT (U/ml): men 66.4 (29.1), women 54.9 (33.1)
Patient characteristics	Population: Liver biopsy confirmed diagnosis of NAFLD. Inclusion: Undergone liver biopsy for the diagnosis and staging of NASH, histologic criterion for the diagnosis of NAFLD is the presence of macrovesicular fatty changes in hepatocytes, with displacement of the nucleus to the edge of the cell. Exclusion: History of hepatic disease, such as chronic hepatitis C or concurrent active hepatitis B (seropositive for hepatitis B surface antigen) infection, autoimmune hepatitis, primary biliary cirrhosis, sclerosing cholangitis, hemochromatosis, α1-antitrypsin deficiency, Wilson disease, or hepatic injury caused by substance abuse and current or past history of the consumption of more than 20 g of alcohol daily. No patients had any clinical evidence of hepatic decompensation, such as hepatic encephalopathy, ascites, variceal bleeding, or elevation of the serum bilirubin level to more than twofold the upper limit of normal.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan; optimal cut-off calculated): Measurements of the right lobe of the liver were performed through the intercostal spaces with the patient lying in the dorsal decubitus position with the right arm in maximal

Study	Yoneda 2010[163]
	abduction—the same site used for the ARFI sonoelastography measurements. Ten successful acquisitions were performed in each patient, and the median value was determined. Presumed to have used appropriate probe for patients' BMI according to manufacturer's instructions (not reported). ARFI (optimal cut-off calculated): Performed by using a Siemens Acuson S2000 US System (Mochida Siemens Medical System, Tokyo, Japan). ARFI sonoelastography was performed with a curved array US probe at 4 MHz for B-mode imaging. The right lobe of the liver was examined through the intercostal space with the patient lying in a dorsal decubitus position with the right arm in maximal abduction. An area where the liver tissue was at least 6 cm thick and free of large blood vessels was chosen. A measurement depth of 2 cm below the liver capsule was chosen. Ten successful acquisitions were performed in each patient, and the median value was determined.
Reference standard	Liver biopsy (Brunt scoring system, 4=cirrhosis): Specimens were obtained by using an 18-gauge needle biopsy apparatus (Pro-Mag; Medical Device Technologies, Gainesville, Fla) with a minimum of 7 portal tracts and a minimum length of 20 mm. Analysed independently by a pathologist with 27 years of experience in pathology who was unaware of the clinical data.
Time between index test and reference standard	TE and ARFI within 12 months of liver biopsy (mean 5.8 months [3.6]).
Prevalence of cirrhosis according to reference standard	6/54
Target condition Cirrhosis	
Results: ARFI AUC (95% CI): 0.976 Optimal cut-off threshold (if calculated): 1.90 m/s Threshold: 1.90 m/s (optimal) Sensitivity: 100 Specificity: 96 PPV: 75 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: 6 FP: Not reported FN: Not reported	

Study	Yoneda 2010[163]
	TN: 46 Results: Fibroscan AUC (95% CI): 0.998 Optimal cut-off threshold (if calculated): 16 kPa Threshold: 16 kPa (optimal) Sensitivity: 100 Specificity: 98 PPV: 86 NPV: 100 +ve/-ve likelihood ratios: Not reported TP: 6 FP: Not reported FN: Not reported TN: 47 Any complications associated with tests reported: Not reported
	General limitations according to QUADAS II: Time period between index test and reference standard up to 12 months. Biopsy length <25 mm.

Study	Zarski 2012[164]
Study type	Multicentre prospective study
Number of studies (number of participants). Recruitment period.	Multicentre. Enrolled n=590 (excluded n=78: 42 biopsies did not conform to criteria; 11 patients without blood sample; 9 patients with HBV co-infection; 5 patients with an excessive consumption of alcohol; 5 patients who received a treatment at the same time as the biopsy or less than 1 month before; 3 patients with unknown HCV status; 1 patient taking

Study	Zarski 2012[164]
	immunosuppressive treatment; 2 patients for whom a lot of data were missing). Fibrosis tests: n=436; Fibroscan: n=382 (not interpretable in 113 patients who were excluded from the analysis, some statistically significant differences were observed between patients included and those with failed Fibroscan). Recruitment November 2006–July 2008.
Countries and Settings	19 French academic hospitals, Fibrostar study cohort.
Funding	French agency for research on AIDS and viral hepatitis (ANRS)
Age, gender, ethnicity	Fibroscan (n=382) Fibrosis tests (n=436) Age, mean (SD): 50.9±10.6 51.2±10.9 Male/female: 60.7%/39.3% 61.5%/38.5% Ethnicity: Not stated Not stated ALT (U/l): 87.9±65.4 88.0±64.9
Patient characteristics	Population: Untreated chronic hepatitis C Inclusion: Time between liver biopsy and other diagnostic tests <3 months. No hepatitis C treatment in past 6 months. All patients had been referred for tests in order to make a decision on treatment strategy. CHC was confirmed by HCV-RNA polymerase chain reaction. Cirrhotic patients were compensated and asymptomatic at time of inclusion. Exclusion: Co-existing liver disease attributed to alcohol, hepatitis B, auto-immune hepatitis, primary biliary cirrhosis, hemochromatosis, alpha-1-antitrypsine deficiency, Wilson's disease, HIV infected, post-transplant.
Index test (including threshold and whether threshold pre-specified)	Transient elastography (Fibroscan) – measurements made on right lobe of liver, through intercostal spaces. At least 10 valid shots obtained/ IQR <30% deemed successful. FibroTest APRI FIB-4
Reference standard	Liver biopsy (METAVIR F4). Performed using Menghini's technique with a 1.6 mm needle, formalin-fixed in the centres and paraffin embedded. Sections were stained with hematoxylin-eosin-saffron and picrosirius red. Evaluated independently by 2 senior liver pathologists blind to clinical and biological data. Minimum length 15mm and/or at least 11 portal tracts (only 2.5% had <15 mm).
Time between index test and reference standard	<3 months (median 5 days, range 0–65 days)
Prevalence of cirrhosis according to	56/382 (14.7%)

Study	Zarski 2012[164]
reference standard	
Target condition	Cirrhosis
Results: FibroTest n=382 (AUC also provided in paper for n=436 sample but without sensitivity and specificity values) AUC (95% CI): 0.87 (0.82, 0.91) Optimal cut-off threshold (if calculated): Not reported Threshold: 0.74 (published) Sensitivity: 71.4% Specificity: 81.0% PPV: 39.2% NPV: 94.3% +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported APRI n=382 (AUC also provided in paper for n=436 sample but without sensitivity and specificity values) AUC (95% CI): 0.87 (0.82, 0.91) Optimal cut-off threshold (if calculated): Not reported Threshold: 2.0 (published) Sensitivity: 7.1 Specificity: 99.7 PPV: 80.0 NPV: 86.2 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported	

Study	Zarski 2012[164]
TN: Not reported FIB-4 n=382 (AUC also provided in paper for n=436 sample but without sensitivity and specificity values) AUC (95% CI): 0.84 (0.77, 0.90) Optimal cut-off threshold (if calculated): Not reported Threshold: Not reported Sensitivity: Not reported Specificity: Not reported PPV: Not reported NPV: Not reported +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported Fibroscan (n=382) AUC (95% CI): 0.93 (0.89, 0.96) Optimal cut-off threshold (if calculated): Not reported Threshold: 12.9 kPa (published) Sensitivity: 76.8 Specificity: 89.6 PPV: 55.8 NPV: 95.7 +ve/-ve likelihood ratios: Not reported TP: Not reported FP: Not reported FN: Not reported TN: Not reported	

Study	Zarski 2012[164]
Other measures reported and conclusions: Contrarily to blood tests, performance of Fibroscan was reduced due to uninterpretable results. Percentage of well classified patients and theoretically avoided liver biopsies according to one or a combination of two tests. For the diagnosis of cirrhosis, no combination was superior to the best blood tests or Fibroscan alone in the 'per-protocol' analysis (382 patients). However, when we considered the population of 436 patients (''intention to diagnose population'') the combination of Fibroscan plus a blood test markedly improved the percentage of well classified patients for the diagnosis of cirrhosis. Any complications associated with tests reported: Not reported	
General limitations according to QUADAS II: Up to 3 months between index test and reference standard. Large number of missing data for Fibroscan (and sensitivity and specificity data for fibrosis tests only provided for n=382 sample). Liver biopsy samples <25 mm.	