Gastroenterology

Transient Elastography, 2D Shear Wave Elastography and Collagen Proportionate Area Diagnose Alcoholic Liver Fibrosis and Cirrhosis with Excellent Accuracy

--Manuscript Draft--

Manuscript Number:	
Full Title:	Transient Elastography, 2D Shear Wave Elastography and Collagen Proportionate Area Diagnose Alcoholic Liver Fibrosis and Cirrhosis with Excellent Accuracy
Article Type:	Clinical - Liver
Section/Category:	Clinical
Corresponding Author:	Maja Thiele, M.D. Odense University Hospital Odense C, DENMARK
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Odense University Hospital
Corresponding Author's Secondary Institution:	
First Author:	Maja Thiele, M.D.
First Author Secondary Information:	
Order of Authors:	Maja Thiele, M.D.
	Sönke Detlefsen
	Linda Sevelsted Møller
	Bjørn Stæhr Madsen
	Janne Fuglsang Hansen
	Annette Dam Fialla
	Aleksander Krag
Order of Authors Secondary Information:	
Abstract:	BACKGROUND & AIMS: Alcohol abuse causes half of all deaths from cirrhosis in the Western world, but data regarding non-invasive diagnostic methods in alcoholic liver disease (ALD) are scarce. We therefore evaluated elastography to diagnose fibrosis and cirrhosis in ALD, with Ishak score and collagen proportionate area as reference. METHODS: During 24 months we included 200 patients with ongoing or prior alcohol abuse, but no known liver disease. We recruited two groups of patients: a group with high pre-test probability of cirrhosis from liver clinics, and a group with a low risk from municipal alcohol rehabilitation centres. All patients had same-day, transient elastography (TE, FibroScan), 2D shear wave elastography (2D-SWE, Supersonic Aixplorer) and liver biopsy. RESULTS: Elastography diagnosed significant fibrosis (Ishak≥3) and cirrhosis (Ishak≥5) with excellent accuracy in both risk groups (AUROC ≥0.92). There was no difference in diagnostic accuracy between TE and 2D-SWE. The optimal TE and 2D-SWE cut-offs were 9.6 and 10.2 kPa for significant fibrosis, and 19.7 and 16.4 kPa for cirrhosis. Negative predictive values were high in both groups, but the predictive value of a positive test for cirrhosis dropped from >66% in the high risk group, to approximately 50% in the low risk group. Alcoholic cholestasis, but not ongoing alcohol abuse, influenced liver stiffness. CPA correlated highly with Ishak grades and was excellent at detecting significant fibrosis and cirrhosis. CONCLUSIONS: Elastography is an excellent tool for detecting significant fibrosis and

	cirrhosis in ALD and can be used in populations with high and low prevalence of cirrhosis. Keywords: Supersonic shear imaging, FibroScan, non-invasive methods, diagnostic test.

TRANSIENT ELASTOGRAPHY, 2D SHEAR WAVE ELASTOGRAPHY AND COLLAGEN PROPORTIONATE AREA DIAGNOSE ALCOHOLIC LIVER FIBROSIS AND CIRRHOSIS WITH EXCELLENT ACCURACY

Short title: Elastography in alcoholic liver disease

Authors: Maja Thiele[1,2,3], Sönke Detlefsen[3,4], Linda Sevelsted Møller[5], Bjørn Stæhr Madsen[1,2,3], Janne Fuglsang Hansen[3,6], Annette Dam Fialla[1,3,7], Aleksander Krag[1,3]

Author affiliations: 1: Department of Gastroenterology and Hepatology, Odense University Hospital, Odense, Denmark. 2: OPEN Odense Patient data Explorative Network, Odense University Hospital, Odense, Denmark. 3: Institute of Clinical Research, University of Southern Denmark, Odense, Denmark. 4: Department of Pathology, Odense University Hospital, Odense, Denmark. 5: Department of Medicine, Odense University Hospital, Svendborg, Denmark. 6: Department of Infectious Diseases, Odense University Hospital, Odense, Denmark. 7: Department of Medicine, Hospital of Southwest Jutland, Esbjerg, Denmark.

Grant support: The study was investigator initiated and partly funded by the Danish National Advanced Technology Foundation and Innovation Fund Denmark. The Supersonic Aixplorer and FibroScan XL-probe were acquired with grants from the A. P.

Moeller Foundation and Toyota Foundation. Working grants from University of Southern Denmark, Odense University Hospital and Region of Southern Denmark supports doctors M.T., B.S.M., J.F.H. and A.D.F..

Abbreviations: 2D-SWE, 2D shear wave elastography; ALD, alcoholic liver disease; BMI, body mass index; CPA, Collagen Proportionate Area; GGT, Gamma-Glutamyltransferase; HCV, chronic viral hepatitis C; ITD, intention-to-diagnose; kPa, kiloPascal; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristics; TE, transient elastography.

Correspondence: Maja Thiele, Department of Gastroenterology and Hepatology, Odense University Hospital, Sdr. Boulevard 29, entrance 128, DK-5000 Odense C, Denmark. Email: maja.thiele@rsyd.dk. Phone: +4565412752. Fax: +4566111328.

Disclosures: Nothing to disclose.

Author contributions: A.K. conceptualised the study. M.T. and A.K. designed the study. All authors acquired data for the study. M.T. analysed the data. M.T., S.D. and A.K. interpreted the data. M.T. drafted the manuscript. All authors revised the work for important intellectual content. All authors have approved the final manuscript.

Word count: 5703

BACKGROUND & AIMS: Alcohol abuse causes half of all deaths from cirrhosis in the Western world, but data regarding non-invasive diagnostic methods in alcoholic liver disease (ALD) are scarce. We therefore evaluated elastography to diagnose fibrosis and cirrhosis in ALD, with Ishak score and collagen proportionate area as reference.

METHODS: During 24 months we included 200 patients with ongoing or prior alcohol abuse, but no known liver disease. We recruited two groups of patients: a group with high pre-test probability of cirrhosis from liver clinics, and a group with a low risk from municipal alcohol rehabilitation centres. All patients had same-day, transient elastography (TE, FibroScan), 2D shear wave elastography (2D-SWE, Supersonic Aixplorer) and liver biopsy.

RESULTS: Elastography diagnosed significant fibrosis (Ishak≥3) and cirrhosis (Ishak≥5) with excellent accuracy in both risk groups (AUROC ≥0.92). There was no difference in diagnostic accuracy between TE and 2D-SWE. The optimal TE and 2D-SWE cut-offs were 9.6 and 10.2 kPa for significant fibrosis, and 19.7 and 16.4 kPa for cirrhosis. Negative predictive values were high in both groups, but the predictive value of a positive test for cirrhosis dropped from >66% in the high risk group, to approximately 50% in the low risk group. Alcoholic cholestasis, but not ongoing alcohol abuse, influenced liver stiffness. CPA correlated highly with Ishak grades and was excellent at detecting significant fibrosis and cirrhosis.

CONCLUSIONS: Elastography is an excellent tool for detecting significant fibrosis and cirrhosis in ALD and can be used in populations with high and low prevalence of cirrhosis.

Keywords: Supersonic shear imaging, FibroScan, non-invasive methods, diagnostic test.

Every year alcoholic liver disease (ALD) cause 500,000 deaths and cost 14.5 million disability-adjusted life years worldwide.[1, 2] Alcohol now accounts for 50% of all deaths from liver cirrhosis.[2] Despite the vast impact on health and society, ALD research is awarded less than three percent of all liver disease funding.[3] That is probably why accurate tools for diagnosis of liver fibrosis and early cirrhosis in ALD are still lacking, and most ALD patients have disproportionately more advanced disease than other liver disease patients.[4] For example, only 25% of patients with alcoholic cirrhosis are diagnosed with compensated disease.[5] Early diagnosis of cirrhosis and its precursor, significant fibrosis, is the key to improve survival in ALD. In contrast to popular belief, early diagnosis motivates more patients to quit drinking and thereby prevents disease progression.[6]

Liver stiffness by transient elastography (TE) is used for non-invasive diagnosis of fibrosis and cirrhosis in chronic hepatitis C (HCV) patients,[7] but only few studies have investigated alcoholic patients to date.[8-10] Additionally, some researchers are concerned that active alcohol abuse may increase liver stiffness, which would limit the applicability of elastography.[11, 12] Consequently, a Cochrane review and the European guidelines on non-invasive liver disease tests recently stressed the need for well-powered, high-quality studies.[13, 14]

Real-time 2D shear wave elastography (2D-SWE), is a novel technique combining real-time visualization of multiple shear-waves with traditional ultrasound imaging (Appendix) with favourable results.[15-18] Some even suggest that 2D-SWE may outperform TE.[19] However, 2D-SWE has never been evaluated for ALD. In lack of a widely recognized histological scoring system for fibrosis in ALD, liver biopsy, “the imperfect gold standard”, is compromised even further.[20] A quantitative measurement of the amount of collagen by digital image analysis of liver biopsies with quantification of the collagen proportionate

area (CPA) is a plausible solution. CPA correlates with fibrosis scores in chronic viral hepatitis and accurately predicts outcomes in cirrhosis, but has never been used for fibrosis quantification in ALD.[21-23]

We therefore aimed to determine the diagnostic accuracy of liver stiffness by TE and 2D-SWE for the diagnosis of significant alcoholic fibrosis and cirrhosis, using liver biopsy as gold standard. Secondary objectives were to compare the diagnostic accuracy of the two elastography methods and to compare CPA with Ishak fibrosis scoring.

PATIENTS AND METHODS

The study was a prospective, cross-sectional, biopsy controlled, single center study. The study protocol was approved by the Regional Ethics Committee (study id S-20120071). The study is reported according to the Liver-FibroSTARD checklist (Appendix).[24]

Patients

All patients gave written informed consent to participate. To ensure that our study covered the entire spectrum of ALD, we consecutively recruited patients from two populations: One with a high- and one with a low pre-test risk of cirrhosis. The high risk group consisted of patients referred for investigation at three liver clinics in Southern Denmark whereas the low risk group consisted of alcohol overusing individuals from municipal alcohol rehabilitation centres and the Danish national public health portal.

Inclusion criteria were: (i) Prior or current chronic alcohol overuse defined as more than 24 grams of alcohol/day for women and 36 g/day for men for >1 year, (ii) age 18-75 years. The exclusion criteria were: (i) Decompensation evidenced by clinical obvious ascites, hepatic encephalopathy, known gastro-intestinal varices or prior variceal bleeding (ii) concurrent liver disease other than alcoholic, (iii) cancer or other debilitating disease with a life expectancy of less than one year, (iv) contraindications for percutaneous liver biopsy, (v) severe alcoholic hepatitis with Maddrey Discriminant Function ≥32, (vi) right heart failure or cholestasis evidence by ultrasound, (vii) human immunodeficiency virus positivity, (viii) ongoing substance abuse other than alcohol, (ix) inability to comply with the study protocol.

Investigations at Inclusion and Blinding

All investigations were performed on the same day, after an overnight fast, in an operator-blinded manner and according to standard operating procedures. During inclusion, patients were investigated with: liver biopsy, TE, 2D-SWE, abdominal ultrasonography, standard biochemical testing, questionnaires on alcohol use, medical history and demographic data. TE and 2D-SWE operators measured liver stiffness independently, blinded to the other's examination. Blinding was abandoned only if the specialist nurse could not obtain a valid TE. In that case, TE was done by the investigator after 2D-SWE. We evaluated alcohol dependency with the CAGE questionnaire.[25] Additionally, patients answered detailed, standardised questions on current and past alcohol consumption. Questions were asked at the end of the inclusion visit, when patient and investigator had familiarized and by assuring the patient of confidentiality. When performed in this manner, questionnaires for self-reporting of drinking habits has high validity and reproducibility.[26]

Elastography

Three experienced specialist nurses with more than 500 scans performed TE (FibroScan 502 Touch, Echosens, France) while two operators (MT or BSM) performed 2D-SWE (Aixplorer, Supersonic Imagine, France). Both operators are experienced ultrasonographers and trained 2D-SWE before the study commenced. Liver stiffness was acquired intercostal, during breath hold, with the patient in the supine position and the right arm above his/her head.

TE was considered reliable if 10 measurements was acquired with a IQR below 30% of the median; except for situations with median <7.1 kPa, where higher IQR/median ratios were accepted.[27]

A 2D-SWE measurement was valid when it fulfilled a set of quality criteria: (i) temporal stability of the viscoelasticity map for at least 3 seconds before image acquisition, (ii) spatial quality evidenced by a homogeneous colour in the region of interest, (iii) a region of interest of at least 15mm and (iv) a standard deviation/mean-ratio ≤30%. We obtained up to three separate 2D-SWE measurements and reported results as the mean of the total number of valid measurements.[28, 29]

Liver Biopsy and Histologic Evaluation

Percutaneous suction needle liver biopsy was performed in the same intercostal space as the elastography (17G Menghini needle, Hepafix, Braun, Germany). Following biopsy, the samples were immediately stored in formalin 4% and paraffin-embedded on 5 µm liver sections. Sirius red stained sections were used for the detection of collagen fibers.

One experienced liver pathologist (SD) assigned CPA and Ishak fibrosis grade to all biopsies with no knowledge of elastography or clinical data. For each specimen, he assigned Ishak scores before the CPA image analysis. A biopsy was of adequate quality if it had a length of at least 10 mm or at least six portal tracks or presence of regeneration nodules.

Tissue sections for CPA measurement were digitized using a Nanozoomer 2.0 HT (Hamamatsu Photonics, Hamamatsu, Japan) and analyzed using the image analysis tool VISIOmorph version 4.3.6 (Visiopharm, Hørsholm, Denmark), which was developed and validated prior to the study (Appendix).

Statistical Analyses

Summary statistics were used to describe patient characteristics. Correlations were investigated with linear and logistic regression analyses using backwards, step-wise elimination. The diagnostic accuracy, pre-test (sensitivity, specificity) and post-test probabilities (positive and negative predictive values, post-test odds) of TE, 2D-SWE and CPA for detecting significant fibrosis (Ishak score ≥3) and cirrhosis (Ishak ≥5) were calculated using receiver operating characteristics (ROC). Optimal cut-off values were decided by maximizing the Youden Index. ROC curves were compared using a non-parametric test for ROC curve comparison. Diagnostic test calculations included both per-protocol and intention-to-diagnose results. For the intention-to-diagnose analyses, cases without a valid elastography were considered as false negatives. The statistical software package STATA 13.1 (College Station, TX, USA) were used for all calculations.

RESULTS

Patients

From May 2013 to April 2015 we included 206 patients. Six patients were excluded after liver biopsy due to of competing liver disease or incomplete biopsy (Figure 1). Of the final 200 patients, the majority were male and the mean age was 55 years (Table 1).

Thirty-seven patients had liver cirrhosis at inclusion. Patients with cirrhosis had early stage disease with a median Child-score of 6 (IQR 2) and the following distribution according to Child-Pugh class: 20 class A, 15 B and two C. Ten patients had small amounts of ascites on ultrasound, not clinically evident. In general, included patients presented free of symptoms, without a history of liver disease and most had normal or only slightly affected liver function tests.

More than half of patients were abstaining from alcohol at inclusion. Of the alcohol abstainers, 70% (74/105) quit drinking less than one year before study inclusion (median 10 weeks of abstinence, IQR 4-20). In contrast, 25% (50/200) had an ongoing overuse (mean alcohol consumption 97±77 g/week) and 5% (11/200) drank more than 120 g/day at inclusion.

Applicability of Transient and 2D Shear Wave Elastography

Elastography was successful in 97% of patients (193/200). The overall median liver stiffness with TE and 2D-SWE was 7.1 kPa (IQR 13.8; range 2.6-75.0) and 8.3 kPa (IQR 9.9; range 4.2-80.4). TE and 2D-SWE covaried in a linear manner (Appendix).

Due to equipment maintenance, nine patients missed TE and 20 patients missed 2D-SWE. TE failed in seven and was unreliable in two patients. All nine patients were overweight

and were investigated before the FibroScan XL-probe was available. 2D-SWE failed in four and was unreliable in three patients. Six patients were overweight and the seventh had a skin-capsule distance above three centimeters. In summary, the TE failure rate was 5%, versus 4% for 2D-SWE (P=0.102).

For each elastography method, we investigated which factors influenced liver stiffness in a multivariate regression model. The factors included in the model were chosen based on previous studies.[12, 30, 31] The model included degree of fibrosis, liver blood tests, drinking pattern (ongoing alcohol overuse, abstinence, number of drinks in the week before inclusion), steatosis evidenced by a hyperechoic liver, smoking, gender, age, BMI, heart rate, mean arterial pressure and operator. Only Ishak fibrosis grades 3-6 and bilirubin were independent predictors of both TE and 2D-SWE, while platelet count and albumin also independently predicted TE (all P<0.050).

To further investigate the role of alcohol on liver stiffness, we compared liver stiffness of the patients with an ongoing daily alcohol abuse above ten drinks per day, with the patients drinking below two drinks per day. There was no difference in liver stiffness by TE or 2D-SWE between heavy and minimal drinkers, both with and without adjustment for degree of fibrosis (all P>0.100).

Elastography to Diagnose Liver Fibrosis and Cirrhosis in Alcoholic Liver Disease

Liver stiffness increased with increasing degree of fibrosis (Table 2 and Figure 2). In the per-protocol analysis, TE and 2D-SWE diagnosed significant fibrosis and cirrhosis with excellent diagnostic accuracies (Table 3). There was no difference in diagnostic accuracies of TE versus 2D-SWE (Figure 3).

The optimal cut-offs for diagnosing significant fibrosis were 9.6 and 10.2 kPa for TE and 2D-SWE. The optimal cut-offs for diagnosing cirrhosis were: 19.7 kPa for TE and 16.4 kPa for 2D-SWE (Table 3).

For cirrhosis, negative predictive values (NPV) were high. Accordingly, the risk of missing a diagnosis of cirrhosis because of false negatives was negligible: Less than 2% of patients (2/145) with liver stiffness below the cirrhosis cut-offs had cirrhosis on biopsy.

In contrast, positive predictive values for cirrhosis (PPV) were considerably lower, with a equivalent high risk of misclassifications due to of false positives: 35% of patients (19/55) with liver stiffness above the cirrhosis cut-offs, did not have cirrhosis on biopsy. In a logistic regression model that included Ishak score, biopsy length, biopsy fragments, alcohol drinking pattern and liver blood tests, only Ishak and Gamma-Glutamyltransferase (GGT) predicted false positive liver stiffness (coefficient 0.002, 95% CI 0.000-0.004, P=0.044). All 19 false positive patients scored Ishak 3 or 4 and 74% had elevated levels of GGT.

To suggest how to stratify patients for further investigation based on their liver stiffness, we determined a rule-in cut-off by optimising the positive likelihood ratio, and a rule-out cut-off by optimising the negative likelihood ratio. Distributional plots with rule-in and rule-out cut-off values were constructed for the diagnosis of fibrosis or cirrhosis (Figure 4 and Appendix). Liver stiffness above the rule-in cut-off suggests that the diagnosis is certain and a liver biopsy can be avoided, while liver stiffness below the rule-out cut-off suggests that the patient is disease-free and examinations can be completed. Liver stiffness between cut-offs warrants further diagnostic investigations.

Intention-to-diagnose analysis

To evaluate the stability of our results, we performed an intention-to-diagnose (ITD) analysis in which per-protocol cut-off values were reused, failures and unreliable results were included as false negatives and missing equipment were excluded.

Diagnostic accuracies decreased to less than 0.85 according to the ITD protocol. The ITD diagnostic accuracy of TE and 2D-SWE for the diagnosis of Ishak ≥3 were: TE = AUROC 0.82 (95% CI 0.76-0.88), and 2D-SWE = AUROC 0.84 (0.78-0.89). And for the diagnosis of cirrhosis: TE = AUROC 0.82 (0.75-0.89), and 2D-SWE = AUROC 0.84 (0.77-0.91). The ITD analysis primarily affected the sensitivities of elastography techniques. ITD sensitivity, specificity, PPV and NPV for the diagnosis of significant fibrosis were 74%, 90%, 84% and 83% for TE; and 75%, 93%, 89% and 83% for 2D-SWE. For the diagnosis of cirrhosis, the same ITD test probabilities were: 74%, 89%, 61% and 94% for TE, and 77%, 91%, 67% and 95% for 2D-SWE.

Differences between high and low pre-test risk groups

Of the total 200 patients, 129 patients belonged to the group with a high pre-test risk of cirrhosis and 71 belonged to the low pre-test risk group. The two risk groups differed on a number of parameters, including age and liver function tests (Table 1). The drinking pattern also differed between risk-groups: Despite similar duration of alcohol overuse (median 11-20 years, P=0.19), patients from the low risk group were to a larger extent abstaining from alcohol at inclusion, but reported a heavier alcohol abuse before inclusion.

While the prevalence of cirrhosis was 26% (34/129) in the high pre-test risk group, only 4% (3/71) in the low risk group had cirrhosis.

In the high and the low risk group, TE and 2D-SWE performed equally well with AUROCs above 0.92 and no difference in diagnostic accuracy between techniques (all P>0.30). Cut-off values for significant fibrosis were considerably lower in the low risk group, than in the high risk group.

Given the different prevalences of fibrosis and cirrhosis in the two groups, PPVs differed substantially. The PPV of TE and 2D-SWE for diagnosing fibrosis were 88% and 92% in the high risk group; compared to 70% and 49% in the low risk group. And for diagnosing cirrhosis, the PPVs decreased from 69% (TE) and 73% (2D-SWE) in the high risk group, to 59% and 49% in the low risk group. The NPV was high regardless of risk-groups.

Collagen Proportionate Area

The median CPA was 4.1% (IQR 4.4; range 0.5-39.9). CPA correlated highly (rho 0.76, *P*<0.001) with Ishak grades (Figure 2). Likewise, CPA was excellent at diagnosing significant fibrosis and cirrhosis (Ishak ≥3, AUROC 0.90, 95% CI 0.86-0.95; Ishak ≥5, AUROC 0.97, 0.94-0.99). The optimal CPA cut-off value for detection of Ishak ≥3 was 4.7%. With this cut-off, the sensitivity and specificity were 82% and 89% with PPV and NPV of 84% and 87%. The optimal CPA cut-off for detection of cirrhosis was 8.4%, and corresponding sensitivity, specificity, PPV and NPV were 88%, 92%, 71% and 97%.

In a uni- and multivariate linear regression model, CPA covaried with Ishak fibrosis grades 2-6, while no other biopsy descriptors predicted CPA.

In the 37 patients with cirrhosis, CPA increased from 11.6% (IQR 6.3, range 5.1-18.9%) for Ishak grade 5, to 20.0% (IQR 12.2, range 5.5-39.9%) for Ishak grade 6 (P=0.002). Additionally, CPA correlated with Child-score independent of Ishak, and only CPA

correlated with presence of ascites (correlation coefficient for Child-score = 0.12, 95% CI 0.02-0.23, P=0.022; and for ascites, coefficient = 0.03, 0.00-0.06, P=0.048).

CPA also correlated with liver stiffness independent of Ishak (TE coefficient = 0.23, 0.20-0.29, P<0.001; 2D-SWE coefficient = 0.39, 0.33-0.45, P<0.001).

DISCUSSION

This biopsy controlled study strongly supports elastography as a clinically useful, non-invasive tool for assessing fibrosis in ALD. For the first time, we also show that 2D-SWE has excellent diagnostic accuracy in patients with ALD and that CPA is a valid supplement to standard histological scoring systems. Finally, we contradict the idea, that active alcohol abuse causes false positive elastography results. Our results thereby fill a longtime knowledge-gap.[14]

The study comprises 200 alcohol-overusing individuals from the entire spectrum of alcoholic liver fibrosis, and cover both primary care and liver clinics. The dual recruitment strategy ensured evaluation of elastography in a background population of asymptomatic alcohol overusers outside a hospital setting. Such patients are not found in other diagnostic test studies. In comparison, the hitherto largest prospective study on elastography in ALD included 54% cirrhotics.[8] A distribution of patients with fibrosis severity across the entire spectrum decrease the risk of over-estimating diagnostic test probabilities due to spectrum bias.[32] We additionally minimized spectrum bias by excluding patients with obvious ascites, varices or hepatic encephalopathy. Patients were recruited from two different settings in which elastography may be used as a screening tool for chronic ALD: In the liver clinic with referral of older, sicker patients that has a high probability of cirrhosis; and the primary health care system with a high number of younger, asymptomatic alcohol overusing individuals of whom only a small percentage have cirrhosis. However, an independent validation cohort is needed to verify the suggested cut-off values for significant alcoholic fibrosis and cirrhosis. Future studies will unravel whether elastography can predict clinical outcomes and mortality in ALD.

Our analyses underline the importance of addressing disease prevalence when interpreting test results: while more than two-thirds of patients with a liver stiffness above the cirrhosis cut-off in the high pre-test risk group actually had cirrhosis on liver biopsy, the same was true for only half of patients in the low pre-test risk group. In contrast, the negative predictive value for cirrhosis was close to 100% in both groups.

The low positive predictive values for cirrhosis may also be caused by sampling error, because biopsy is an imperfect gold standard. However, GGT rather than biopsy length correlated with false positive liver stiffness measurements. Likewise, bilirubin correlated with liver stiffness independently of liver fibrosis, while alcohol alone did not influence liver stiffness. We therefore suggest, that alcoholic cholestasis was the predominant cause of false positive misclassifications in our cohort, and that patients should not be excluded from liver stiffness evaluation simply because of ongoing alcohol abuse.

We did not find any difference in diagnostic accuracy or applicability of TE and 2D-SWE, which suggests that both methods may be used in the non-invasive work-up of alcohol overusing patients. Despite a strong liniear correlation, it is important to emphasize, that the values from one method cannot be directly translated to the other. And while TE equipment is cheaper and can be operated by specialist nurses without ultrasound training, the main advantage of 2D-SWE is the possibility of a concurrent conventional ultrasound examination.

The ALD cut-offs in our study are substantially higher than for HCV fibrosis and HCV cirrhosis, which is most likely because fibrosis differ between etiologies.[7, 33] While fibrosis in ALD is perivenular and pericellular with central expansion, HCV fibrosis starts periportally and extends in a portal manner.[20] Since Ishak is developed for HCV, there is naturally a larger amount of fibrosis in alcoholic patients compared to HCV patients, within

the same Ishak stage and consequently higher liver stiffness. For example, our TE cut-off of 19.7 kPa for the diagnosis of alcoholic cirrhosis is close to the cut-off proposed for clinically significant portal hypertension in HCV patients.[34] This underlines the importance of single-etiology assessment of elastography performance and may explain why our AUROCs are higher than studies with mixed etiologies.[35]

Difficulties in semiquantitative histological staging of fibrosis may be overcome by using an automated method such as CPA. CPA measures the amount of collagen in a continuous manner and is probably more stable to observer variance, than classical scoring systems.[21] CPA may therefore be a useful additional tool for the liver pathologist when evaluating liver biopsies from patients suspected of ALD. Additionally, CPA has been suggested as a tool to subclassify patients with HCV cirrhosis.[23] Our results support CPA to subdivide alcoholic cirrhosis, but the number of patients with cirrhosis in our cohort is too small to draw any definite conclusions.

In conclusion, liver stiffness measurement by elastography is a reliable marker of significant fibrosis and cirrhosis in ALD. When applying cut-off values to stratify the risk of fibrosis and cirrhosis in patients with ALD, the prevalence in the screened population must be taken into account.

Acknowledgments:

The specialist nurses at Odense University Hospital outpatient liver clinic contributed immensely to this study: Trine Møller, Charlotte Damby Jensen, Minna Ingham and Birgit Mathiasen.

We also wish to thank the entire staff at Odense Municipality Alcohol Rehabilitation Center for their contribution to the patient recruitment; and management and colleagues at the University Hospitals in Odense, Svendborg and Esbjerg for support to the study.

FIGURES

Figure 1. Study flow diagram.

Figure 2. Distribution of (A) Collagen Proportionate Area, (B) TE and (C) 2D-SWE according to Ishak fibrosis stage in 200 alcohol overusing patients.

Figure 3. Receiver Operating Characteristic curves for transient elastography (TE) and 2D shear wave elastography (2D-SWE) to diagnose significant fibrosis (≥Ishak 3, AUROC = 0.946 for TE and 0.942 for 2D-SWE) and cirrhosis (≥Ishak 5, AUROC = 0.958 for TE and 0.955 for 2D-SWE).

Figure 4. Distributional plots of liver stiffness measurements with TE and 2D-SWE according to significant fibrosis and cirrhosis, with rule-in and rule-out cut-offs marked as a line. (A) and (B) display the distribution of liver stiffness in patients with-or-without significant fibrosis, while (C) and (D) display the distribution of liver stiffness in patients with-or-without cirrhosis.

(A) TE>13.7 kPa can be used to rule in significant fibrosis, as only one patient with Ishak 0-2 had a higher liver stiffness. TE<5.8 kPa can be used to rule out significant fibrosis, as only one patient with Ishak ≥3 had lower liver stiffness.

(B) 2D-SWE>14.6 kPa can be used to rule in significant fibrosis, as no one with Ishak 0-2 had higher liver stiffness. 2D-SWE<7.0 kPa can be used to rule out significant fibrosis, as only one patient with Ishak ≥3 had lower liver stiffness.

(C) TE>51.4 kPa can be used to rule in cirrhosis, as only three patients without cirrhosis had higher liver stiffnesses. TE<20.2 kPa can be used to rule out cirrhosis, as only one patient with cirrhosis had lower liver stiffness.

(D) 2D-SWE>27.3 kPa can be used to rule in cirrhosis, as only four patients without cirrhosis had higher liver stiffnesses. 2D-SWE<12.1 kPa can be used to rule out cirrhosis, as only one patient with cirrhosis had lower liver stiffness.

Appendix:

1. Principle of 2D Shear Wave Elastography (2D-SWE).
2. Description of the VISIOmorph tool to measure collagen proportionate area
3. Correlation between transient elastography and 2D shear wave elastography.
4. Table with diagnostic values for rule-in and rule-out cut-off values.
5. STARD checklist.
6. Liver-FibroSTARD checklist.

TABLES

Table 1. Patient characteristics

	All	High Risk Group	Low Risk Group	P*
Number of patients	200	129 (65%)	71 (35%)	
Male gender	148 (74%)	96 (74%)	51 (72%)	0.716
Age	55 ±11	57 ±10	50 ±12	<0.001
Heart rate (bpm)	76 ±16	79 ±17	71 ±12	0.001
Mean arterial pressure (mmHg)	96 ±13	96 ±13	96 ±13	0.969
ALT (U/L)	33 ±29	36 ±31	24 ±22	<0.001
Alkaline phosphatase (U/L)	91 ±59	103 ±78	79 ±33	<0.001
Bilirubin (µmol/L)	11 ±9	13 ±13	9 ±5	<0.001
Albumin (g/L)	41 ±5	40 ±6	43 ±4	<0.001
Platelet count (10^9/L)	230 ±108	216 ±145	259 ±92	<0.001
CRP (mg/L)	5 ±6	6 ±8	4 ±6	0.016
Ferritin (µg/L)	135 ±245	199 ±302	76 ±103	<0.001
Drinking pattern and medical history				
Abstaining from alcohol	105 (53%)	53 (41%)	51 (73%)	<0.001
Ongoing alcohol overuse[1]	51 (26%)	42 (33%)	9 (13%)	0.002
Moderate drinkers	44 (22%)	33 (26%)	11 (14%)	0.002
Average alcohol intake during overuse (grams/day)	]180 ±216	144 ±216	240 ±216	0.009
CAGE score	3 ±2	3 ±3	4 ±1	<0.001
Prior hepatic events[2]	46 (23%)	42 (33%)	4 (6%)	<0.001
Any comorbidity	136 (68%)	98 (76%)	38 (54%)	<0.001
Smokers /prior smokers /non smokers	117 (63%) /48 (26%) /20 (11%)	71 (56%) /44 (35%) /12 (9%)	46 (68%) /14 (20%) /8 (12%)	0.204
Histology				
Ishak fibrosis score 0/1/2/3/4/5/6	12/52/51/35/13/15/22	4/15/37/28/11/13/21	8/37/14/7/2/2/1	<0.001
Collagen Proportionate Area	6.4 ±6.8	8.1 ±7.8	3.4 ±2.5	<0.001
Biopsy length (mm)	30.0 ±9.4	29.9 ±9.8	30.1 ±8.9	0.921

*P-value reports equality test between high- and low risk groups. Count numbers are stated as n(%). Summary statistics are reported as mean±SD and median±IQR for normal and non-normal distributed data. 1: >24g/d for women and >36g/d for men. 2: Any case of upper gastrointestinal bleeding, variceal bleeding, alcoholic hepatitis, icterus, ascites, hepatic encephalopathy, hepatorenal syndrome and/or spontaneous bacterial peritonitis.

Table 2. Median liver stiffness according to degree of fibrosis

Transient elastography	Ishak 0	Ishak 1	Ishak 2	Ishak 3	Ishak 4	Ishak 5	Ishak 6
Median (kPa)	4.1	5.4	5.9	11.5	21.6	27.0	65.2
IQR (kPa)	2.0	1.8	3.8	12.3	20.7	26.2	22.1
Min-max (kPa)	2.6-5.8	3.1-13	3.6-29.9	4.9-70.6	8.6-75	11.8-75	20.2-75
Number of patients	12	50	47	29	13	13	18
P-values*	-	0.038	0.003	<0.001	0.049	0.143	0.027
2D Shear Wave Elastography	Ishak 0	Ishak 1	Ishak 2	Ishak 3	Ishak 4	Ishak 5	Ishak 6
Median (kPa)	5.8	6.0	7.6	13.3	18.0	25.7	35.7
IQR (kPa)	1.0	1.0	2.9	11.6	12.1	15.5	14.1
Min-max (kPa)	4.9-8.4	4.2-12.0	5.6-14.6	6.5-48.8	7.2-51.5	9.7-43.4	12.1-80.4
Number of patients	12	48	44	26	10	11	22
P-values*	-	0.267	<0.001	<0.001	0.163	0.181	0.056

*Non-parametric test of equality between single grades of fibrosis: Ishak 0 versus Ishak 1; Ishak 1 versus Ishak 2, etc.

Table 3. Diagnostic test results, per protocol analyses

Overall	**Fibrosis (Ishak ≥3)**		**Cirrhosis (Ishak ≥5)**	
	Transient Elastography	2D Shear Wave Elastography	Transient Elastography	2D Shear Wave Elastography
Prevalence	85/200 (43%)		37/200 (19%)	
AUROC	0.95 (0.92-0.98)	0.94 (0.91-0.97)	0.96 (0.93-0.98)	0.95 (0.92-0.98)
Optimal cut-off value	9.6 kPa	10.2 kPa	19.7 kPa	16.4 kPa
Correctly classifies	87%	88%	93%	93%
Sensitivity	84%	83%	97%	94%
Specificity	91%	93%	90%	91%
Pos. predictive value	87%	90%	69%	71%
Neg. predictive value	88%	88%	99%	99%
Pre-test odds	0.74		0.23	
Post-test odds (+)	6.67	9.07	2.20	2.49
Post-test odds (-)	0.13	0.14	0.01	0.02
High Risk Group	Transient Elastography	2D Shear Wave Elastography	Transient Elastography	2D Shear Wave Elastography
Prevalence	73/129 (57%)		34/129 (26%)	
AUROC	0.94 (0.90-0.98)	0.94 (0.89-0.98)	0.94 (0.90-0.98)	0.93 (0.87-0.98)
Optimal cut-off value	10.5 kPa	10.1 kPa	19.7 kPa	16.4 kPa
Correctly classifies	85%	89%	91%	91%
Sensitivity	86%	89%	97%	93%
Specificity	84%	89%	85%	88%
Pos. predictive value	88%	92%	69%	73%
Neg. predictive value	82%	87%	99%	97%
Pre-test odds	1.30		0.36	
Post-test odds (+)	7.10	10.96	2.23	2.75
Post-test odds (-)	0.22	0.15	0.02	0.03
Low Risk Group	Transient Elastography	2D Shear Wave Elastography	Transient Elastography	2D Shear Wave Elastography
Prevalence	12/71 (17%)		3/71 (4%)	
AUROC	0.92 (0.82-1.00)	0.92 (0.84-0.99)	0.97 (0.93-1.00)	0.98 (0.95-1.00)
Optimal cut-off value	7.8 kPa	7.2 kPa	18.8 kPa	15.8 kPa
Correctly classifies	87%	86%	98%	98%
Sensitivity	82%	92%	100%	100%
Specificity	93%	81%	97%	96%
Pos. predictive value	70%	49%	59%	49%
Neg. predictive value	96%	98%	100%	100%
Pre-test odds	0.20		0.04	
Post-test odds (+)	2.37	0.97	1.46	0.97
Post-test odds (-)	0.04	0.02	0.00	0.00

AUROC: Area Under the Receiver Operating Characteristics curve. kPa: kilo Pascal

REFERENCES

1. Rehm J, Samokhvalov AV, Shield KD. Global burden of alcoholic liver diseases. J Hepatol 2013;59:160-168.
2. WHO. Global status report on alcohol and health. http://www.who.int/substance_abuse/publications/global_alcohol_report/en/, 2014.
3. Bataller R. Personal communication, 2015.
4. Wong RJ, Aguilar M, Cheung R, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology 2015;148:547-55.
5. Fialla AD, de Muckadell OBS, Touborg Lassen A. Incidence, etiology and mortality of cirrhosis: a population-based cohort study. Scand J Gastroenterol 2012;47:702-709.
6. Sheron N, Moore M, O'Brien W, et al. Feasibility of detection and intervention for alcohol-related liver disease in the community: the Alcohol and Liver Disease Detection study (ALDDeS). Br J Gen Pract 2013;63:e698-705.
7. Castéra L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology 2005;128:343-350.
8. Nahon P, Kettaneh A, Tengher-Barna I, et al. Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. J Hepatol 2008;49:1062-8.
9. Nguyen-Khac E, Chatelain D, Tramier B, et al. Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: prospective comparison with seven non-invasive laboratory tests. Aliment Pharmacol Ther 2008;28:1188-98.
10. Mueller S, Millonig G, Sarovska L, et al. Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. World J Gastroenterol 2010;16:966-72.
11. Trabut JB, Thepot V, Nalpas B, et al. Rapid decline of liver stiffness following alcohol withdrawal in heavy drinkers. Alcohol Clin Exp Res 2012;36:1407-11.
12. Bardou-Jacquet E, Legros L, Soro D, et al. Effect of alcohol consumption on liver stiffness measured by transient elastography. World J Gastroenterol 2013;19:516-22.
13. Pavlov CS, Casazza G, Nikolova D, et al. Transient elastography for diagnosis of stages of hepatic fibrosis and cirrhosis in people with alcoholic liver disease. Cochrane Database Syst Rev 2015;1:CD010542.
14. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015.
15. Ferraioli G, Tinelli C, Dal Bello B, et al. Accuracy of real-time shear wave elastography for assessing liver fibrosis in chronic hepatitis C: a pilot study. Hepatology 2012;56:2125-33.
16. Leung VY-f, Shen J, Wong VW-s, et al. Quantitative Elastography of Liver Fibrosis and Spleen Stiffness in Chronic Hepatitis B Carriers: Comparison of Shear-Wave Elastography and Transient Elastography with Liver Biopsy Correlation. Radiology 2013;269:910-918.
17. Poynard T, Munteanu M, Luckina E, et al. Liver fibrosis evaluation using real-time shear wave elastography: applicability and diagnostic performance using methods without a gold standard. J Hepatol 2013;58:928-35.

18. Deffieux T, Gennisson J-L, Bousquet L, et al. Investigating liver stiffness and viscosity for fibrosis, steatosis and activity staging using shear wave elastography. J Hepatol 2015;62:317-324.
19. Cassinotto C, Lapuyade B, Mouries A, et al. Noninvasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and Fibroscan. J Hepatol 2014;61:550-557.
20. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology 2009;49:1017-1044.
21. Calvaruso V, Burroughs AK, Standish R, et al. Computer-assisted image analysis of liver collagen: relationship to Ishak scoring and hepatic venous pressure gradient. Hepatology 2009;49:1236-44.
22. Huang Y, de Boer WB, Adams LA, et al. Image analysis of liver biopsy samples measures fibrosis and predicts clinical outcome. J Hepatol 2014;61:22-27.
23. Tsochatzis E, Bruno S, Isgro G, et al. Collagen proportionate area is superior to other histological methods for sub-classifying cirrhosis and determining prognosis. J Hepatol 2014;60:948-54.
24. Boursier J, de Ledinghen V, Poynard T, et al. An extension of STARD statements for reporting diagnostic accuracy studies on liver fibrosis tests: The Liver-FibroSTARD standards. J Hepatol 2015;62:807-15.
25. Aertgeerts B, Buntinx F, Kester A. The value of the CAGE in screening for alcohol abuse and alcohol dependence in general clinical populations: a diagnostic meta-analysis. J Clin Epidemiol 2004;57:30-9.
26. Roy M, Dum M, Sobell LC, et al. Comparison of the quick drinking screen and the alcohol timeline followback with outpatient alcohol abusers. Subst Use Misuse 2008;43:2116-23.
27. Boursier J, Zarski J-P, de Ledinghen V, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 2013;57:1182-1191.
28. Sporea I, Gradinaru-Tascau O, Bota S, et al. How many measurements are needed for liver stiffness assessment by 2D-Shear Wave Elastography (2D-SWE) and which value should be used: the mean or median? Med Ultrason 2013;15:268-72.
29. Procopet B, Berzigotti A, Abraldes JG, et al. Real-time shear-wave elastography: applicability, reliability and accuracy for clinically significant portal hypertension. J Hepatol 2015;62:1068-75.
30. Nascimbeni F, Lebray P, Fedchuk L, et al. Significant Variations in Elastometry Measurements Made Within Short-term in Patients With Chronic Liver Diseases. Clin Gastroenterol Hepatol 2015;13:763-771.
31. Macaluso FS, Maida M, Cammà C, et al. Steatosis Affects the Performance of Liver Stiffness Measurement for Fibrosis Assessment in Patients with Genotype 1 Chronic Hepatitis C. Journal of Hepatology 2014;61:523-9.
32. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926-30.
33. Afdhal NH, Bacon BR, Patel K, et al. Accuracy of Fibroscan, Compared With Histology, in Analysis of Liver Fibrosis in Patients With Hepatitis B or C: A United States Multicenter Study. Clin Gastroenterol Hepatol 2015;13:772-779.
34. Lemoine M, Katsahian S, Ziol M, et al. Liver stiffness measurement as a predictive tool of clinically significant portal hypertension in patients with compensated hepatitis C virus or alcohol-related cirrhosis. Aliment Pharmacol Ther 2008;28:1102-1110.

35. Friedrich-Rust M, Ong MF, Martens S, et al. Performance of Transient Elastography for the Staging of Liver Fibrosis: A Meta-Analysis. Gastroenterology 2008;134:960-974.

Figure 1
Click here to download Figure: Figure_1_studyflowdiagram.png

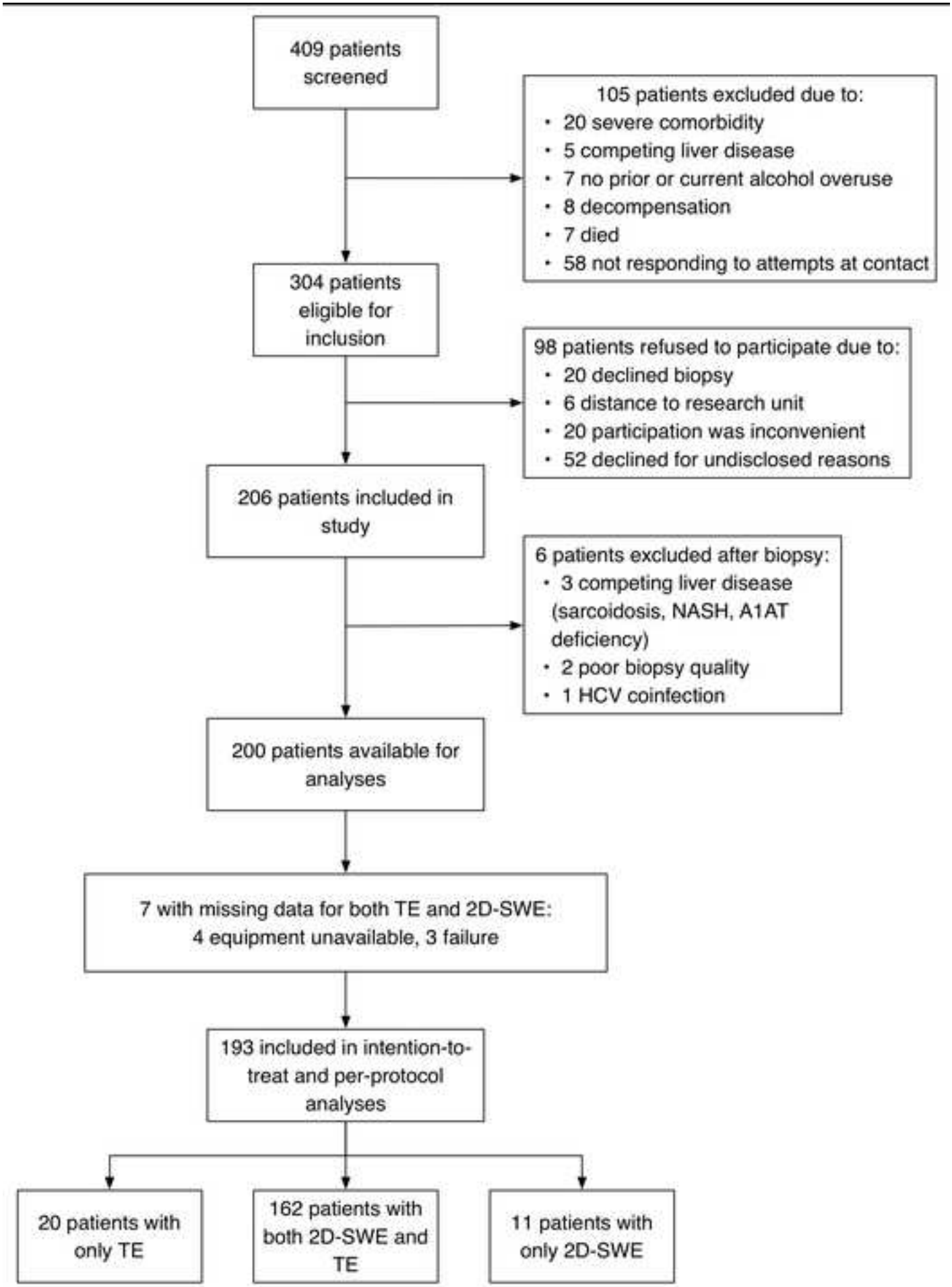


Figure 2
Click here to download Figure: Figure_2_CPA_TE_SWEoverISHAK.png

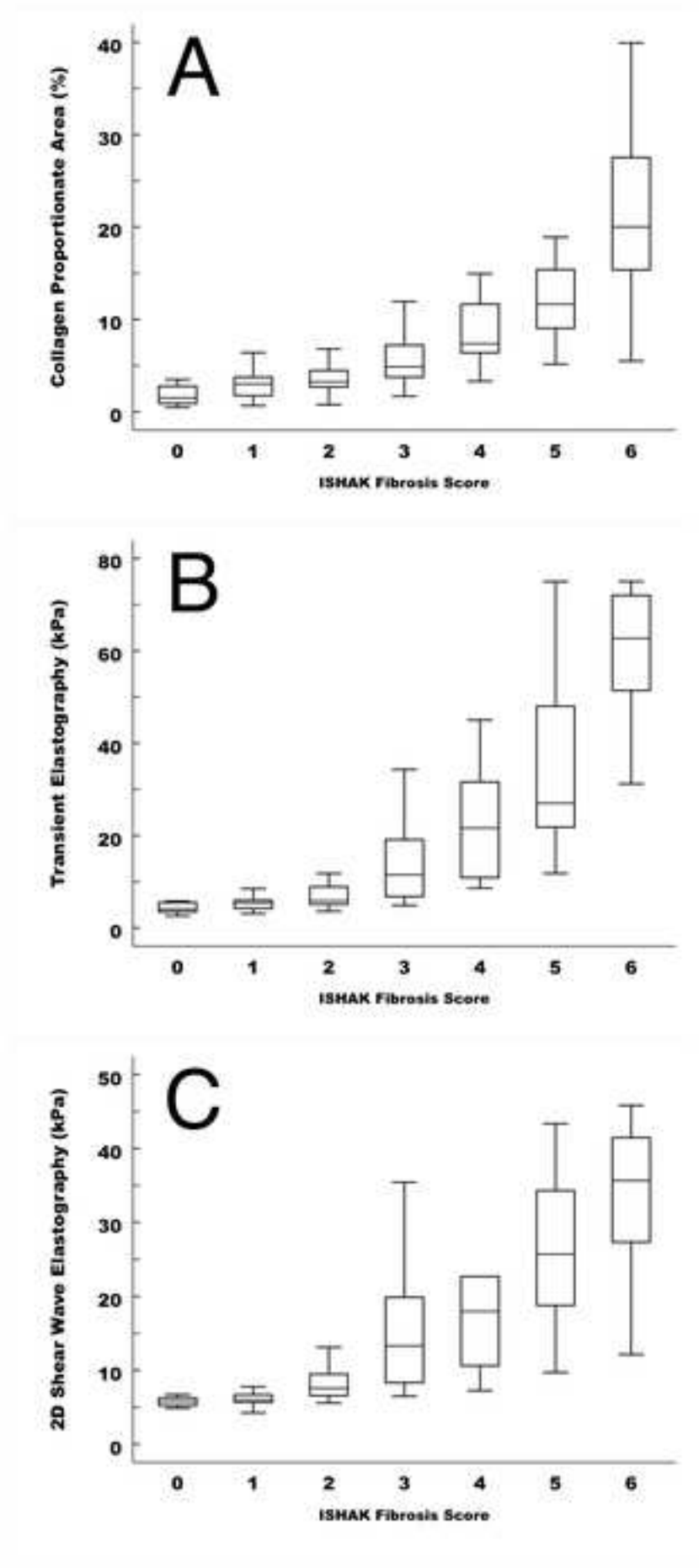


Figure 3
Click here to download Figure: Figure_3_ROC_I3_I5.png

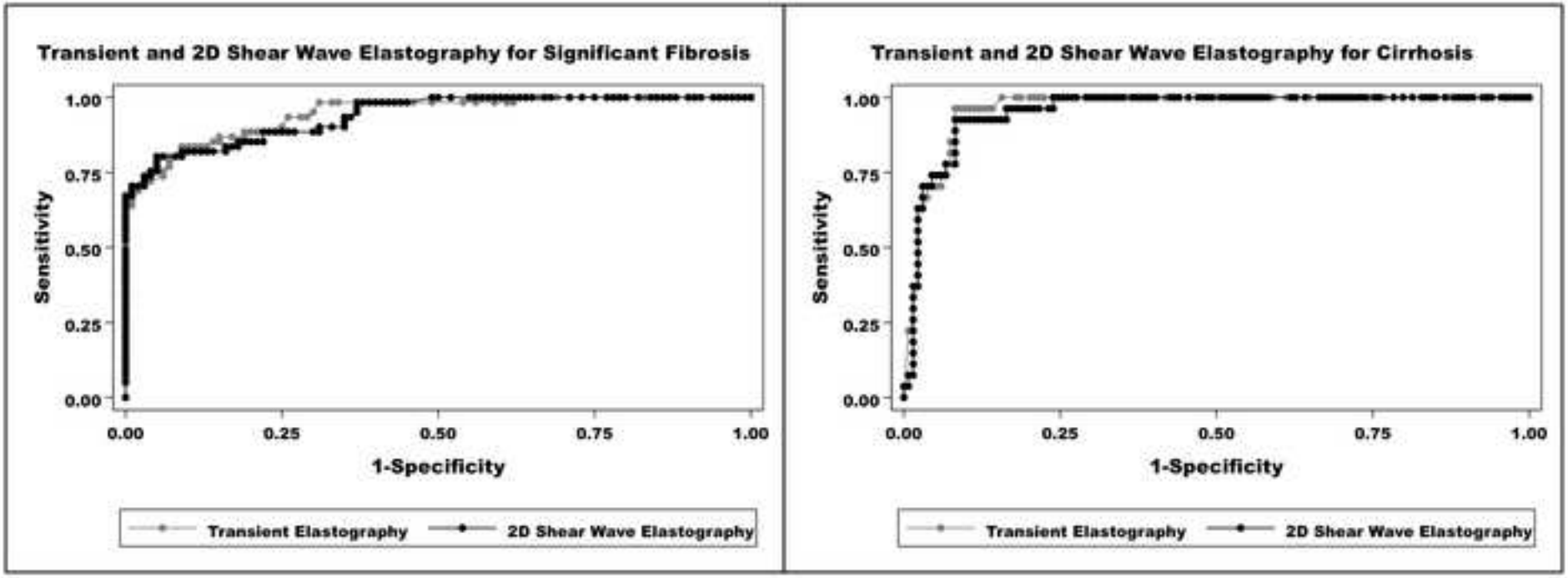
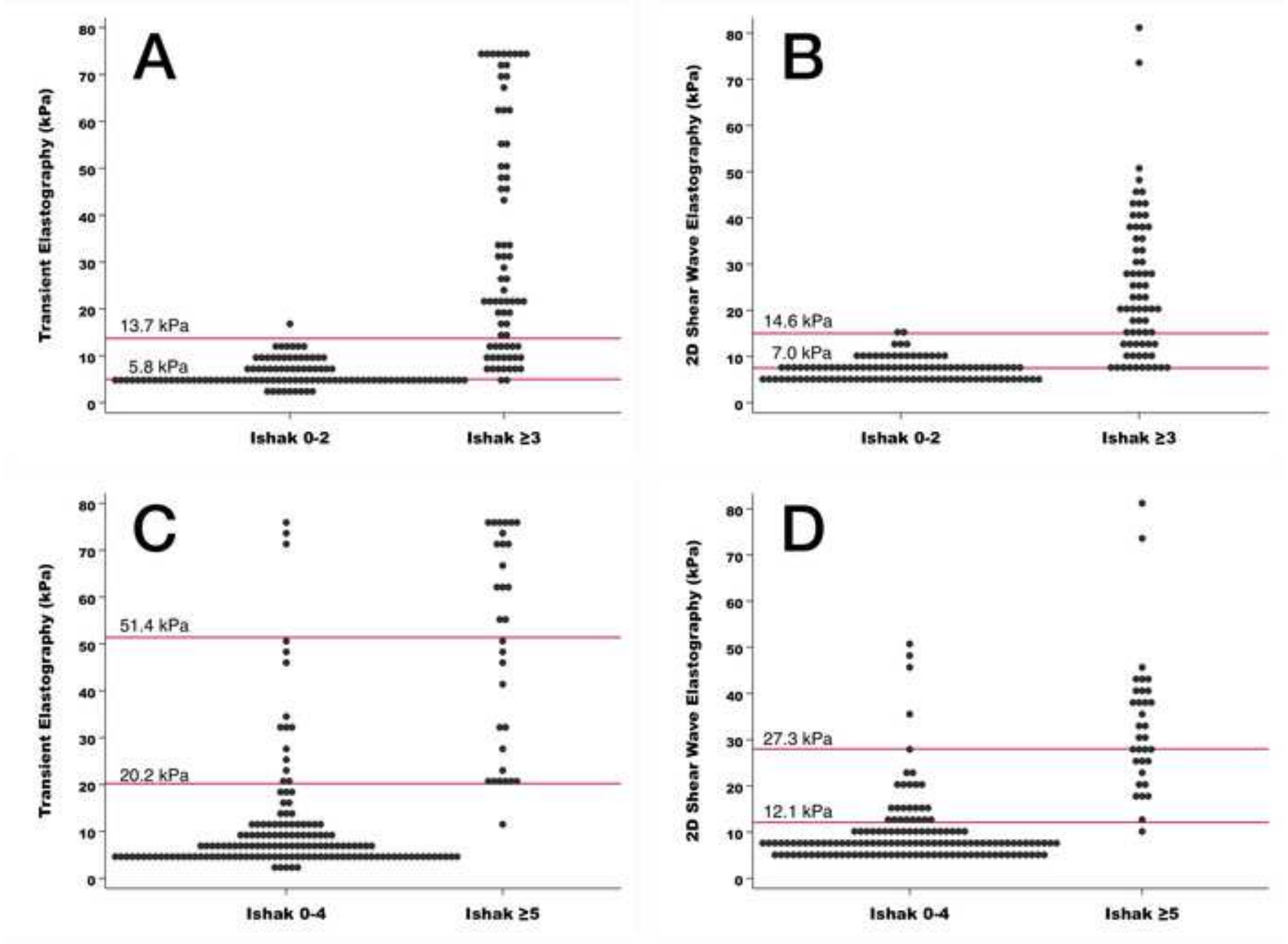


Figure 4

Click here to download Figure: Figure_4_Distributional_plots_LS_Ishak.png

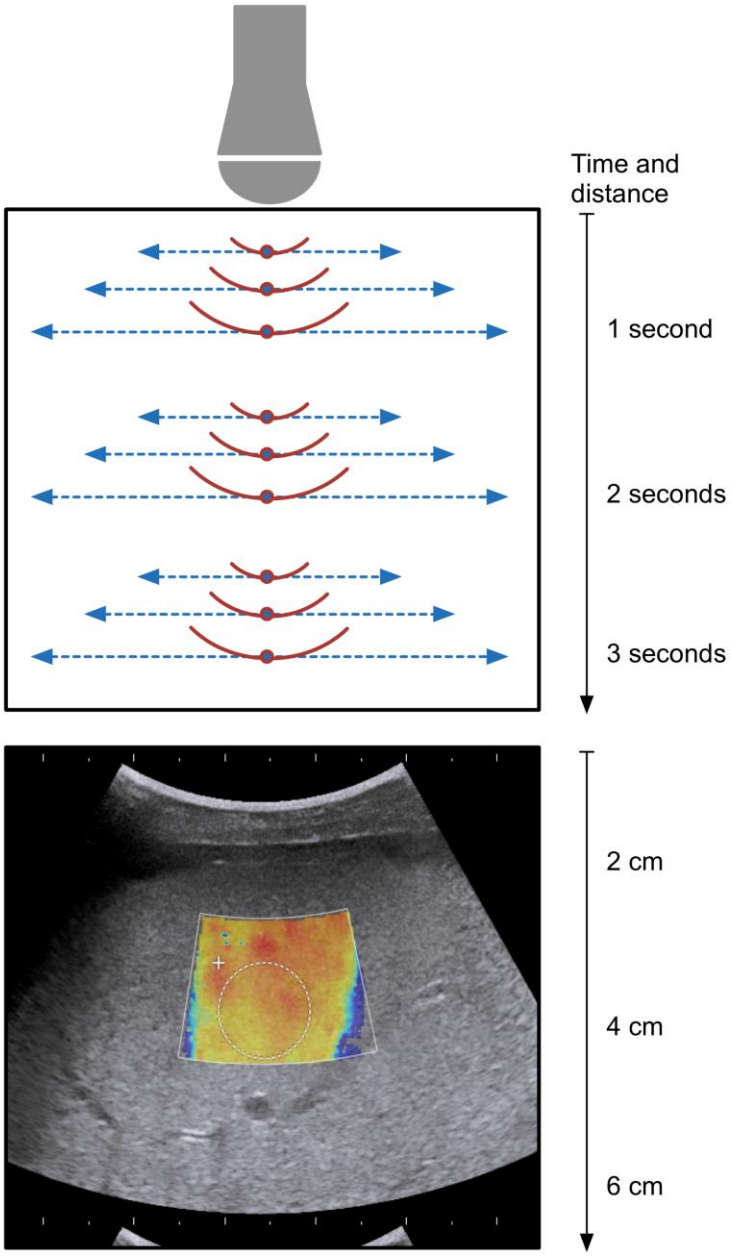


APPENDIX

1. Figure: Principle of 2D Shear Wave Elastography (2D-SWE).
2. Detailed description of the software to measure collagen proportionate area.
3. Figure: Correlation between transient elastography and 2D shear wave elastography.
4. Table: Diagnostic values for rule-in and rule-out cut-off values.
5. Liver-FibroSTARD checklist.

1. Principle of 2D Shear Wave Elastography (2D-SWE).

2D-SWE displays a colour-coded viscoelasticity map of the liver from which the elastography can be obtained using a operator-adjustable, circular ROI (the Q-box). Because the viscoelasticity map is generated by the propagation of multiple shear waves over time, the colour needs to be stable for 2-4 seconds before 2D-SWE is measured.

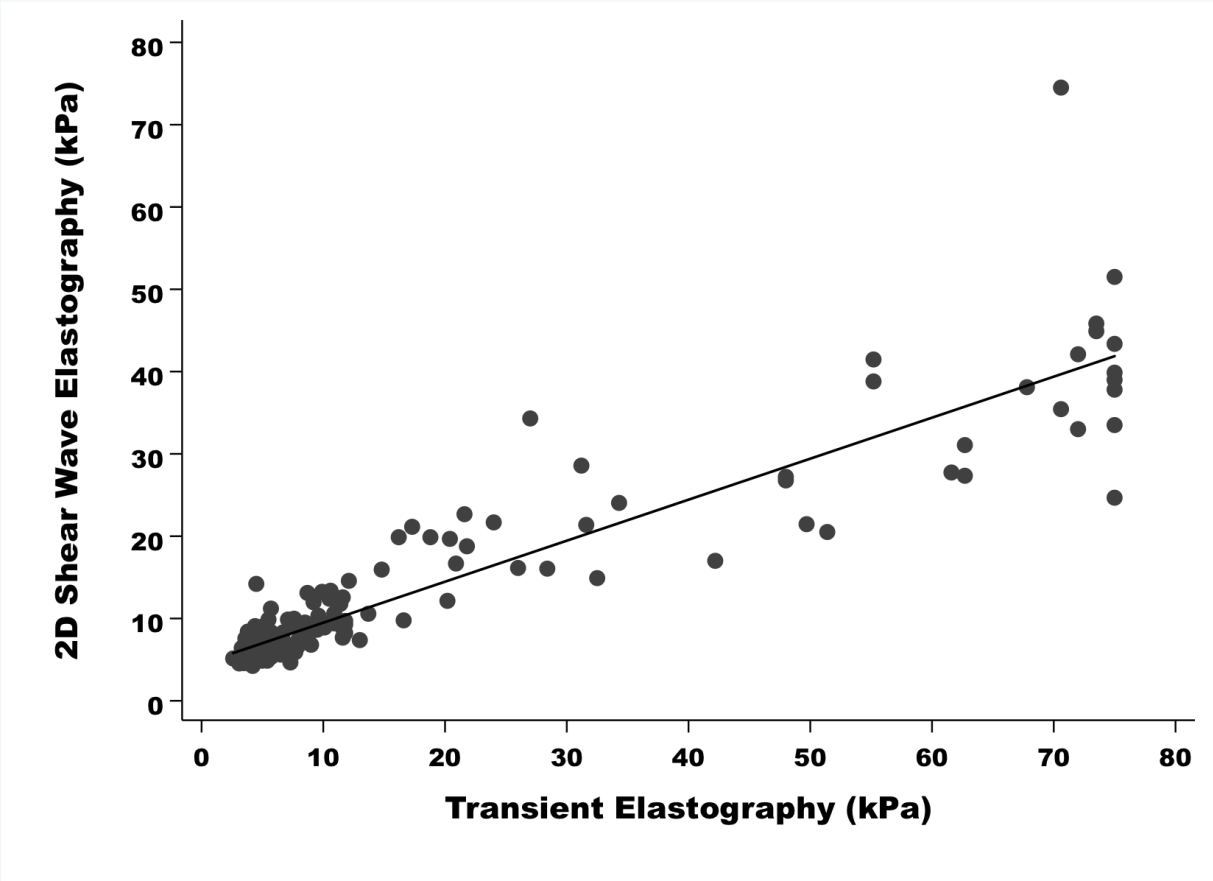


2. Description of the software to measure collagen proportionate area

Our liver pathologist (SD) and Visiopharm (Visiopharm, Hørsholm, Denmark) collaborated in the development of a software application ("APP"), using the image analysis tool VISIOmorph version 4.3.6 (Visiopharm, Hørsholm, Denmark). With the VISIOmorph tool, an algorithm was designed to separate the liver samples into five spaces: fat vacuoles, other vacuolar spaces that could be perceived as fat vacuoles, nuclei, collagen and remaining tissue. The first image processing step involves segmentation of the liver tissue from the background at a low digital magnification. After the first image processing, high-resolution analysis is performed. The high-resolution analysis uses a Bayesian classifier trained on preprocessing steps that highlight the green-blue contrast, red-green contrast and local linear areas. This detects hepatocyte area, fat vacuoles, sinusoids, blood vessels and collagen. Following classification, post processing steps are applied to remove non-collagen objects and areas of tear, holes etc. that could affect the calculated collagen ratio. In some instances, large vessels and other collagenous structures not representing liver fibrosis are manually excluded.

To validate the software, we randomly selected five images from the complete dataset. The pathologist assessed the selected images, to ensure that they represented the expected variations in regards to staining intensity, tissue structures and artifacts. Next, the VISIOmorph software measured collagen proportionate area measurements on this training set. The automated measurements were appraised by our pathologist, with subsequent software modifications and reiterations, e.g. to handle erythrocytes and other structures initially appearing as small collagen fragments. Once approved in performance and expected outcome, the algorithm was locked (APP ID: 10086) and applied to the remainder of the patient cohort.

3. Correlation between transient elastography and 2D shear wave elastography.



Correlation coefficient rho = 0.84, P<0.001

4. Table with diagnostic values for rule-in and rule-out cut-off values.

Decision table to rule-in or rule-out significant fibrosis (Ishak≥3) and cirrhosis by means of low and high cut-off values for transient elastography (TE) and 2D shear wave elastography (2D-SWE)							
Diagnosis	Decision	Elastography technique	Cut point (kPa)	Sens. (%)	Spec. (%)	PPV (%)	NPV (%)
ISHAK≥3	Rule out	**TE**	**5.8**	**99**	**60**	**64**	**98**
		SSI	7.0	99	62	65	98
	Rule in	**TE**	**13.7**	**69**	**99**	**98**	**81**
		SSI	14.6	68	100	100	81
ISHAK≥5	Rule out	**TE**	**20.2**	**97**	**90**	**69**	**99**
		SSI	12.1	97	82	55	99
	Rule in	**TE**	**51.4**	**55**	**98**	**86**	**91**
		SSI	27.3	64	97	84	92
Abbreviations: 2D-SWE, 2D shear wave elastography; kPa, kiloPascal; NPV, negative predictive value; PPV, positive predictive value; sens., sensitivity; spec., specificity; TE, transient elastography.							

5. Liver-Fibro STARD checklist

LIVER-FIBROSTARD CHECKLIST

The Liver-FibroSTARD checklist summarizes the important information that must be present in the manuscripts of diagnostic studies on non-invasive tools for liver fibrosis evaluation. Compared to STARD, the Liver-FibroSTARD checklist includes 2 additional items (#12 and #26) and 44 sub-items. The sub-items correspond to those proposals that clearly depicted, within the items, each of the particular features of diagnostic studies on liver fibrosis tests. Finally, Liver-FibroSTARD presents as a complementary module of the STARD checklist.

Some items or sub/items include several criteria; major criteria are indicated by an asterisk (*). Example: item #3: "The study population: The inclusion and exclusion criteria, setting, and locations* where data were collected". If a major item is missing, the corresponding criterion has to be rated absent. Some items/sub-items (#12.1 and #23.1, #13.10 and #22.2) are redundant since they can be found in different locations of the article.*

TITLE/ABSTRACT/KEYWORDS	**1. Identify the article as a study of diagnostic accuracy (recommend MeSH heading "sensitivity and specificity").**
1.1. Identify the article, especially in the title, as a study of the diagnostic performance of liver fibrosis/cirrhosis biomarker(s)/test(s).	Cover page
1.2. Recommended key words (choose the most appropriate): "liver fibrosis", "cirrhosis", "diagnosis", "biomarker", "diagnostic test", "noninvasive diagnosis".	Abstract page
INTRODUCTION	**2. State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.**
In study aims, specify:	
2.1. If the aim is to identify new marker(s)/develop new test(s), or to evaluate published marker(s)/test(s).	Page 5
2.2. Whether the study is performed in a single or multiple cause(s) of chronic liver disease.	Title and pages 5 and 6.
2.3. The reference used for fibrosis diagnosis in the study.	Page 5, 6 and 7

2.4. The diagnostic target used as the primary aim of the study and, if appropriate, other diagnostic targets used as secondary aims.	Page 5
METHODS	*Describe:*
3. The study population: The inclusion and exclusion criteria*, setting, and locations* where data were collected.	Page 6-7

4. Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?

Page 6

4.1. State if healthy subjects without chronic liver disease are included or not in the study.	Positive controls – i.e. at-risk patients with alcohol overuse, but no liver disease on biopsy – were included. No healthy controls were included.
4.2. State if patients were selected by one abnormal or several discordant fibrosis test(s).	No. Selected based on at-risk profile.
4.3. State if patients were selected according to the availability of reference or index test(s) result(s).	No
5. Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Page 6

6. Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?

Prospective

6.1. The chronology between patient inclusion, data collection (reference/index tests)*, and data analysis is well described.*	Page 7
6.2. Has the study population been previously used/published for the evaluation of the studied fibrosis test(s)?	No
7. The reference standard and its rationale.	Page 4-5

8. Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.

For the reference and index test(s), specify characteristics with sufficient detail to permit exact reoperation, when appropriate:	
8.1. Center: standardization of procedures across centers.	Page 7
8.2. Patient: fasting conditions, time, posture, etc. (give information about the influence of conditions on the intra-individual variability).*	Page 7
8.3. Delay: time interval between reference and index test(s).	None. All tests performed on the same day.
8.4. Material: technical specifications (name, generation, manufacturer, instrument), method of measurement, applicability (failure/reliability criteria). Specifically for liver biopsy, indicate material used per center, i.e. percutaneous/transjugular/other, needle diameter.*	Page 7-9
8.5. Biological samples: description of method of collection, transport, storage.*	Page 8-9
8.6. Specify how the index tests were calculated.	Page 8
8.7. Specify how the risk for false negative/positive results was taken into account.	Page 6-7: Fasting conditions; exclusion of patients with right heart failure, cholestasis and severe hepatitis
Specifically for liver biopsy:	
8.8. How sample bias was limited: minimal biopsy size (length), number of portal tracts required, number of fragments.*	Page 8
8.9. Methods for histological assessment: human/automated reading, local/central reading*, number and expertise of pathologists*, single/double reading*, consensus methods.*	Page 8-9
8.10. Scoring system used (Metavir, Ishak, Scheuer, etc.).	Page 8
9. Definition of and rationale for the units, cut-offs*, and/or categories of the results of the index tests and the reference standard.	Pages 4 and 9
10. The number*, training and expertise* of the persons executing and reading the index tests and the reference standard.	Page 7

11. Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	Page 7
12. State if the study is conducted on an intention-to-diagnose basis or if the analysis is per-protocol (i.e. with exclusion of failed/unreliable fibrosis test(s)/reference measurements).	Both per-protocol and intention-to-diagnose are used.
12.1. If intention-to-diagnose analysis, specify how failure and unreliable test(s)/reference are taken into account in the analysis. [a]	Pages 9 and 13
13. Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	
Specify:	
13.1. Detailed sample size calculation.	Not done
13.2. Statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	95% CI
13.3. Control of multiple comparisons that increases type I error: multiple comparisons of tests (e.g. Bonferroni correction, etc.), multiple diagnostic targets.	No multiple comparisons
13.4. Method for calculation of fibrosis test(s) diagnostic cut-offs.	Maximizing Youden Index. For the rule-in and rule-out cut offs, LR+ and LR- were maximized.
13.5. Method for validation of new test(s) or new calculated diagnostic cut-off(s) (e.g. external validation set, internal validation by bootstrapping, etc.).	Page 18. Not done
13.6. Method for control of center/operator effect.	Page 11
13.7. Method for control of spectrum effect if unrepresentative prevalence of fibrosis stages (e.g. Obuchowski index, DANA, etc.).	Page 6
13.8. Method for control of misclassification errors by the reference test.	Page 11-13; no method to correct misclassifications
13.9. Use of a reference without gold standard.	Discussed on page 17

13.10. Analysis of discordances between reference/index test(s). b	Pages 12-13 and 17
14. Methods for calculating test reproducibility.	Not done
RESULTS	
15. When study was performed, including beginning and end dates of recruitment.	Page 10
16. Clinical and demographic characteristics of the study population (e.g. age*, sex*, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centers).	Table 1
16.1. For liver biopsy: size (length), number of portal tracts, number of fragments.*	Table 1
16.2. For index test(s): confounding factors that potentially influence the test(s) results (flare-up, inflammation, other liver lesions, intrinsic characteristics, etc.).	Page 11, table 1
17. The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard*; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	Figure 1, page 10-11
17.1. If per-protocol analysis, report comparisons between patients excluded due to failed/unreliable test(s)/reference and patients with reliable fibrosis test(s)/reference.	ITD done
18. Time-interval* between the index tests and the reference standard, and any treatment administered between.	Same-day index and reference
19. Distribution of severity of disease (define criteria) in those with the target condition*; other diagnoses in participants without the target condition.	Pages 10 and 14-15, Table 1
19.1. Specify the prevalence of the diagnostic condition (spectrum effect).*	Pages 10 and 14, Table 1

20. A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	
20.1. Presentation of contingency tables, box/scatter plots.	Figures 2 and 4
21. Any adverse events from performing the index tests or the reference standard.	None for index. Biopsy AE's is well known
22. Estimates of diagnostic accuracy* and measures of statistical uncertainty (e.g. 95% confidence intervals).	95% CI for AUROCS and regression coefficients.
22.1. Specify sensitivity and specificity* with 95% confidence intervals; ROC analysis.*	Table 3, pages 12-15
22.2. Analyzing discordances between fibrosis tests(s)/reference. b	Page 12-13
23. How indeterminate results, missing data and outliers of the index tests were handled.	
23.1. How missing/failure/unreliable results of index test(s)/reference were handled (intention-to-diagnose/per-protocol analysis). a	ITD and per-protocol.
23.2. How outliers of the index tests were handled.	Outliers included in analyses
24. Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Page 14-15, table 3
25. Estimates of test reproducibility, if done.	Not done
26. Estimates of cost-benefit.	Not done
DISCUSSION	**27. Discuss the clinical applicability of the study findings.**
27.1. Discuss the representativeness of the study sample and recruiting centers (i.e. spectrum effect, etc.).	Pages 16 and 18

27.2. Discuss the interpretation of fibrosis test(s) results in clinical practice. Pages 16-19

27.3. Discuss the clinical relevance of the study results. Pages 16-19

[b] *Items 13.10 and 22.2 are redundant but retained since they can be located in different paragraphs within an article*