Jüngst Christoph (Orcid ID: 0000-0001-5048-390X)

Performance of two-dimensional shear wave elastography and transient elastography compared to liver biopsy for staging of liver fibrosis

Running title: Elastography for staging of liver fibrosis

Audrey Kovatsch[1*], Dr. Hanna Honcharova-Biletska[2*], Dr. Daniel Segna[1,3], Klaus Steigmiller[4], Dr. Sena Blümel[1], Dr. Rudolf A Deibel[1], Tobias Kühlewindt[1], Dr. Georg Leinenkugel[1], Dr. Sandra Müller[1], Dr. Eva Furrer[4], Dr. Khoschy Schawkat[5,6], Prof. Cäcilia S Reiner [6], Prof. Achim Weber[2], Prof. Beat Müllhaupt[1], Prof. Michael Scharl[1], Prof. Christoph Gubler[1,7], Dr. Christoph Jüngst[1,8].

**Both authors contributed equally*

[1]Department of Gastroenterology and Hepatology, University Hospital Zurich, Zurich, Switzerland

[2]Department of Pathology, University Hospital Zurich, Zurich, Switzerland

[3]Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland

[4]Institute of Epidemiology, Biostatistics and Prevention, University of Zurich, Zurich, Switzerland

[5]Department of Radiology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA

[6]Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Zurich, Switzerland

[7]Division of Gastroenterology &Hepatology, Triemli Hospital, Zurich, Switzerland

[8]Department of Medicine II, Saarland University Medical Center, Saarland University, Homburg, Germany

Corresponding author: PD Dr. Christoph Jüngst, Department of Gastroenterology and Hepatology, University Hospital Zurich, Rämistrasse 100, CH-8091Zürich, Switzerland, +41 44 255 8548, christoph.juengst@usz.ch

Acknowledgements

Financial support statement: The study was founded by internal grants

Word count main body of the manuscript:3428

Abstract

Background:

Staging of liver fibrosis traditionally relied on liver histology, however transient elastography (TE) and more recently two-dimensional shear wave elastography (2D-SWE) evolved to non-invasive alternatives. Hence, we evaluated the diagnostic accuracy of 2D-SWE assessed by the Canon Aplio i800 ultrasound system using liver biopsy as reference and compared the performance to TE.

Methods:

In total, 108 adult patients with chronic liver disease undergoing liver biopsy, 2D-SWE and TE were enrolled prospectively at the University Hospital Zurich. Diagnostic accuracies were evaluated using the area under the receiver operating characteristic (AUROC) analysis, and optimal cutoff values by Youden's index

Results:

Diagnostic accuracy of 2D-SWE was good for significant (≥F2; AUROC 85.2%, 95% confidence interval (95%CI):76.2-91.2%) as well as severe fibrosis (≥F3; AUROC 86.8%, 95%CI:78.1-92.4%) and excellent for cirrhosis (AUROC 95.6%, 95%CI:89.9-98.1%), compared to histology. TE performed equally well (significant fibrosis: 87.5%, 95%CI:77.7-93.3%; severe fibrosis: 89.7%, 95%CI:82.0-94.3%; cirrhosis: 96%, 95%CI:90.4-98.4%), and accuracy was not statistically different to 2D-SWE. 2D-SWE optimal cutoff values were 6.5, 9.8 and 13.1 kPa for significant fibrosis, severe fibrosis, and cirrhosis, respectively.

Conclusions:

Performance of 2D-SWE was good to excellent and well comparable with TE, supporting the application of this 2D-SWE system in the diagnostic workup of chronic liver disease.

Key words:

2D-SWE, chronic liver disease, non-invasive, liver stiffness measurement, shear wave elastography, transient elastography

Introduction

Chronic liver diseases account for approximately two million deaths per year worldwide, mostly due to complications of cirrhosis, viral hepatitis, and hepatocellular carcinoma [1]. The burden of liver disease is increasing with most common underlying etiologies being excessive alcohol consumption, non-alcoholic fatty liver disease (NAFLD) and viral hepatitis B (HBV) or C (HCV) infection [2,3]. Chronic liver diseases cause an inflammatory process resulting in the development of liver fibrosis which may progress to liver cirrhosis which is associated with complications of portal hypertension, hepatocellular carcinoma, increased morbidity and mortality [4,5]. The process of fibrogenesis is dynamic and therapeutic decisions, as well as prognosis depend on the actual fibrosis stage, which emphasizes the need for repetitive staging of liver fibrosis [6-9]. For many years, liver histology has been the gold standard for staging of liver fibrosis. However it is costly, invasive and its universal application is not feasible considering the large amount of patients with chronic liver disease requiring repeated assessment of liver fibrosis [7,10]. To overcome these limitations, non-invasive methods including serum-based tests and tools for liver stiffness measurement (LSM) have been developed for the staging of liver fibrosis [11-13]. Ultrasound-based methods for assessment of liver stiffness are the most extensively studied non-invasive procedures. Among those, the most frequently evaluated technique is transient elastography (TE) which has shown good to excellent diagnostic accuracy for staging liver fibrosis compared to liver histology in numerous studies and for different underlying etiologies of liver disease [14-18]. However, there are some disadvantages related to the application of TE given that patients with ascites cannot be examined and during a measurement no specific region of interest (ROI) can be selected [7]. Furthermore, due to the need of a specific device for TE-measurements the

resulting acquisition cost may limit its utilization to high volume and tertiary care centers [7,19].

In recent years, a new ultrasound-based technique, the two-dimensional shear wave elastography (2D-SWE) has been introduced, which promises to resolve some of the limitations of TE. 2D-SWE is implemented in a standard ultrasound system and showed good diagnostic accuracy for staging liver fibrosis in different studies [7,11,18]. Several technical approaches for 2D-SWE on various ultrasound systems are currently available and cutoff values for fibrosis stages cannot be interchanged between different systems [11,13]. Therefore, all of these techniques have to be validated for accuracy in different patient cohorts with chronic liver diseases, before utilization in clinical practice can be recommended. The Canon Aplio i800 system is a recently developed 2D-SWE system using a propagation map, which allows precise placement of a ROI for detection of liver stiffness, however data on the diagnostic performance is limited.

Therefore, we aimed to investigate the diagnostic accuracy of 2D-SWE using the Canon Aplio i800 system in relation to the current gold standard of histopathological staging of liver fibrosis, determine cutoff values for different fibrosis stages and compare its diagnostic accuracy to TE.

Patients and methods

Patients

This prospective cohort study was performed at the Department of Gastroenterology and Hepatology, University Hospital Zurich, Zurich, Switzerland. From January 2018 through July 2020, consecutive adult patients with chronic liver disease of different etiologies scheduled for liver biopsy were included. Exclusion criteria were acute liver disease, such as acute alcoholic or viral hepatitis, and patients with ascites. Upon written informed consent, we performed LSM by means of TE and 2D-SWE prior to liver biopsy. LSM was performed by specifically trained and experienced physicians. In most cases the stiffness measurements obtained from 2D-SWE and TE were performed on the same day prior liver biopsy. In few exceptions, only 2D-SWE was performed directly before liver biopsy and respective TE measurements were used from a previous visit no longer than six months ago. In addition to the patients undergoing a liver biopsy, some patients with liver cirrhosis as diagnosed clinically and by imaging studies were included in the group of cirrhotic patients (F4).

Furthermore, we recorded clinical parameters and demographic data (sex, age, body mass index (BMI), history of diabetes and smoking). The following biochemical parameters were included whenever available: serum gamma-glutamyltranspeptidase (GGT), aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP), albumin, total bilirubin levels, creatinine, platelet count and the international normalized ratio (INR).

The study was approved by the local ethics committee (Kantonale Ethikkommission Zürich, KEK-ZH-Nr. 2013-0024) and the study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

Liver biopsy and histological examination

For all liver biopsies, we used a Biopince™ Needle 16G (Argon Medical, Texas, TX) [20]. Patients were placed in supine position, close and parallel to the edge of the bed, with the patient's right arm behind the head. Experienced examiners used an ultrasound-assisted approach for the liver biopsy. A diagnostic ultrasound was performed immediately before the biopsy to identify the optimal biopsy site, distant to any large hepatic vessels or biliary ducts [21].

The liver biopsy specimen was fixed in formalin and embedded in paraffin. All biopsy specimens were stained with hematoxylin-eosin and Sirius red. The histological slides were then analyzed by specialized hepatopathologists blinded to the non-invasive LSM. The liver biopsies were graded and staged histologically using METAVIR [22,23], NAFLD Activity Score (NAS) [24] or Batts-Ludwig score [23,25] as appropriate. After assessment of liver fibrosis stage and inflammatory activity, the results were standardized according to a harmonized scale (staging: F0-F1-F2-F3-F4, grading: A0, A1, A2, A3) [26], to have the opportunity to compare the aetiologically different cases. This histopathological staging of liver fibrosis was used as the reference method.

Two-dimensional shear wave elastography

For 2D-SWE measurements, we used the Aplioi800 ultrasound system (Canon Medical System, Japan) with a convex broadband probe. All patients were placed in a supine position with the right arm behind the head and the measurements taken at the right liver lobe through the intercostal space [11]. All patients had fasted for at least six hours. 2D-SWE is a real-time ultrasound method whose technical principle is based on acoustic radiation force impulse imaging (ARFI). The shear wave propagation is measured in a user-defined visualized and color-coded shear wave box, which superimposes an ultrasound B-mode picture (Figure 1). Within the box, a circular ROI

of 10 mm diameter was placed in a vascular free area with parallel lines in the propagation map one to five centimeter below the liver capsule [27,28]. The acquisitions were performed during breath hold. Following the manufacturer´s technical guidelines and the recommendations by the European Federation for Ultrasound in Medicine and Biology (EFSUMB) [11] and the European Association for the Study of the Liver (EASL) [7], all measurements were obtained by trained examiners with extensive experience in ultrasound and blinded to the histopathological results. The median value (Med) of five to ten 2D-SWE acquisitions in kPa and the interquartile/median ratio (IQR/Med) were calculated for statistical analysis [7,11].

Transient elastography

TE was performed with the Fibroscan© (Echosens, Paris, France) by specifically trained physicians according to the EFSUMB and EASL recommendations [7,11]. The patients were placed in a supine position with the right arm behind the head after a fasting period of at least six hours. The M probe (3.5 MHz, standard) was used for the measurements in a region in the right hepatic lobe for patients with BMI < 30 kg/m^2 or the XL probe (2.5 MHz) for BMI ≥ 30 kg/m^2 [6,29,30]. Reliable results were defined as ≥10 valid measurements with a success rate ≥ 60% and an IQR/Med ≤ 30% as suggested by current guidelines [7,11]. For each patient, the median value and the IQR/Med ratio were used for statistical analysis.

Statistical analysis

For descriptive analyses, mean and standard deviation (SD) were reported for approximately normal continuous variables, median and interquartile range (IQR) for skewed continuous variables and frequency and percentage for categorical variables.

Distributional assumptions were visually investigated by inspection of, inter alia, boxplots.

The optimal cutoffs for 2D-SWE were estimated by maximizing Youden's index [31] weigthing the cost for false positive and false negative equally important.

We compared the diagnostic accuracy of 2D-SWE and TE to liver histology in terms of area under the receiver operating characteristic (AUROC) via DeLong et al. [32], using the R package biostatUZH (version 1.8.0). We used the Obuchowski-measure [33] with proportional penalty function and weighting scheme based on the prevalence of the fibrosis stages in this study for the overall diagnostic accuracy comparison of 2D-SWE and TE to take the ordinal scale and the spectrum bias into account using the R package nonbinROC (version 1.0.1). Moreover, 95% logit-Wald CI for AUROC and the Obuchowski-measure and 95% Wald CI for the paired difference of AUROC and the Obuchowski-measure were calculated. The AUROC was considered as good for values between 80% and 90% and as excellent for values above 90% [18].

No adjustment for multiple testing was done, also no subgroup or sensitivity analyses were performed. Indeterminate reference test (liver biopsy) results did not occur, observations with indeterminate index test (2D-SWE and TE) results were dropped. Due to the low number of missing values, missing data was handled with complete case analysis for the corresponding methods assuming data missing completely at random (MCAR).

All statistical analyses were performed in R version 4.0.3 by an independent team of statisticians at the Institute of Epidemiology, Biostatistics and Prevention at University of Zurich. Results were reported according to the STARD guideline [34] and reporting confirms to the broad EQUATOR guidelines [35].

Results

Patient cohort

During the study period, 123 patients were initially included, thereof nine patients (7,3%) were excluded as these patients suffered from acute liver disease, for instance acute drug-induced liver injury. Of the remaining 114 patients, six had to be excluded because 2D-SWE acquisitions did not yield reliable results, corresponding to a failure rate of 2D-SWE of 5,3%. In these six patients, the acquisition of a proper propagation map was not achieved in a sufficient number of measurements due to distortion of propagation lines indicating artifacts or insufficient color filling of the box. Finally, a total of 108 patients were analyzed. The main clinical and demographic characteristics of the study cohort are reported in Table 1. The most frequent etiology was non-alcoholic fatty liver disease/steatohepatitis (NAFLD/NASH) (47.2%). The etiologies of the 15 patients (13.9%) in the miscellaneous group comprised idiopathic hepatopathy (six patients), autoimmune hepatitis (AIH, four patients), hereditary hemochromatosis (three patients) and chronic drug-induced liver injury (two patients). Four patients (3.7%) were included without any liver disease in the F0 group, as they were evaluated as donor for living donor liver transplantation and histopathological examination showed regular liver parenchyma without fibrosis. One patient with a small sized hemangioma and focal nodular hyperplasia not receiving a liver biopsy without any clinical, laboratory or imaging signs of liver injury was regarded as having a healthy liver and therefore was included in the F0 group. In our study population including 108 patients, the vast majority of 100 patients (92.6%) underwent a transcutaneous liver biopsy, whereas seven patients were included in the F4 group based on combined clinical and imaging signs of liver cirrhosis and concomitant portal hypertension. The number of patients in the different fibrosis stages was as follows: F0 17 patients

(15.7%), F1 19 patients (17.6%), F2 31 patients (28.7%), F3 26 patients (24.1%), F4 15 patients (13.9%).

In all but one patient 2D-SWE was performed on the same day of the liver biopsy, the latter with an interval of three months. Three patients (2.8%) had missing or incomplete TE measurements. In 25 patients (23.8%), TE was performed using the XL probe.

Diagnostic accuracy of 2D-SWE and TE for staging liver fibrosis as compared to liver histology

In 56 patients (52%), 10 measurements of 2D-SWE were performed, in the remaining patients at least five acquisitions. In 105 patients (97%), the IQR/median ratio was ≤ 0.3, while the IQR/median ratio was slightly above 0.3 in three patients, however these measurements were deemed reliable due to a proper propagation map. The median values of liver stiffness according to fibrosis stages for 2D-SWE and TE are presented in Supplementary Table 1. Individual patients' data and a boxplot of 2D-SWE are shown for the different fibrosis stages in Figure 2.

The optimal 2D-SWE cutoffs by maximizing Youden's index are shown in Table 2. Detailed statistical measures of accuracy and prevalence are given in Supplementary Table 2. Cross tabulations of 2D-SWE and TE by the results of the reference test liver biopsy for each fibrosis stage are shown in Supplementary Table 3 and Supplementary Table 5.

The diagnostic accuracy of 2D-SWE for staging liver fibrosis was good for discriminating "F0-F1 vs F2-F4" and "F0-F2 vs F3-F4", and was excellent for discriminating "F0-F3 vs F4", as compared to liver histology. The diagnostic accuracy for TE vs. liver histology showed comparable results (Figure 3A-C; Table 3). Due to a large overlap of liver stiffness values for the groups F0 and F1 detected by 2D-SWE as well as TE, these two groups cannot be distinguished by either method. Accuracy

measures for “F0 vs F1-F4” are given in Supplementary Table 2 and cross tabulations are shown in Supplementary Table 3.

Comparison of the diagnostic accuracy of 2D-SWE and TE

The diagnostic accuracy of 2D-SWE and TE for staging liver fibrosis was compared in all 105 patients with available data for both methods. There was no statistical evidence for a difference between these two non-invasive techniques in discriminating “F0-F1 vs F2-F4”, “F0-F2 vs F3-F4”, and “F0-F3 vs F4”, as compared to liver histology as the reference respectively, as presented in detail in Table 3. In addition, there was also no evidence for an overall difference between the diagnostic accuracy of 2D-SWE and TE using the Obuchowski-measure (Supplementary Table 4).

Discussion

The presented study is the first evaluating the Canon Aplio i800 2D-SWE system with liver biopsy as the reference standard including a direct comparison with TE. In our prospective cohort of 108 patients with chronic liver diseases, we observed a good to excellent diagnostic accuracy of 2D-SWE compared to liver histology, which was similar to TE. In addition, we derived cutoff values across all fibrosis stages.

The diagnostic accuracy of the 2D-SWE was good for diagnosing significant fibrosis (F2-F4; AUROC 85.2%), as well as severe fibrosis (F3-F4, AUROC 86.8%), and excellent for detecting cirrhosis (F4, AUROC 95.6%), as compared to liver histology.

The derived 2D-SWE cutoff from our cohort for classifying significant fibrosis was 6.5 kPa, for severe fibrosis 9.8 kPa and for cirrhosis 13.1 kPa, respectively. These results are similar to a large retrospective individual patient data based meta-analysis by Herrmann et al. using liver biopsy as a reference, which estimated AUROC of 86%, 91%, and 93% for the diagnosis of significant fibrosis, severe fibrosis, and liver cirrhosis, respectively, across all liver disease etiologies [18]. However, this study was performed on a different 2D-SWE device based on supersonic shear imaging (Aixplorer), the most frequently studied 2D-SWE technique. There are also some recently reported studies applying the Canon Aplio i800 2D-SWE system. In the single-center study by Ferraioli et al. including 367 patients this 2D-SWE system was utilized, however the diagnostic accuracy and internally derived cutoffs for staging of fibrosis were derived using TE as the reference method [28]. In contrast to our study, most of the patients had a lower stage of liver fibrosis and there was a higher proportion of chronic HCV infection, although NALFD/NASH was the most prevalent etiology of chronic liver disease in both studies. Despite differences in study designs and study populations, the 2D-SWE cutoff values were similar in both studies for significant fibrosis (7 kPa vs.

6.5 kPa) and severe fibrosis (9 kPa vs. 9.8 kPa), respectively. Since there was no specific cutoff value evaluated for the stage of liver cirrhosis, no comparison with our study results is possible. The multicenter study by Ronot et al. was also performed with the 2D-SWE Canon Aplio system, however a previous version (Aplio 500) was used and therefore cutoff values are not directly comparable with the Aplio i800 system. Of the 537 patients included in the study, most (89.6%) had chronic viral hepatitis as underlying etiology of chronic liver disease, and TE was used as reference method [36]. The authors found a correct classification of liver fibrosis in the majority of patients. In a study from Korea including 114 patients undergoing liver biopsy, 2D-SWE derived from the Aplio i800 system showed a good diagnostic accuracy to detect higher stages of fibrosis [37]. Compared to our study, populations differed in ethnicity, many patients had autoimmune hepatitis (32,4%) and the proportion of patient with higher stages of fibrosis (≥F3) was relatively low (21,1%). However, similar cutoff values for severe fibrosis (9.4kPa vs. 9.8 kPa) and cirrhosis (12.2 kPa vs. 13.1 kPa) were found. A head-to-head comparison of 2D-SWE with the diagnostic accuracy of TE was not provided. The same group investigated this 2D-SWE system in 102 patients with histopathologically confirmed NAFLD/NASH and reported cutoff values for significant fibrosis (≥F2) of 7.6 kPa and for severe fibrosis (≥F3) of 9 kPa [37]. A limitation of the study stated by the authors was the very low number of eight patients with advanced fibrosis/cirrhosis (≥F3), which hampered the derivation of cutoff values for cirrhosis. In our study, the relatively even distribution of fibrosis stages within the study population allowed us to evaluate also cutoff levels for advanced fibrosis and cirrhosis.

The Aplio i800 2D-SWE system includes a propagation map to visualize shear wave generation and propagation in real time to assist in selecting an optimal ROI placement for assessment of liver stiffness. Due to limited data, there is no general agreement on

objective quality criteria concerning valid 2D-SWE measurements and following manufactures technical guidelines is recommended. For TE, 10 valid readings and an IQR/median ≤ 30% need to be achieved for a reliable assessment. The number of required valid readings differs between studies evaluating 2D-SWE, though a minimum of three measurements is recommended [36]. In our study, in half of the patients ten measurements were performed and in the remaining patients at least five acquisitions. In total, six patients (5.3%) had to be excluded from our study due to unreliable 2D-SWE acquisitions. This failure rate corresponds to reported rates from other studies [11,36,37], and demonstrates the high applicability of 2D-SWE in these patients.

TE is by far the most widely studied technique for evaluation of liver stiffness and has been incorporated in diagnostic algorithms of different major liver diseases [7]. 2D-SWE overcomes some of the limitations of TE with the option of selecting a precise ROI in a B-mode ultrasound window and the possible application in patients with ascites. 2D-SWE exhibits the major advantage of being integrated in a standard ultrasound system and therefore, there is no need to purchase an additional device such as TE. In several individual studies, 2D-SWE has shown an equal or even higher diagnostic accuracy for staging liver fibrosis compared to TE. Hermann et al. described in their meta-analysis a significantly larger AUROC for the supersonic shear imaging system (Aixplorer) than that of TE for diagnosing significant fibrosis and cirrhosis, respectively [7,8,11,18,29,38]. In our study, the diagnostic accuracy of staging liver fibrosis was not different between 2D-SWE and TE (Table 3), suggesting that the utilized 2D-SWE system can be applied as an alternative tool to TE for non-invasive staging of liver fibrosis. The cutoff values derived with 2D-SWE and TE were similar (Table 2), which may facilitate the implementation of this 2D-SWE technique in clinical practice.

Next to the prospective direct comparison of 2D-SWE to TE, another strength of our study is the evaluation of diagnostic accuracy with histopathology as reference standard. Histopathological assessment of liver fibrosis has also some limitations such as sampling error due to the heterogenic distribution of fibrosis, costs and risk of complications [39]. In addition, interobserver variability in biopsy interpretation has to be taken into account, which may be less present when specialized hepatopathologists assess the liver biopsy [40]. Despite these limitations, it is still considered the current gold-standard enabling an elaborate determination of localization and amount of fibrosis [7]. For some major chronic liver diseases specific scores are recommended for the histopathological staging of liver fibrosis. In our study, the METAVIR-Score was used for patients with chronic viral hepatitis and the NAS-Score for patients with NAFLD as appropriate, liver biopsies of all other patients were scored with the Batts-Ludwig score [22-25]. Finally, the results from individual scoring were standardized according to a harmonized in-house developed scale, which allowed a comparison across different etiologies.

There are some limitations in this study. Due to the rather strict indication for liver biopsy, the sample size was not large enough to stratify for different etiologies of chronic liver disease, which may yield disease specific cutoffs for fibrosis stages and differences in diagnostic accuracies. The Fibroscan© XL probe was used in patients with a BMI $\geq$ 30 kg/m^2 as formerly suggested [40], however more recent guidelines recommend the use in patients with a skin-to-liver capsule distance of > 25mm [11]. Also, the proposed 2D-SWE cutoffs for different fibrosis stages need to be externally validated in other cohorts.

In conclusion, our study showed a good to excellent performance of 2D-SWE for staging significant, severe fibrosis and especially cirrhosis using the recently developed Canon Aplio i800 system with liver histology as reference standard. Importantly, these results were comparable with TE, the to date most frequently used non-invasive technique for LSM, however, unlike the latter 2D-SWE does not require additional equipment. Therefore, this system can be applied for evaluation of patients with chronic liver disease in clinical practice as an alternative to TE. Differences in diagnostic accuracy and cutoffs of 2D-SWE according to different etiologies of chronic liver disease still have to be elucidated.

References

1. Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. J Hepatol 2019;70(1):151-171.
2. Pimpin L, Cortez-Pinto H, Negro F, et al. Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. J Hepatol 2018;69(3):718-735.
3. Sepanlou SG, Safiri S, Bisignano C, et al. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Gastroenterology & Hepatology 2020;5(3):245-266.
4. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44(1):217-31.
5. Paternostro R, Reiberger T, Bucsics T. Elastography-based screening for esophageal varices in patients with advanced chronic liver disease. World journal of gastroenterology 2019;25(3):308-329.
6. Barr RG, Ferraioli G, Palmeri ML, et al. Elastography Assessment of Liver Fibrosis: Society of Radiologists in Ultrasound Consensus Conference Statement. Ultrasound Q 2016;32(2):94-107.
7. Berzigotti A, Tsochatzis E, Boursier J, et al. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis – 2021 update. Journal of Hepatology 2021;75(3):659-689.
8. Ferraioli G, Tinelli C, Dal Bello B, Zicchetti M, Filice G, Filice C. Accuracy of real-time shear wave elastography for assessing liver fibrosis in chronic hepatitis C: a pilot study. Hepatology 2012;56(6):2125-33.

9. Dulai PS, Singh S, Patel J, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology 2017;65(5):1557-1565.

10. Castera L, Friedrich-Rust M, Loomba R. Noninvasive Assessment of Liver Disease in Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2019;156(5):1264-1281.e4.

11. Dietrich CF, Bamber J, Berzigotti A, et al. EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). Ultraschall Med 2017;38(4):e16-e47.

12. Lim JK, Flamm SL, Singh S, Falck-Ytter YT. American Gastroenterological Association Institute Guideline on the Role of Elastography in the Evaluation of Liver Fibrosis. Gastroenterology 2017;152(6):1536-1543.

13. Ferraioli G, Wong VW, Castera L, et al. Liver Ultrasound Elastography: An Update to the World Federation for Ultrasound in Medicine and Biology Guidelines and Recommendations. Ultrasound Med Biol 2018;44(12):2419-2440.

14. Pavlov CS, Casazza G, Nikolova D, Tsochatzis E, Gluud C. Systematic review with meta-analysis: diagnostic accuracy of transient elastography for staging of fibrosis in people with alcoholic liver disease. Aliment Pharmacol Ther 2016;43(5):575-85.

15. Mikolasevic I, Orlic L, Franjic N, Hauser G, Stimac D, Milic S. Transient elastography (FibroScan(®)) with controlled attenuation parameter in the assessment of liver steatosis and fibrosis in patients with nonalcoholic fatty liver disease - Where do we stand? World J Gastroenterol 2016;22(32):7236-51.

16. Anstee QM, Castera L, Loomba R. Impact of non-invasive biomarkers on hepatology practice: Past, present and future. Journal of Hepatology 2022;76(6):1362-1378.

17. de Lédinghen V, Vergniol J, Foucher J, Merrouche W, le Bail B. Non-invasive diagnosis of liver steatosis using controlled attenuation parameter (CAP) and transient elastography. Liver Int 2012;32(6):911-8.

18. Herrmann E, de Lédinghen V, Cassinotto C, et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. Hepatology 2018;67(1):260-272.

19. Foucher J, Castéra L, Bernard PH, et al. Prevalence and factors associated with failure of liver stiffness measurement using FibroScan in a prospective study of 2114 examinations. Eur J Gastroenterol Hepatol 2006;18(4):411-2.

20. Hall TC, Deakin C, Atwal GS, Singh RK. Adequacy of percutaneous non-targeted liver biopsy under real-time ultrasound guidance when comparing the Biopince™ and Achieve™ biopsy needle. Br J Radiol 2017;90(1080):20170397-20170397.

21. Al Knawy B, Shiffman M. Percutaneous liver biopsy in clinical practice. Liver Int 2007;27(9):1166-73.

22. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. Hepatology 1996;24(2):289-293.

23. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. Journal of Hepatology 2007;47(4):598-607.

24. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology 2011;53(3):810-20.

25. Batts KP, Ludwig J. An Update on Terminology and Reporting. The American Journal of Surgical Pathology 1995;19(12):1409-1417.

26. Schawkat K, Ciritsis A, von Ulmenstein S, et al. Diagnostic accuracy of texture analysis and machine learning for quantification of liver fibrosis in MRI: correlation with MR elastography and histopathology. European Radiology 2020;30(8):4675-4685.

27. Cassinotto C, Lapuyade B, Guiu B, et al. Agreement Between 2-Dimensional Shear Wave and Transient Elastography Values for Diagnosis of Advanced Chronic Liver Disease. Clin Gastroenterol Hepatol 2020;18(13):2971-2979.e3.

28. Ferraioli G, Maiocchi L, Dellafiore C, Tinelli C, Above E, Filice C. Performance and cutoffs for liver fibrosis staging of a two-dimensional shear wave elastography technique. Eur J Gastroenterol Hepatol 2021;33(1):89-95.

29. Cassinotto C, Boursier J, de Lédinghen V, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology 2016;63(6):1817-27.

30. Cassinotto C, Lapuyade B, Mouries A, et al. Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan®. J Hepatol 2014;61(3):550-7.

31. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its Associated Cutoff Point. Biometrical Journal 2005;47(4):458-472.

32. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics 1988;44(3):837-845.

33. Obuchowski NA. Estimating and Comparing Diagnostic Tests' Accuracy When the Gold Standard Is Not Binary1. Academic Radiology 2005;12(9):1198-1204.

34. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Bmj 2015;351:h5527.

35. Simera I, Moher D, Hoey J, Schulz KF, Altman DG. A catalogue of reporting guidelines for health research. Eur J Clin Invest 2010;40(1):35-53.

36. Ronot M, Ferraioli G, Müller HP, et al. Comparison of liver stiffness measurements by a 2D-shear wave technique and transient elastography: results from a European prospective multi-centre study. Eur Radiol 2021;31(3):1578-1587.

37. Lee DH, Lee ES, Lee JY, et al. Two-Dimensional-Shear Wave Elastography with a Propagation Map: Prospective Evaluation of Liver Fibrosis Using Histopathology as the Reference Standard. Korean J Radiol 2020;21(12):1317-1325.

38. Deffieux T, Gennisson J-L, Bousquet L, et al. Investigating liver stiffness and viscosity for fibrosis, steatosis and activity staging using shear wave elastography. Journal of Hepatology 2015;62(2):317-324.

39. Sumida Y, Nakajima A, Itoh Y. Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. World J Gastroenterol 2014;20(2):475-85.

40. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. Journal of Hepatology 2015;63(1):237-264.

Figure Legends

Figure 1: 2D-SWE in a patient with NAFLD and stage 2 fibrosis in liver biopsy. The split-screen image displays a color-coded 2D-SWE image (left side) and the propagation map (right side) overlaying the B-mode image. The ROI is placed in an area of parallel lines in the propagation map.

NAFLD, non-alcoholic fatty liver disease; 2D-SWE, two-dimensional shear wave elastography; ROI, region of interest.

Figure 2: Distribution of 2D-SWE measurements per fibrosis stage are shown in a box-and-whisker plot. The box represents the first and the third quartile, respectively, the median is depicted by the bar in the box, whiskers indicate values within 1.5 times the interquartile range. Each point indicates a measured median value of a patient.

2D-SWE, two-dimensional shear wave elastography.

Figure 3: Comparison of AUROC for 2D-SWE and TE for different fibrosis stages (A-C). **A**: AUROC for significant fibrosis, **B**: AUROC for severe fibrosis, **C**: AUROC curve for cirrhosis. The optimal cutoffs are displayed together with the corresponding specificity and sensitivity at this point. All available measurements were used.

AUROC, area under receiver operating characteristic; 2D-SWE, two-dimensional shear wave elastography; TE, transient elastography.

Abbreviations

TE, transient elastography

2D-SWE, two-dimensional shear wave elastography

AUROC, area under the receiver operating characteristic

95%CI, 95% confidence interval

NAFLD, non-alcoholic fatty liver disease

HBV, hepatitis B virus

HCV, hepatitis C virus

LSM, liver stiffness measurement

ROI, region of interest

BMI, body mass index

GGT, gamma-glutamyltranspeptidase

AST, aspartateaminotransferase

ALT, alanine aminotransferase

ALP, alkaline phosphatase

INR, international normalized ratio

NAS, NAFLD Activity Score

ARFI, acoustic radiation force impulse imaging

EFSUMB, European Federation for Ultrasound in Medicine and Biology

EASL, European Association for the Study of the Liver

Med, median

IQR/Med, interquartile/median ratio

SD, standard deviation

IQR, interquartile range

MCAR, missing completely at random

NASH, non-alcoholic steatohepatitis

AIH, autoimmune hepatitis

Authors contributions: Conception and design: CG, MS, CJ. Data acquisition: AK, HHB, DS, KS, SB, AD, TK, GL, SM, EF, KSch, CSR, AW, BM, MS, CG, CJ. Analysis and interpretation of the data: AK, HHB, DS, KS, EF, AW, BM, CG, CJ. Drafting of the article: AK, DS, KS, CJ. Critical revision of the article for important intellectual content: AK, HHB, DS, KS, SB, AD, SM, EF, KSch, CSR, BM, MS, CG, CJ. Final approval of the article: all authors.

Funding

SB was supported by two grants from the University of Zurich (Filling the Gap – Walter and Gertrud Siegentaler Fellowship; Fonds zur Förderung des Akademischen Nachwuchses) and by a grant from Hartmann Müller Foundation.

Conflict of interest

All authors declare that they have no conflict of interest.

Table 1: Demographic and clinical characteristics of the patient cohort

Parameter	Number
n	108
Sex = male (%)	69 (63.9)
Age (range)	50 (18-81)
BMI (range)	28.3 (17.9 – 51.2)
Diabetes (%)	29 (26.9)
Etiology group	**n (%)**
NAFLD/NASH	51 (47.2)
Alcoholic liver disease	10 (9.3)
HBV	16 (14.8)
HCV	4 (3.7)
Cholestatic hepatopathy	7 (6.5)
Miscellaneus	15 (13.9)
Normal Liver	5 (4.6)
Histopathological data	**n (%)**
Fibrosis stage	
F0	16 (16)
F1	19 (19)
F2	31 (31)
F3	26 (26)
F4	8 (8)
Standardised scale activity	

0	23 (23.0)
1	58 (58.0)
2	19 (19.0)
Laboratory parameters	**median [IQR]**
ALT (U/L)	56.0 [5.0, 96.0]
AST (U/L)	40.0 [30.2, 63]
GGT (U/L)	77.5 [35.8, 162.0]
ALP (U/L)	75.0 [58.0, 101.0]
Total bilirubin (µmol/L)	9.0 [7.0, 12.0]
Albumin (g/L)	44.0 [42.8, 47.0]
Creatinine (µmol/L)	71.5 [59.0, 85.0]
Platelet count (G/L)	217 [177.5, 264]
INR	1.1 [1.1, 1.1]

Values are given as median and range or interquartile range (IQR). Percentage is given in relation to number of patients with available data on specific parameter. Histopathological data is given for the 100 patients who received liver biopsy, in addition 7 patients were included in the F4 group based on clinical and imaging data consistent with cirrhosis.

Abbreviations: ALT, alanine aminotransferase; ALP, alkaline phosphatase; AST, aspartate aminotransferase; BMI, body mass index; GGT, gamma-glutamyltranspeptidase; HBV, hepatitis B virus; HCV hepatitis C virus; INR, international normalized ratio; IQR, interquartile range; n, numbers; NASH, Non-alcoholic steatohepatitis; NAFLD, Non-alcoholic fatty liver disease.

Table 2: Cutoff points for 2D-shear wave elastography and transient elastography for different fibrosis stages.

Fibrosis stages	Method	Cutoff (kPa)	Sensitivity (%)	Specificity (%)
F0-F1 vs. F2-F4	2D-SWE	6.5	76.4	86.1
F0-F1 vs. F2-F4	TE	6.2	90.1	76.5
F0-F2 vs. F3-F4	2D-SWE	9.9	75.6	88.1
F0-F2 vs. F3-F4	TE	9.0	87.5	78.5
F0-F3 vs. F4	2D-SWE	13.1	93.3	91.4
F0-F3 vs. F4	TE	12.1	100.0	83.5

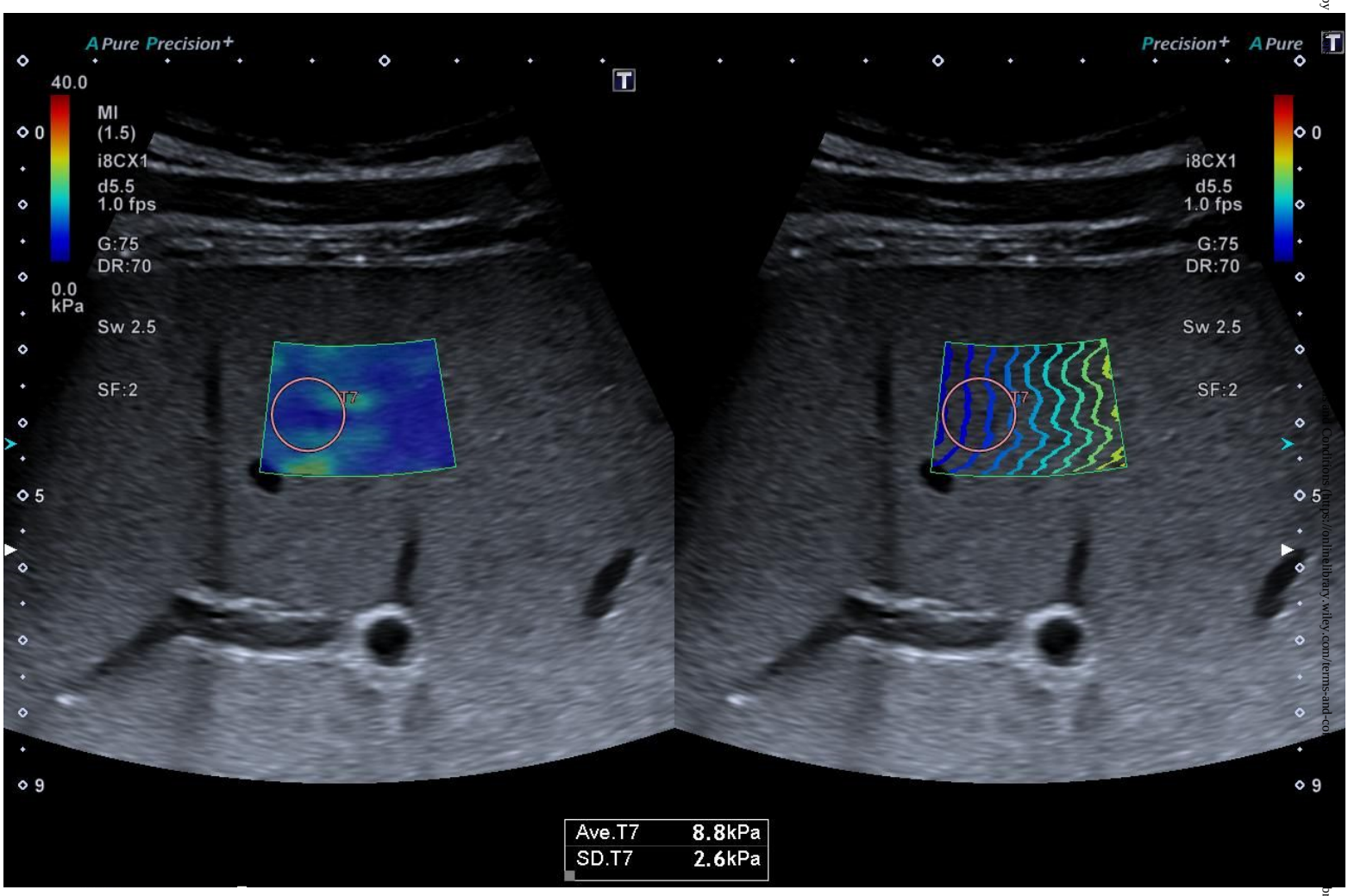
The cutoffs were estimated by maximizing Youden's Index.

2D-SWE, two-dimensional shear wave elastography; TE, transient elastography.

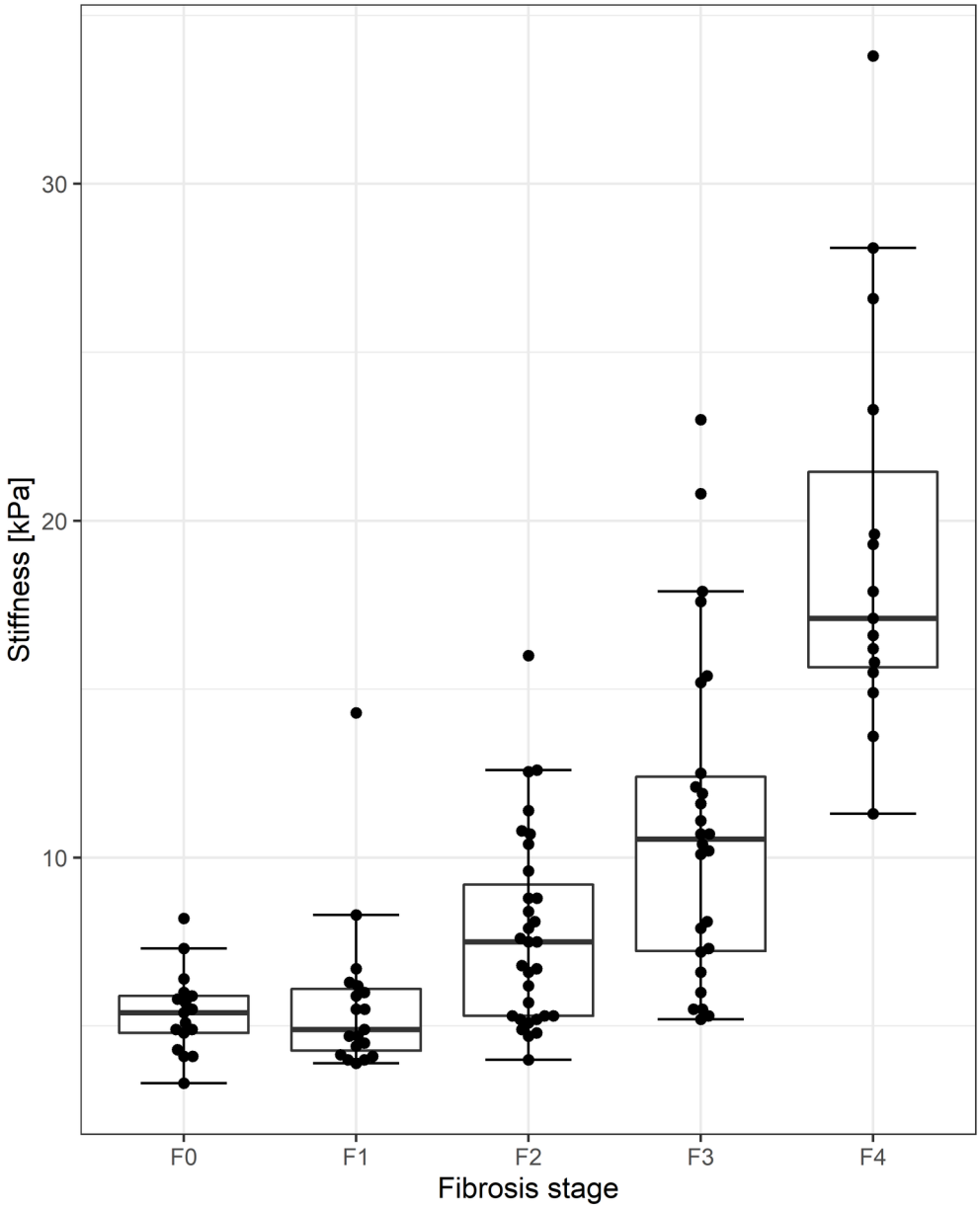
Table 3: Diagnostic accuracy of 2D-shear wave elastography and transient elastography

Fibrosis stage		AUROC (95% CI)	p-value
F0-F1 vs. F2-F4	2D-SWE	85.2 (76.2 - 91.2)	
F0-F1 vs. F2-F4	TE	87.5 (77.7 - 93.3)	
F0-F1 vs. F2-F4	Difference	-2.3 (-9.5 - 5.0)	0.54
F0-F2 vs. F3-F4	2D-SWE	86.8 (78.1 - 92.4)	
F0-F2 vs. F3-F4	TE	89.7 (82.0 - 94.3)	
F0-F2 vs. F3-F4	Difference	-2.9 (-8.1 - 2.4)	0.28
F0-F3 vs. F4	2D-SWE	95.6 (89.9 - 98.1)	
F0-F3 vs. F4	TE	96.0 (90.4 - 98.4)	
F0-F3 vs. F4	Difference	-0.4 (-4.4 - 3.5)	0.83

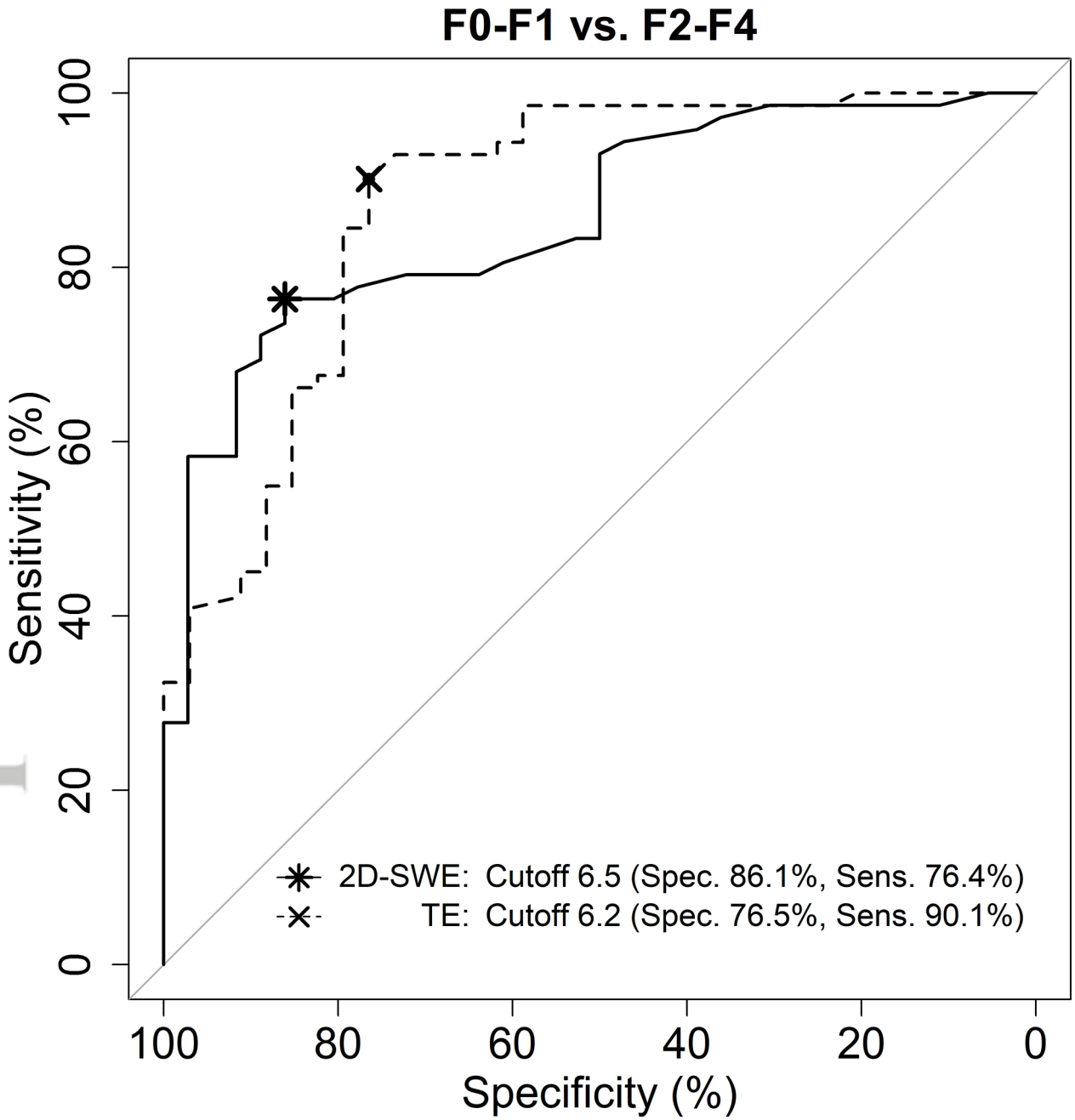
AUROC (area under the receiver operating characteristic) values including 95% confidence interval and estimated paired difference are given. 2D-SWE, two-dimensional shear wave elastography; TE, transient elastography. P-values were derived by DeLong-test.



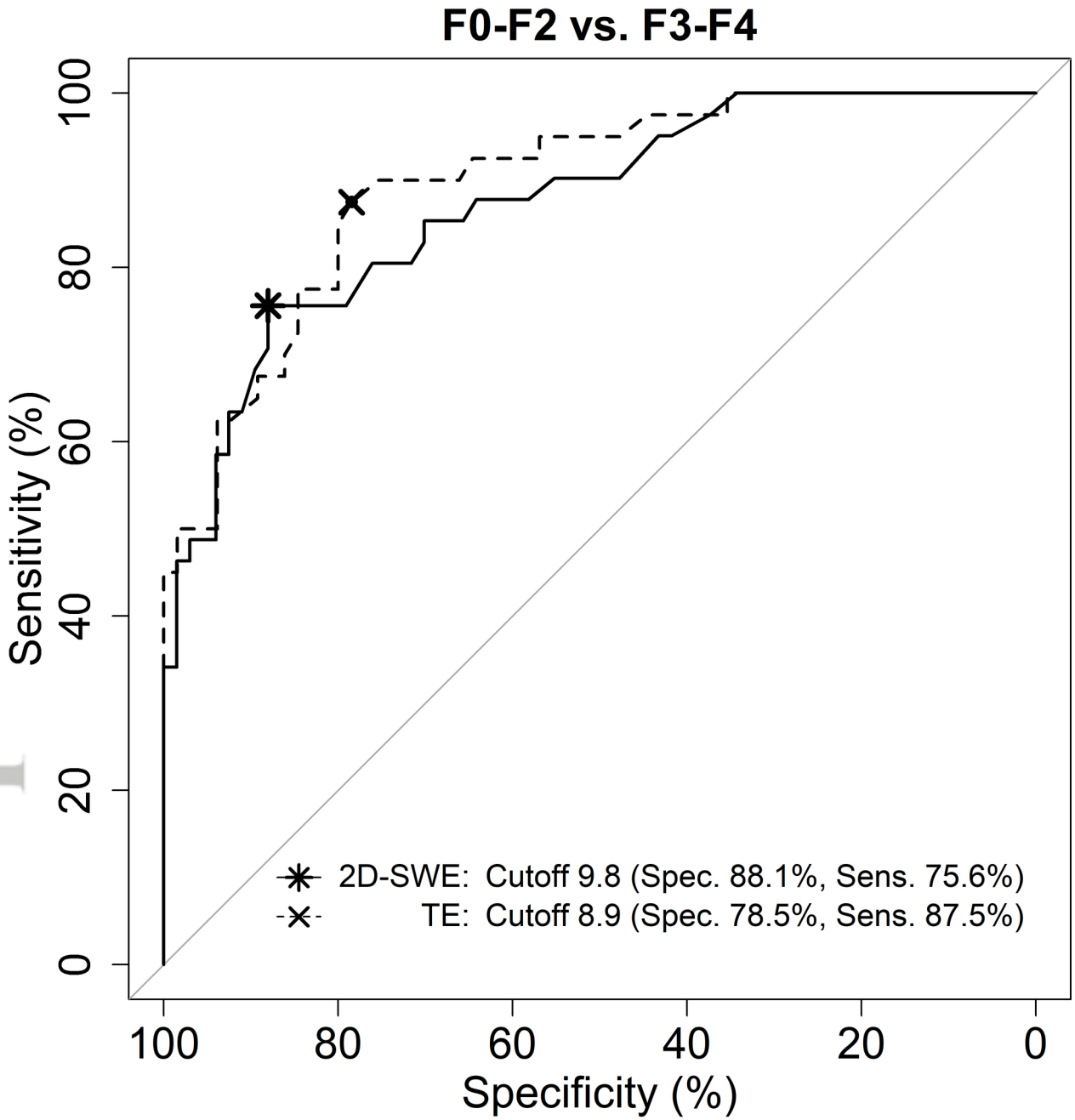
ECI_13980_Figure 1 Kovatsch et al.tiff



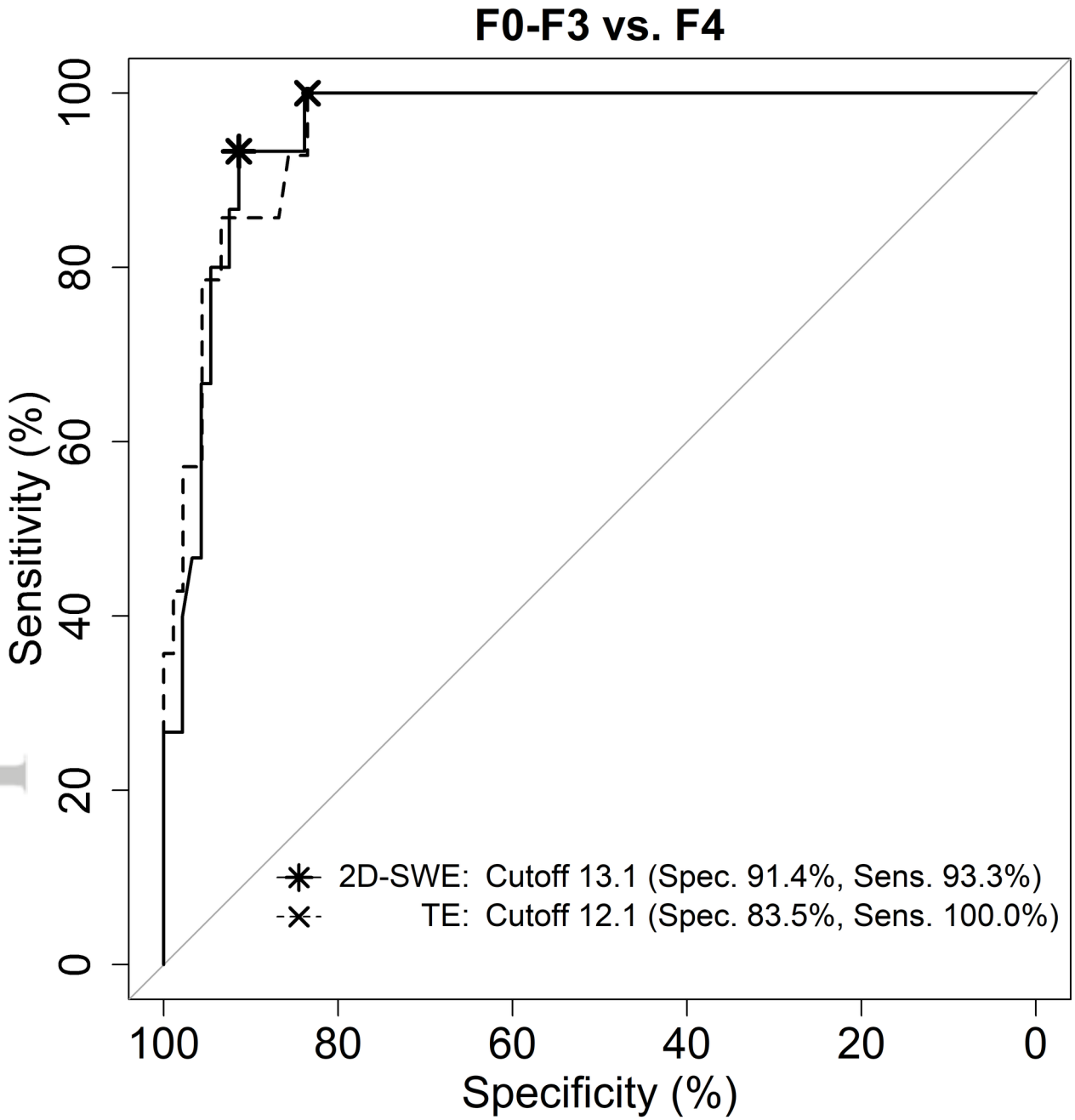
ECI_13980_Figure 2 Kovatsch et al.tiff



ECI_13980_Figure 3A Kovatsch et al.tiff



ECI_13980_Figure 3B Kovatsch et al.tiff



ECI_13980_Figure 3C Kovatsch et al.tiff