Vibration-controlled Transient Elastography to Assess Fibrosis and Steatosis in Patients With Nonalcoholic Fatty Liver Disease

[1]Mohammad S. Siddiqui, [1]Raj Vuppalanchi, [3]Mark L. Van Natta, [3]Erin Hallinan, [4]Kris V. Kowdley, [5]Manal Abdelmalek, [6]Brent A Neuschwander-Tetri, [7]Rohit Loomba, [8]Srinivasan Dasarathy, [9]Danielle Brandman, [10]Edward Doo, [3]James A. Tonascia, [11]David E. Kleiner, MD; [2]Naga Chalasani[¶]; [1]Arun J. Sanyal[¶], for the NASH Clinical Research Network.
[¶]Both are corresponding authors

SHORT TITLE: Vibration Controlled Transient Elastography in NAFLD

[1] Virginia Commonwealth University, Richmond, VA; [2]Indiana University School of Medicine, Indianapolis, IN; [3]The Johns Hopkins University School of Public Health, Baltimore, MD; [4]Swedish Medical Center, Seattle, WA; [5]Duke University, Durham, NC; [5]Saint Louis University, St Louis, MO; [7]University of California San Diego, San Diego, CA; [8]Cleveland Clinic Foundation, Cleveland, OH; [9]University of California at San Francisco, San Francisco, CA; [10]Liver Disease Branch, NIDDK, National Institutes of Health, Bethesda, MD; [11]National Cancer Institute, Bethesda, MD

Key words: Fibroscan, Vibration Controlled Transient Elastography, Controlled Attenuation Parameter, NAFLD, Fibrosis, Steatosis.
Abstract word count: 250
Word count: 3981 (including references)
Number of Tables: 4
Number of Figures: 2

AUTHOR CONTRIBUTIONS:

Patient Enrollment: MSS, RV, KVK, MA, BANT, RL, SD, DB, NC, AJS
Data Analysis: MLVN, EH, JAT
Manuscript Preparation: MSS, NC, AJS
Manuscript Review: MSS, RV, KVK, MA, BANT, RL, SD, DB, NC, AJS, MLVN, EH, JAT, ED, DEK,

Correspondence may be addressed to

Arun J Sanyal, MD at arun.sanyal@vcuhealth.org or Naga Chalasani, MD at nchalasa@iu.edu

Email addresses:

Raj Vuppalanchi: rvuppala@iu.edu
Naga Chalasani: nchalasa@iu.edu
Mark Van Natta: mvnatta@jhu.edu

LAY SUMMARY

Vibration controlled transient elastography is a non-invasive method of detecting liver fat and fibrosis in patients with nonalcoholic fatty liver disease.

Danielle Brandman: Danielle.Brandman@ucsf.edu
Arun J Sanyal: asanyal@mcvh-vcu.edu
Mohammad S. Siddiqui: mosiddiq@gmail.com
Brent Neuschwander-Tetri: brent.tetri@health.slu.edu
David E Kleiner: kleinerd@mail.nih.gov
Rohit Loomba: roloomba@ucsd.edu
Edward Doo: Dooe@extra.niddk.nih.gov
Erin Hallinan: erin.hallinan@jhu.edu
James Tonascia: James.tonascia@jhu.edu
Srinivas Dasarathy: dasaras@ccf.org
Manal Abdelmalek manal.abdelmalek@duke.edu
Kris V. Kowdley kris.kowdley@swedish.org

Abbreviations: NAFLD: Nonalcoholic Fatty Liver Disease; NASH: Nonalcoholic Steatohepatitis; NASH CRN: NASH Clinical Research Network; VCTE: Vibration Controlled Transient Elastography; LSM: Liver Stiffness Measurement; CAP: Controlled Attenuation Parameter. IQR: Interquartile range

Contributions:
Drs. Vuppalanchi, Chalasani, Brandman, Sanyal, Siddiqui, Neuschwander-Tetri, Loomba, Dasarathy, Abdelmalek, and Kowdley participated in study design, study conduct, data analysis, manuscript preparation and revision. Dr. Tonascia, Dr. Kleiner, Dr. Doo, Erin Hallnan, and Mark Van Natta participated in data analysis, manuscript preparation, and its revision.

Conflicts of Interests: Mr. Van Natta and Drs. Tonascia, Hallman, Dasarathy, Doo, Brandman, and Kleiner report no conflicts of interests. Drs. Chalasani, Abdelmalek, Sanyal, Kowdley, Neuschwander-Tetri, Vuppalanchi and Loomba have consulting agreements and/or research grants from various pharmaceutical companies but none have any consulting agreements with Echosens. The Fibroscan machines were provided by Echosens to the NASH CRN adult clinical centers through a Clinical Trial Agreement with the NIDDK. Echosens had no input into study design or data analysis but had the opportunity to review this manuscript ahead of its submission.

Source of funding (adult and pediatric)

The Nonalcoholic Steatohepatitis Clinical Research Network (NASH CRN) is supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (grants U01DK061718, U01DK061728, U01DK061731, U01DK061732, U01DK061734, U01DK061737, U01DK061738, U01DK061730, U01DK061713). Additional support is received from the National Center for Advancing Translational Sciences (NCATS) (grants UL1TR000439, UL1TR000077, UL1TR000436, UL1TR000150, UL1TR000424, UL1TR000006, UL1TR000448, UL1TR000040, UL1TR000100, UL1TR000004, UL1TR000423, UL1TR000058, UL1TR000454). This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute.

ABSTRACT

Background & Aims: Vibration-controlled transient elastography (VCTE), which measures liver stiffness, has become an important tool for evaluating patients with nonalcoholic fatty liver disease (NAFLD). We aimed to determine the diagnostic accuracy of VCTE in detection of NAFLD in a multicenter cohort of patients.

Methods: We performed a prospective study of 393 adults with NAFLD who underwent VCTE within 1 year of liver histology analysis (median time, 49 days; interquartile range, 25–78 days), from July 1, 2014 through July 31, 2017. Liver stiffness measurement (LSM) cutoffs for pairwise fibrosis stage and controlled attenuation parameter (CAP) cutoffs for pairwise steatosis grade were determined using cross-validated area under the receiver operating characteristics curve (AUROC) analyses. Diagnostic statistics were computed at sensitivity fixed at 90% and specificity fixed at 90%.

Results: LSM identified patients with advanced fibrosis with an AUROC of 0.83 (95% CI, 0.79– 0.87) and patients with cirrhosis with an AUROC of 0.93 (95% CI, 0.90–0.97). At fixed sensitivity, a cutoff LSM of 6.5 kPa excluded advanced fibrosis with a negative predictive value of 0.91; a cut-off LSM of 12.1 kPa excluded cirrhosis with a negative predictive value of 0.99. At fixed specificity, LSM identified patients with advanced fibrosis with a positive predictive 0.71 and patients with cirrhosis with a positive predictive value of 0.41. CAP analysis detected steatosis with an AUROC of 0.76 (95% CI, 0.64–0.87). In contrast, the VCTE was less accurate in distinguishing lower fibrosis stages, higher steatosis grades, or presence of NASH.

Conclusion: In a prospective study of adults with NAFLD, we found VCTE to accurately distinguish advanced vs earlier stages of fibrosis, using liver histology as the reference standard.

KEY WORDS: NAFLD, VCTE, Fibrosis, Steatosis

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is the most prevalent chronic liver disease in the U.S[1]. NAFLD exists as two predominant histological subtypes: nonalcoholic fatty liver (NAFL) and nonalcoholic steatohepatitis (NASH) [2]. NAFL is associated with a relatively benign clinical course, while NASH is associated with increased risk of progressive fibrosis and cirrhosis[3]. In NAFLD, liver biopsy remains the gold standard for diagnosis, assessing activity and staging fibrosis. However, routine use of liver biopsy is limited by its invasive nature, risk of complications, cost, sampling error, and poor patient acceptance[4,5]. This underscores an urgent need for non-invasive and accurate methods for disease detection and staging. Although, there are currently no reliable non-invasive means of differentiating NAFL from NASH, non-invasive models that correlate with individual histological parameters have been developed[6,7]. Hepatic steatosis and fibrosis are two of the most studied histological parameters as they are essential in disease diagnosis and staging, respectively. While several non-invasive methods for assessing steatosis and fibrosis have been evaluated, these all have major limitations[8].

Vibration controlled transient elastography (VCTE) measures the speed of a mechanically generated shear wave across the liver to derive a liver stiffness measurement (LSM), a marker of hepatic fibrosis[9]. Measuring the attenuation of ultrasound signal through the liver is used to derive the Controlled Attenuation Parameter (CAP), which is measured simultaneously with LSM as a marker of hepatic steatosis[10]. The performance of VCTE using the standard M probe in NAFLD was limited by high failure rates in patients with higher body mass index (BMI) and skin to liver capsule distance[11]. To circumvent the high failure rate in obese patients, an XL probe was developed[12]. To further reduce the failure rate and standardize methodology, Fibroscan 502 Touch®, a probe selection software tool that automatically determines the choice of the probe based on skin to capsule distance, has been developed. With these improvements, the failure rate of VCTE was reported to be <5%[13]. Despite the growing literature with VCTE in NAFLD, there are only a few single center studies evaluating the accuracy of both M and XL probes in American cohorts[14,15]. The aim of the current study is to examine the diagnostic accuracy of VCTE in assessing steatosis and fibrosis in a multi-center cohort of American adults with biopsy proven NAFLD.

METHODS

Study Design:

All subjects included in this study were prospectively enrolled as part of the NIH funded NASH Clinical Research Network (NASH CRN) NAFLD Database 2 study with inclusion and exclusion criteria as previously reported[13]. Eligible adult subjects (age ≥ 18 years) were enrolled across eight medical centers in the United States[13]. All subjects had biopsy-proven NAFLD within twelve months of the VCTE examination. Data were stored, monitored and analyzed at the Data Coordinating Center at the John Hopkins Bloomberg School of Public Health. The Institutional Review Boards at participating centers approved the study (NCT01030484) and all participants provided written informed consent prior to enrollment. All authors reviewed and approved the manuscript prior to submission. This study was conducted according to Transparent Reporting of a multivariate prediction model for Individual Prognosis or Diagnosis for biomarker development (see supplementary material) [16].

Study Visit and Procedures:

All subjects were evaluated at their respective medical center by a study investigator and research nurse after an overnight fast. Protocol driven anthropometric measurements, study-specific questionnaires, and blood tests were collected. All eligible subjects underwent VCTE examinations between July 1, 2014 and July 31, 2017.

Liver Biopsy:

All liver biopsies were scored for features of NAFLD using the NASH CRN scoring system by the Pathology Committee of the NASH CRN, who were blinded to the VCTE and clinical data[2]. Hepatic steatosis was graded ordinally from 0-3 [grade 0=<5% steatosis; grade 1=5-33% steatosis, grade 2=34-66% steatosis; grade 3=≥67% steatosis]. Hepatic fibrosis was quantified from stages 0-4 and for the purposes of this analysis advanced fibrosis was defined as fibrosis stage≥3 with cirrhosis as stage 4. The presence of definite NASH was defined according the NASH CRN criteria[2]. Portal inflammation, lobular inflammation and cytological ballooning was graded ordinally according the NASH CRN histological scoring system.

Vibration Controlled Transient Elastography (VCTE):

VCTE was performed using Fibroscan® 502 Touch, which were provided by Echosens (Paris, France) to all the NASH-CRN sites through a Clinical Trial Agreement with the NIDDK.

Trained study coordinators or principal investigators performed all VCTE examinations using a standardized protocol[13]. Subjects were placed in supine position with the right arm in maximal abduction and measurements were taken over the right hepatic lobe through an intercostal space[13]. All studies were started using the M probe with transition to the XL probe only if prompted by the device's automatic probe selection tool. Only cases with ≥10 valid acquisitions were used. Either the same or a different certified technician repeated the VCTE exam at the same session. The mean of the two VCTE exams was used to obtain higher statistical power due to lower variability when using mean as opposed to a single measurement. To evaluate the impact of using the first reading compared to the mean of the two VCTE examination, summary statistics between the first and second examination were compared. Unreliability of LSM was defined as IQR/Median >30% and technical failure was defined by the inability to obtain 10 valid measurements. The LSM and CAP measurements used for this analysis were the mean of the medians obtained with the 2 exams. If one exam was missing or had unreliable data, the data from the completed exam was used[13].

Statistical Analysis

Summary statistics include means, standard deviations and percentages. Diagnostic statistics include sensitivity, specificity, positive predictive value, negative predictive value and cross-validated (using jack-knife procedure) using area under the ROC (AUROC) and 95% confidence intervals. Diagnostic statistics and liver stiffness measurement (LSM) cut-offs for increasing pairwise fibrosis stages (0 vs 1-4, 0-1 vs 2-4, 0-2 vs 3-4 and 0-3 vs 4) and controlled attenuation parameter (CAP) cutoffs for increasing pairwise steatosis grades (0 vs 1-3, 0-1 vs 2-3 and 0-2 vs 3) were estimated at (1) optimized sensitivity and specificity (via Youden Index), (2) sensitivity fixed at 90% and (3) specificity fixed at 90%. Similarly, diagnostic statistics for detecting presence of NASH using LSM, CAP and the combination of CAP and LSM were determined. To evaluate the impact of the time interval between liver biopsy and VCTE, the cohort was sub-divided into those who had a liver biopsy and VCTE within versus greater than 30 days. The diagnostic accuracy of VCTE in those two cohorts was evaluated by

comparing AUROC. Finally sensitivity analysis was performed to assess the performance of VCTE between first and second measurements.

To evaluate impact of the liver histology on LSM, multiple linear regression models were constructed with steatosis, lobular and portal inflammation, ballooning, fibrosis and body mass index as candidate covariates and LSM as the outcome variable. To evaluate the impact of the liver histology on CAP, multiple linear regression models were constructed with steatosis, portal and lobular inflammation, ballooning, fibrosis and body mass index as candidate covariates and CAP as the outcome variable. Final model selection was based on Akaike's Information Criteria. Analyses were conducted using SAS (Version 9.3 of the SAS System for Windows, Cary, NC: SAS Institute Inc., 2002-2004) and Stata (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP).

RESULTS:

Study Population

A total of 393 subjects met inclusion criteria and were included in the analysis. Thirty-five subjects had missing CAP data while using the XL probe at the beginning of the study as software to compute CAP values was not available on XL probe. The median [quartiles] absolute value of time from liver biopsy to VCTE was 49 (25, 78) days. The mean (±SD) age and BMI of the cohort was 51±11 years and 34±6kg/m^2, respectively (**Table 1**). The distribution of biopsy fibrosis stage 0, stage 1, stage 2, stage 3, and stage 4 was 24%, 25%, 19%, 23%, and 9%, respectively. The distribution of biopsy steatosis grade for grade 0, 1, 2, and 3 was 5%, 38%, 30% and 27%, respectively. Twenty-one (2.7%) of the 786 LSM measurements had unreliable results, and the failure rate was 3.7% (reasons for failure: 7 subjects had skin-to-capsule distance >3.5cm; 4 cases where the machine was not working or available and 4 cases where the patient stopped or refused).

Performance Diagnostics of Liver Stiffness Measurements

The median LSM scores for fibrosis stages 0, 1, 2, 3, and 4 were 5.5[4.5, 7.4], 6.5[5.0, 8.8], 7.7[6.6, 10.6], 11.2[8.3, 13.8], and 23.2[14.8, 45.8] kPa, respectively (**Figure 1**). There were two participants with stage 0 fibrosis but outlier LSM values of 69.2kPa and 45.1kPa. The first patient's examinations had LSM values of 69.2kPa and 45.0kPa on first and second exam with IQR/median of 15% and 44%, respectively by the

same performer. According to the study design, the results of the second exam were excluded since IQR/median was >30%. On histology, this participant had NAFLD with NAS=1 and had BMI of 32.9kg/m^2. The second patient's examinations had LSM of 19.6kPa (IQR/median=12%) and 70.6kPa (IQR/median=15%) using different examiners with the average value of 45.1kPa. On histology, the participant had NAFLD with NAS=2 and had BMI of 45.0kg/m^2.

The cross-validated AUROC for classifying fibrosis stage 0 from stages 1-4 was: 0.74 (95% CI 0.68, 0.79); fibrosis stages 0-1 from stages 2-4 was: 0.79 (0.74-0.83); fibrosis stages 0-2 from stages 3-4 was: 0.83 (0.79, 0.87); and fibrosis stages 0-3 was: 0.93 (0.90, 0.97) (**Table 2**). The LSM cutoff values with sensitivity fixed at 90% for differentiating between dichotomous fibrosis stages are as follows: 4.9kPa for stages 0 vs. stages 1-4; 5.6kPa for stage 0-1 vs. stages 2-4; 6.5kPa for stages 0-2 vs. stages 3-4; and 12.1kPa for stages 0-3 vs. stage 4. Using these LSM cutoff values, the PPV was 0.80, 0.62, 0.45 and 0.34 and NPV was 0.48, 0.80, 0.91, and 0.99 for discriminating between stage 0 vs. stages 1-4, stage 0-1 vs. stages 2-4, stages 0-2 vs. stages 3-4, and stage 0-3 vs. stage 4, respectively (**Table 2**). In contrast, with specificity fixed at 90%, the LSM cutoff values for discriminating fibrosis stage 0 vs. stages 1-4, stages 0-1 vs. stages 2-4, stages 0-2 vs. stages 3-4, and stages 0-3 vs. stage 4 were 9.4kPa, 11.9kPa, 12.1kPa and 14.9kPa, respectively. The PPV was 0.93, 0.80, 0.71, and 0.41, respectively for differentiating between stage 0 vs. stages 1-4, stage 0-1 vs. stages 2-4, stages 0-2 vs. stages 3-4, and stage 0-3 vs. stage 4, while corresponding NPV were 0.34, 0.59, 0.80, and 0.97 (**Table 2**). Finally, the cutoff value optimizing sensitivity and specificity for differentiating stage 0 from stages 1-4 was 8.6kPa; stages 0-1 vs. stages 2-4 was 8.6kPa; stages 0-2 vs. stages 3-4 was 8.6kPa; and stages 0-3 vs. stage 4 was 13.1kPa (**Table 2**). The diagnostic accuracy of LSM was not altered by the time interval between liver biopsy and VCTE (**Table 3**). Finally, sensitivity analysis showed no difference between LSM measurements from first and second exam (**Supplemental Table 1 and 2**).

Performance Diagnostics of Controlled Attenuation Parameter

The median CAP scores for steatosis grade 0, 1, 2 and 3 were 274[244, 281], 306[270, 338], 340[312, 369], and 340[311, 360] dB/m (**Figure 2**). The cross-validated AUROC for classifying steatosis grade 0 vs. grade 1-2, steatosis grade 0-1 vs. 2-3, and steatosis grade 0-2 vs. 3 were 0.76 (95% CI: 0.64, 0.89), 0.70 (0.64, 0.75), and 0.58

(0.51, 0.64), respectively (**Table 4**). At sensitivity fixed at 90%, a cutoff value 263dB/m provided 0.35 specificity, 0.96 PPV, and 0.15 NPV for detecting presence of ≥5% steatosis. When the specificity was fixed at 90%, a cutoff value 353dB/m provided sensitivity of 0.29, PPV of 0.98 and NPV of 0.06. The cutoff values for differentiating between steatosis grade 0-1 vs. 2-3 and steatosis grade 0-2 vs. 3 at 90% fixed sensitivity were 280dB/m and 274dB/m and at 90% fixed specificity were 367dB/m and 380dB/md specificity. The cutoff values optimizing sensitivity and specificity for differentiating steatosis grade 0 vs. grade 1-3 was 285dB/m; grade 0-1 vs. grade 2-3 was 311dB/m; and grade 0-2 vs. grade 3 was 306dB/m (**Table 3**). Finally, the diagnostic accuracy of CAP was similar whether the time interval between liver biopsy and VCTE was less than 30 days or more than 30 days (**Table 3**). Using sensitivity analysis, there was no difference between CAP measurements obtained between first and second exam (**Supplemental Table 1 and 2**).

Regression Models

In regression analysis, fibrosis (β-coefficient 4.3kPa/stage [95% CI: 3.4, 5.2], P<0.001) and body mass index (β-coefficient 0.12kPa/kg/m^2 [-0.03, 0.27], p=0.10) were directly related to LSM, while an inverse relationship between steatosis grade (β-coefficient -1.8 kPa/grade [-2.9, -0.7], P=0.001) and ballooning (β-coefficient -1.1kPa/grade [-2.5, 0.4], p=0.16) were found. Portal and lobular inflammation were not related to LSM. A direct and significant relationship between CAP and steatosis (β-coefficient of 17dB/m/grade [12, 22], P<0.001), portal inflammation (β-coefficient -5.9dB/m/grade [-13.0, 1.2], P=0.10) and body mass index (β-coefficient 2.8dB/m/kg/m^2 [2.1, 3.5], p<0.001) were found (**Supplemental Table 3**).

Although BMI was significantly related to both LSM and CAP, the diagnostic performance of LSM for assessing fibrosis and CAP for assessing steatosis did not vary by BMI category (**Supplemental Table 4).** The relationship between steatosis grade and LSM did not vary by presence or absence of advance fibrosis. Similarly, after adjusting for BMI, no significant relationship between LSM and CAP was noted (data not shown).

Diagnostic Accuracy of VCTE in Predicting NASH

Among 358 subjects with definite NASH, the cross-validated AUROC for LSM was 0.74 (95% CI: 0.68, 0.79) with OR= 1.078 (1.034, 1.123) per kPa (P<0.001) for detecting the presence of NASH. The cross-validated AUROC for CAP was 0.58 (0.52,

0.64) in detecting NASH with OR=1.007 (1.002, 1.011) per dB/m (P=0.003). Finally, the model with both LSM and CAP had an AUROC of 0.71 (0.66, 0.76) in diagnosing NASH with LSM OR=1.071 (1.028, 1.115) per kPa (P=0.001) and CAP OR=1.006 (1.001, 1.011) per dB/m (P=0.02).

DISCUSSION:

An important unmet need in NAFLD is a point of care test that can aid in detection and identification of advance fibrosis. VCTE can simultaneously detect steatosis and fibrosis, but there is paucity of data defining optimal use of VCTE in American cohorts[14,15]. The current study evaluates the diagnostic accuracy of VCTE in a multicenter cohort with histologically confirmed NAFLD to assess parameters for clinical use by identifying threshold that are highly sensitive or specific.

Early detection of NAFLD is vital to allow sufficient time to implement strategies aimed at favorably altering the natural history of the disease. The CAP value is positively associated with severity of hepatic steatosis and the cross-validated AUROC is 76% for classifying patients with ≥5% steatosis on histology. This cutoff (CAP 263dB/m) is similar to the previously proposed cutoff in U.S. cohort[17]. In addition to clinical care, the CAP value may also be used as an adjunct tool in regulatory science to allow for subject enrichment in early phase clinical trials with non-histological endpoints. A CAP value < 274 dB/m has 84% NPV for grades 0-2 steatosis (i.e., excludes grades 3 steatosis) suggesting that cut-off may offer some clinical and research utility. In contrast, the accuracy of CAP in separating steatosis grade, particularly grade 2 and 3, was sub-optimal, a finding that confirms prior reports[10,14].

In NAFLD, hepatic fibrosis is a key predictor of liver related outcomes[3,18] and VCTE can be used to detect fibrosis, especially in its advance stage. Although VCTE is not a confirmatory test, it can help identify patients in whom additional histological assessment maybe warranted, while avoiding liver biopsies in patients with none to minimal fibrosis. Identifying optimal cutoff values of VCTE depends on the context of use for VCTE. Non-invasive biomarkers aim to either to minimize false negatives (i.e. high sensitivity) or to minimize false positives (i.e. high specificity) depending on whether VCTE is being used as screening modality or a tool to identify NAFLD patients with fibrosis with great degree of certainty. Moderate fibrosis is linked to liver related outcomes and mortality[18], and a LSM <5.6kPa has a NPV of 80% for excluding moderate fibrosis. Similarly, a less invasive approach can be employed in patients with a LSM <6.5kPa since the presence of advance fibrosis can be excluded with at least 91% certainty. While higher LSM values allow for greater specificity and can be used to identify individuals in whom additional confirmatory histological assessment maybe warranted. Furthermore, we also applied cutoffs proposed by Baveno IV consensus for detection of advance fibrosis in our cohort and the published data[15,19]. The cutoff values

of >9.9kPa had a PPV of 46% and 64% for detecting advance fibrosis in the cohorts studied by Tapper et. al. and the NASH CRN, respectively (**Supplemental Table 3**). The higher PPV observed in the NASH CRN cohort is likely due to higher prevalence of advance fibrosis within the NASH CRN cohort (32% vs. 18%). Conversely, using a cutoff value >15kPa yielded a NPV of 75% with in the NASH CRN cohort. These findings is in line with the assertion that VCTE has good accuracy at extremes with low LSM essentially ruling out advanced disease and higher LSM values ruling in cirrhosis[20]. An interesting inverse relationship between LSM and steatosis grade and cytological ballooning was noted as has been reported previously[21]. This likely represents disappearance of classic histological components of NAFLD as patients progress to advance fibrosis[22]. Although inflammation has been shown to impact LSM in patients with chronic liver disease, no such association was noted in the current study[23,24]. This is likely due to the fact that inflammation in NAFLD is often less severe than is found in viral hepatitis. Finally, the diagnostic accuracy of VCTE for distinguishing NAFL from NASH was also poor.

There are several notable strengths of the current study. This multicenter study evaluated the accuracy of VCTE using both M and XL probes and Fibroscan® 502 Touch software in a US cohort using a standardized and uniform protocol. Due to the multicenter design, the results are more generalizable than previously reported single center experiences[14,15]. The sample size of the current study is also larger than prior U.S. studies with more equal distribution of histological parameters, particularly steatosis and fibrosis. Finally, we found that a single patient scan for both LSM (S.D.=10.9 kPa) and CAP (S.D.-50 dB/m) are nearly as precise as the average of the two scans LSM (S.D.=11.0 kPa) and CAP (S.D.= 48 dB/m) with no bias between the first and second scans, thus a single scan can be used, unless there is some reason other than increased precision to do so.

A potential limitation of the study is that VCTE and liver biopsy were not performed simultaneously. However, since fibrosis evolves slowly, it is unlikely that the relatively short delay between biopsy and VCTE had any significant impact on LSM. Although, the delay between liver biopsy and VCTE did not impact the diagnostic accuracy of LSM or CAP, the power to detect such interactions was low. The current study evaluated patients enrolled in an observational research study, and the diagnostic performance of VCTE cannot be extrapolated primary care clinics where the prevalence

and the severity of disease may be different. Thus, the PPV and NPV reported in the NASH CRN cohort maybe different than in primary care clinics.

In summary, VCTE is a non-invasive point of care tool that can be used in clinical practice for identifying steatosis and advance fibrosis in patients with NAFLD. VCTE may be useful in identifying patients in whom additional histological assessment may be warranted due to the presence of advance fibrosis, while excluding patients without significant fibrosis in whom a liver biopsy may be unnecessary.

REFERENCES

1. Younossi ZM, Stepanova M, Afendy M, et al. Changes in the prevalence of the most common causes of chronic liver diseases in the united states from 1988 to 2008. *Clin Gastroenterol Hepatol*. 2011;9(6):524-530.e1; quiz e60.

2. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology*. 2005;41(6):1313-1321.

3. Ekstedt M, Franzen LE, Mathiesen UL, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology*. 2006;44(4):865-873.

4. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver Diseases. Liver biopsy. *Hepatology*. 2009;49(3):1017-1044.

5. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology*. 2005;128(7):1898-1906.

6. Siddiqui MS, Patidar KR, Boyett S, Luketic VA, Puri P, Sanyal AJ. Performance of non-invasive models of fibrosis in predicting mild to moderate fibrosis in patients with non-alcoholic fatty liver disease. *Liver Int*. 2016;36(4):572-579.

7. Cui J, Ang B, Haufe W, et al. Comparative diagnostic accuracy of magnetic resonance elastography vs. eight clinical prediction rules for non-invasive diagnosis of advanced fibrosis in biopsy-proven non-alcoholic fatty liver disease: A prospective study. *Aliment Pharmacol Ther*. 2015;41(12):1271-1280.

8. European Association for Study of Liver, Asociacion Latinoamericana para el Estudio del Higado. EASL-ALEH clinical practice guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol*. 2015;63(1):237-264.

9. Wong VW, Vergniol J, Wong GL, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology*. 2010;51(2):454-462.

10. Ledinghen V, Wong GL, Vergniol J, et al. Controlled attenuation parameter for the diagnosis of steatosis in non-alcoholic fatty liver disease. *J Gastroenterol Hepatol*. 2016;31(4):848-855.

11. Castera L, Foucher J, Bernard PH, et al. Pitfalls of liver stiffness measurement: A 5-year prospective study of 13,369 examinations. *Hepatology*. 2010;51(3):828-835.

12. Wong VW, Vergniol J, Wong GL, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. *Am J Gastroenterol*. 2012;107(12):1862-1871.

13. Vuppalanchi R, Siddiqui MS, Van Natta ML, et al. Performance characteristics of vibration-controlled transient elastography for evaluation of non-alcoholic fatty liver disease. *Hepatology*. 2017.

14. Park CC, Nguyen P, Hernandez C, et al. Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease. *Gastroenterology*. 2017;152(3):598-607.e2.

15. Tapper EB, Challies T, Nasser I, Afdhal NH, Lai M. The performance of vibration controlled transient elastography in a US cohort of patients with nonalcoholic fatty liver disease. *Am J Gastroenterol*. 2016;111(5):677-684.

16. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63.

17. Park CC, Nguyen P, Hernandez C, et al. Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease. *Gastroenterology*. 2017;152(3):598-607.e2.

18. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2015;149(2):389-397.e10.

19. de Franchis R, Baveno VI Faculty. Expanding consensus in portal hypertension: Report of the baveno VI consensus workshop: Stratifying risk and individualizing care for portal hypertension. *J Hepatol*. 2015;63(3):743-752.

20. Boursier J, Vergniol J, Guillet A, et al. Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. *J Hepatol*. 2016;65(3):570-578.

21. Petta S, Maida M, Macaluso FS, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease. *Hepatology*. 2015;62(4):1101-1110.

22. Caldwell SH, Lee VD, Kleiner DE, et al. NASH and cryptogenic cirrhosis: A histological analysis. *Ann Hepatol*. 2009;8(4):346-352.

23. Kim SU, Kim JK, Park YN, Han KH. Discordance between liver biopsy and fibroscan(R) in assessing liver fibrosis in chronic hepatitis b: Risk factors and influence of necroinflammation. *PLoS One*. 2012;7(2):e32233.

24. Lim SG, Cho SW, Lee YC, et al. Changes in liver stiffness measurement during antiviral therapy in patients with chronic hepatitis B. *Hepatogastroenterology*. 2011;58(106):539-545.

Table 1. Selected Characteristics of the Study Population

	Mean ± SD or n (%)
N	393
Age –years	51±11
Gender - male	127 (32%)
Race - white	314 (80%)
Ethnicity - Hispanic	49 (13%)
LABORATORY	
AST (U/L)	49±37
ALT (U/L)	64±44
Alkaline phosphatase (U/L)	83±32
GGT (U/L)	70±83
Bilirubin, total (mg/dL)	0.7±0.6
International normalized ratio	1.04±0.13
Platelet count (1000 cells/uL)	235±72
METABOLIC FACTORS	
Body mass index (kg/m^2) Diabetes Severe obesity (BMI ≥35 kg/m2) Dyslipidemia	34.4±6.4 170 (44%) 163 (42%) 221 (57%)
HISTOLOGY	
Ballooning mean grade Grade 0 Grade 1 Grade 2	0.9±0.8 143 (36%) 132 (34%) 118 (30%)
Lobular inflammation – mean grade Grade 0 Grade 1 Grade 2 Grade 3	1.6±0.7 5 (1%) 211 (54%) 130 (33%) 47 (12%)
Steatosis – mean grade Grade 0 Grade 1 Grade 2 Grade 3	1.8±0.9 19 (5%) 150 (38%) 119 (30%) 105 (27%)
NAFLD Activity Score (NAS)	4.3±1.7
Portal inflammation – mean grade Grade 0 Grade 1 Grade 2	1.2±0.6 45 (11%) 234 (60%) 114 (29%)
Fibrosis – mean stage	1.7±1.3
Stage 0	94 (24%)
Stage 1	99 (25%)
Stage 2	73 (19%)
Stage 3	91 (23%)
Stage 4	36 (9%)
Definite NASH	225 (57%)
Time from biopsy to VCTE – absolute value (days)	
Mean±SD	64±64
Median [IQR]	49 [25, 78]

Table 2. Performance diagnostics of liver stiffness measurement assessing liver fibrosis stage

Fibrosis stage: Non-event vs event	Prevalence of event	Cross-validated AUROC (95% CI)	Cutoff Criteria	Cutoff (kPa)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
0 vs 1-4	76%	0.74 (0.68, 0.79)	Sensitivity = 90%	4.9	0.90	0.31	0.80	0.48
			Specificity = 90%	9.4	0.46	0.90	0.93	0.34
			Youden's index	8.6	0.53	0.87	0.93	0.37
0-1 vs 2-4	51%	0.79 (0.74, 0.83)	Sensitivity = 90%	5.6	0.90	0.44	0.62	0.81
			Specificity = 90%	11.9	0.40	0.90	0.80	0.59
			Youden's index	8.6	0.66	0.80	0.78	0.70
0-2 vs 3-4	32%	0.83 (0.79, 0.87)	Sensitivity = 90%	6.5	0.90	0.47	0.45	0.91
			Specificity = 90%	12.1	0.52	0.90	0.71	0.80
			Youden's index	8.6	0.80	0.74	0.59	0.89
0-3 vs 4	9%	0.93 (0.90, 0.97)	Sensitivity = 90%	12.1	0.90	0.82	0.34	0.99
			Specificity = 90%	14.9	0.69	0.90	0.41	0.97
			Youden's index	13.1	0.89	0.86	0.39	0.99

Abbreviations: AUROC; area under the receiver operating characteristic

Table 3. AUROCs by length of time between VCTE exam and biopsy

***not cross-validated**

Predictor	Outcome	AUROC*		P-value from test of independence of AUROCs
		VCTE exam and biopsy within 30 days (n=119)	VCTE exam and biopsy outside of 30 days (n=274)	
LSM	Fibrosis stage 0 vs 1-4	0.76	0.75	0.85
	Fibrosis stage 0-1 vs 2-4	0.76	0.80	0.44
	Fibrosis stage 0-2 vs 3-4	0.87	0.82	0.27
	Fibrosis stage 0-3 vs 4	0.95	0.93	0.56
CAP	Steatosis grade 0 vs 1-3	0.86	0.75	0.33
	Steatosis grade 0-1 vs 2-3	0.77	0.67	0.09
	Steatosis grade 0-2 vs 3	0.63	0.58	0.41

Table 4: Performance diagnostics of controlled attenuation parameter in assessing steatosis grade

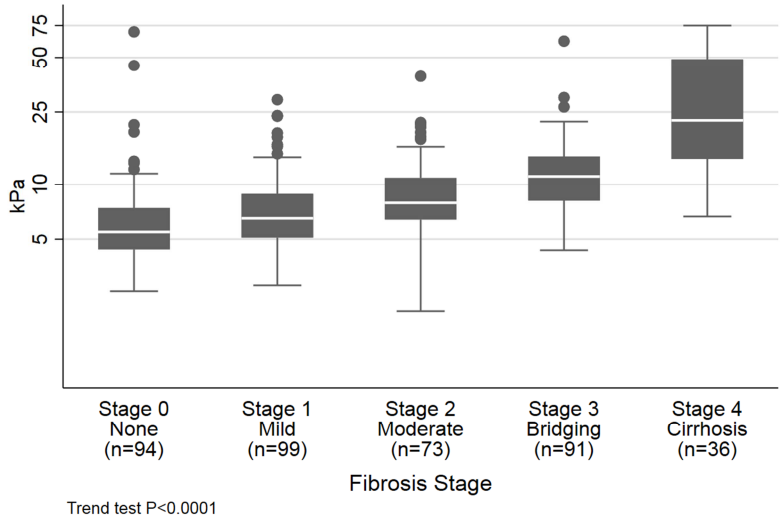
Steatosis grade: Non-event vs event	Prevalence of event	Cross-validated AUROC (95% CI)	Cutoff Criteria	Cutoff (dB/m)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
0 vs 1-3	95%	0.76 (0.64, 0.89)	Sensitivity = 90%	263	0.90	0.35	0.96	0.15
			Specificity = 90%	353	0.29	0.90	0.98	0.06
			Youden's index	285	0.80	0.77	0.99	0.16
0-1 vs 2-3	58%	0.70 (0.64, 0.75)	Sensitivity = 90%	280	0.90	0.35	0.64	0.72
			Specificity = 90%	367	0.20	0.90	0.70	0.46
			Youden's index	311	0.77	0.57	0.70	0.66
0-2 vs 3	27%	0.58 (0.51, 0.64)	Sensitivity = 90%	274	0.90	0.20	0.29	0.84
			Specificity = 90%	380	0.03	0.90	0.10	0.72
			Youden's index	306	0.80	0.40	0.32	0.85

Abbreviations: AUROC; area under the receiver operating characteristic,

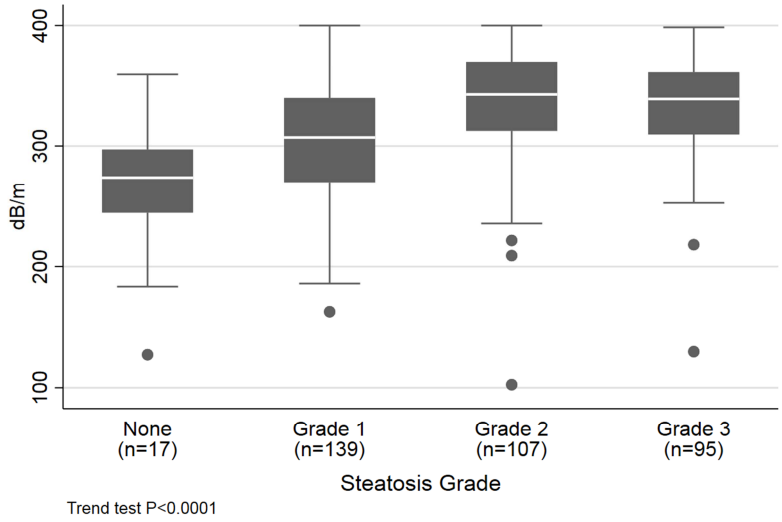
FIGURE LEGEND:

Figure 1. Liver Stiffness Measurement According to Fibrosis Stage

Figure 2. Controlled Attenuation Parameter According to Steatosis Grade



Trend test P<0.0001



Appendix Table 1. Comparison of summary statistics between first vs. second VCTE exam

	First exam	Second exam	Mean (First, Second) exam	Difference (First – Second) exam
LSM – kPa				
N	385	375	393	367
Mean	10.9	11.0	11.0	-0.2*
SD	10.9	11.4	11.0	4.4
CAP – dB/m				
N	358	352	358	352
Mean	319	320	319	0†
SD	50	52	48	36

*P-value from t-test of Difference=0 is 0.32

†P-value from t-test of Difference=0 is 0.97

Appendix Table 2. Comparison of diagnostic performance between first vs. second VCTE exam

Predictor	Outcome	AUROC		P-value
		First exam	Second exam	
LSM	Fibrosis stage 0 vs 1-4	0.74	0.76	0.72
	Fibrosis stage 0-1 vs 2-4	0.80	0.79	0.78
	Fibrosis stage 0-2 vs 3-4	0.84	0.83	0.70
	Fibrosis stage 0-3 vs 4	0.93	0.94	0.53
CAP	Steatosis grade 0 vs 1-3	0.78	0.76	0.86
	Steatosis grade 0-1 vs 2-3	0.70	0.68	0.64
	Steatosis grade 0-2 vs 3	0.61	0.56	0.25

Appendix Table 3. Linear regressions of Liver Stiffness Measurement (LSM) on NAFLD Activity Score (NAS) stratified by fibrosis stage

		kPa / NAS		
Fibrosis stage	**N**	**Slope**	**95% CI**	**P-value**
0	94	-1.7	-3.1, -0.3	0.02
1	99	0.6	0.1, 1.0	0.01
2	73	0.2	-0.4, 0.8	0.56
3	91	1.2	0.3, 2.2	0.01
4	36	-4.1	-7.8, -0.3	0.03

Note: P-value for test of interaction of fibrosis stage by NAS on LSM < 0.001

Appendix Table 4. Area under the receiver operating characteristic (AUROC) for liver stiffness measurement assessing fibrosis stage and controlled attenuation parameter assessing steatosis grade by body mass index* (BMI)

		AUROC			
Outcome	**Non-event vs event comparison**	**BMI < 30 kg/m^2**	**BMI ≥ 30 & < 35 kg/m^2**	**BMI ≥ 35 kg/m^2**	**P-value†**
Fibrosis stage		N=107	N=118	N=163	
	0 vs 1-4	0.75	0.81	0.68	0.22
	0-1 vs 2-4	0.80	0.84	0.72	0.08
	0-2 vs 3-4	0.89	0.85	0.77	0.06
	0-3 vs 4	0.96	0.91	0.94	0.50
Steatosis grade		N=104	N=104	N=145	
	0 vs 1-3	0.79	0.90	0.68	0.20
	0-1 vs 2-3	0.80	0.64	0.71	0.07
	0-2 vs 3	0.68	0.55	0.61	0.29

*5 participants had missing bmi data

†Based on test of equality of AUROCs across 3 bmi categories (ROC analysis of independent samples; Stata 15.1, 2017)

TRIPOD Checklist: Prediction Model Validation

Section/Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	3
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4-5
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5-7
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	6
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	6
	5b	Describe eligibility criteria for participants.	5-6
	5c	Give details of treatments received, if relevant.	n/a
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	Report any actions to blind assessment of the outcome to be predicted.	6
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	7-8
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	n/a
Sample size	8	Explain how the study size was arrived at.	n/a
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	5-7
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	7
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	7
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	n/a
Risk groups	11	Provide details on how risk groups were created, if done.	n/a
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	n/a
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	7-8
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	6-7
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	7-8
Model performance	16	Report performance measures (with CIs) for the prediction model.	8-11
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	n/a
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	13
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	11-12
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	11-13
Implications	20	Discuss the potential clinical use of the model and implications for future research.	13
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	n/a
Funding	22	Give the source of funding and the role of the funders for the present study.	2

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

EDITOR'S NOTE:

BACKGROUND AND CONTEXT

Nonalcoholic fatty liver disease (NAFLD) is common in the United States and hepatic fibrosis is a key predictor of liver related outcomes in NAFLD. Vibration controlled transient elastography (VCTE) is a non-invasive biomarker that utilizes shear wave elastography to estimate hepatic fibrosis.

NEW FINDINGS

VCTE has high diagnostic accuracy for identifying presence of advance fibrosis and cirrhosis in patients with NAFLD.

LIMITATIONS

The study did not evaluate the impact of VCTE on clinical outcomes.

IMPACT

VCTE can be used as a clinical tool in management of patients with NAFLD.