Using vibration controlled transient elastography and FIB-4 to assess liver cirrhosis in a hepatitis C virus infected population

Nabil Saleem, MD<sup>a</sup>,<sup>∗</sup>, Lesley S. Miller, MD<sup>a</sup>, Alia S. Dadabhai, MD<sup>a</sup>, Emily J. Cartwright, MD<sup>a,b</sup>

**Abstract**

We assessed the performance characteristics of the Fibrosis-4 (FIB-4) score in a veteran population with chronic hepatitis C virus (HCV) infection and used vibration controlled transient elastography (VCTE) as the gold standard.

All VCTE studies were performed by a single operator on United States veterans with HCV infection presenting for care at the Atlanta VA Medical Center (AVAMC) over a 2 year period. VCTE liver stiffness measurements (LSM) were categorized as cirrhotic if LSM was $>12.5$ kPa and non-cirrhotic if LSM was $\leq 12.5$ kPa. FIB-4 scores $<3.25$ were considered non-cirrhotic and scores $>3.25$ were considered cirrhotic. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for the FIB-4 score. A second analysis was done which identified and excluded indeterminate FIB-4 scores, defined as any value between 1.45 and 3.25.

When FIB-4 was used to screen for liver cirrhosis using VCTE as the gold standard, sensitivity was 42%, specificity was 88%, PPV was 62%, and NPV was 76%. When indeterminate FIB-4 scores were excluded from the analysis, sensitivity was 95%, specificity was 61%, PPV was 62%, and NPV was 94.4%. In a veteran population with chronic HCV infection, we found the sensitivity of the FIB-4 score to be unacceptably low for ruling out liver cirrhosis when using a binary cutoff at 3.25. Using a second staging method like VCTE may be an effective way to screen for liver cirrhosis in persons with chronic HCV, especially when the FIB-4 score is in the indeterminate range.

**Abbreviations:** CPRS = computerized patient record system, FIB-4 = fibrosis 4 index, HCV = hepatitis C virus, HIV = human immunodeficiency virus, LSM = liver stiffness measurement, VCTE = vibration controlled transient elastography.

**Keywords:** fibrosis 4 index, fibrosis, HCV treatment, hepatitis C, liver fibrogenesis, vibration controlled transient elastography, viral hepatitis

1. Introduction

Management of chronic hepatitis C virus (HCV) infection relies on accurate staging of liver fibrosis to select an appropriate antiviral regimen and duration. Accurate staging is particularly important for those with advanced liver fibrosis, as it informs decisions about hepatocellular carcinoma (HCC) surveillance and esophageal varices screening in this population. While liver biopsy has traditionally been the gold standard for staging liver fibrosis, noninvasive methods of assessing liver fibrosis, including direct serologic markers, indirect serologic markers, and elastography are increasingly utilized in clinical practice.

The fibrosis-4 (FIB-4) score is an indirect serologic marker of liver fibrosis that is calculated using readily available clinical laboratory tests and patient age. FIB-4 scores have been found to predict liver fibrosis in a variety of populations, including those with HCV monoinfection or Human Immunodeficiency Virus (HIV)/HCV co-infection. Vibration Controlled Transient Elastography (VCTE) is a noninvasive technique that uses a handheld probe to deliver ultrasound shear waves into the liver. Wave propagation speed is measured by a receiver in the probe and is used to calculate liver stiffness measurements ([LSM] in units of kilopascal, kPa) using Hook law. Liver stiffness measurements have been validated against liver biopsy results to create cut-off ranges for fibrosis stages; cut-off ranges vary based on the underlying etiology of liver disease.

In late 2019, the Infectious Diseases Society of America (IDSA) and the American Association for the Study of Liver Diseases (AASLD) released updated guidance on the evaluation and treatment of persons with HCV infection. Liver biopsy is not required in the pretreatment liver cirrhosis assessment, and classification of probable liver cirrhosis status informs further treatment and surveillance strategies. According to this guidance, cirrhosis can be determined using FIB-4 score alone. A person is
presumed to have cirrhosis if the pretreatment FIB-4 score is greater than 3.25; conversely, a patient is presumed non-cirrhotic if the pretreatment FIB-4 score is less than or equal to 3.25.\[9\]

In light of this updated guidance, we undertook an assessment evaluating the performance of FIB-4 scores to accurately identify liver cirrhosis in a cohort of veterans with untreated chronic HCV infection, using VCTE measurements as a gold standard.\[10\]

2. Subjects and methods

2.1. Design and patient selection

The study population consists of United States veterans with chronic HCV infection who underwent VCTE studies between September 1, 2015 and June 20, 2017 as part of their routine HCV clinical care at a single location, the Atlanta VA Medical Center (AVAMC). AVAMC is a level 1a tertiary care facility that provides comprehensive medical and surgical specialty services to more than 130,000 enrolled veterans living in 50 counties across northeast Georgia. All VCTE studies were performed by a single physician operator on the FibroScan 502 Touch machine (Echosens; Paris, France). VCTE study success rate was defined as valid measurements divided by valid plus invalid measurements, and is an established measure of VCTE study integrity.\[7\]

VCTE studies with a success rate less than 60% or studies that had less than 10 valid measurements were excluded. Additionally, any studies obtained when the VCTE probes were due for calibration were excluded. Lastly, for each VCTE study, variance is calculated as the interquartile range of stiffness divided by the median stiffness. Any VCTE studies with a variance ≥30% were excluded, consistent with published guidance.\[11\] All studies were conducted prior to HCV treatment. Only one VCTE study was included per person; in cases of duplicate studies, the study with a lower variance was selected. Liver Stiffness Measurement (LSM) less than or equal to 12.5 kPa were considered non-cirrhotic while those values greater than 12.5 kPa were considered to predict cirrhosis.\[12,13\] Probe size (either Medium or XL) was chosen by the VCTE operator using standard indicators.

For each VCTE study, a corresponding FIB-4 score was calculated using available laboratory values obtained up to 365 days preceding the VCTE study. Age, AST, ALT, platelet count, body mass index (BMI), and HCV RNA values were assessed through individual chart review using the Veterans Administration electronic health record, known as the Computerized Patient Record System (CPRS). FIB-4 scores were calculated as follows:

\[
\text{FIB-4} = \frac{\text{age (years)} \times \text{AST level (U/L)} / \text{platelet count (10^9/L) \times \sqrt{\text{ALT (U/L)}}}}{	ext{2}}
\]

All values in the FIB-4 calculation were obtained on the same calendar day and the laboratory testing date closest to the VCTE study date was selected. FIB-4 scores greater than 3.25 were considered to predict cirrhosis, while those less than or equal to 3.25 were considered non-cirrhotic. For the purpose of additional analyses, scores greater than 1.45 and less than 3.25 were considered “indeterminate.”

2.2. Sensitivity, specificity, positive predictive value, negative predictive value

A FIB-4 score of greater than 3.25 and a VCTE greater than 12.5 kPa was a true positive, while false positives were defined as FIB-4 scores greater than 3.25 but VCTE values less than or equal to 12.5 kPa. True negatives were defined as FIB-4 scores less than or equal to 3.25 and VCTE less than or equal to 12.5 kPa, while false negatives were defined as FIB-4 scores less than or equal to 3.25 and VCTE greater than 12.5 kPa.

Race, HIV status, and alcohol use disorder were ascertained through the Veteran Affairs Clinical Case Registry. An alcohol use disorder was defined by the presence of any ICD-9 codes corresponding to alcoholism in the patient’s CPRS record. The degree of concordance between FIB-4 score and VCTE study was assessed, along with sources of discordance in these 2 noninvasive staging modalities.

2.3. Statistical analysis

Data were stored in a secure research drive on the VA network and analyzed in Microsoft Excel 2016. MatLab was used to calculate normality using the Anderson-Darling test, and Open-Epi was used to calculate ANOVA and Welch t test when appropriate to determine significance with a 2 tailed P value of .05.\[15\] This study was approved by the Emory IRB, VA research and development, and has a waiver for the documentation of consent given its retrospective nature.

3. Results

Of 390 VCTE studies identified, 315 met the pre-determined inclusion criteria for quality (Fig. 1). Subsequently, 2 studies were excluded because they were performed during HCV treatment, 15 studies lacked complete laboratory data for FIB-4 calculation, and 28 studies were duplicates, which left 270 total studies. VCTE LSM, VCTE variance data, and FIB-4 scores were all found to be distributed normally. The average difference between VCTE study date and corresponding laboratory date was 84.4 days (range: 0–356). Ninety four percent of subjects were male (Table 1). The average age of our subjects was 63 years (range: 24–75). Of the 270 subjects, 205 (76%) were Black or African American (Table 1). Average AST, ALT, and platelet count were significantly different across VCTE subcategories and across FIB-4 subcategories (Table 2).

3.1. Liver fibrosis staging

Of the 270 VCTE studies, 184 (68.1%) were non-cirrhotic with LSM ranging from 4.1 to 12.5 kPa (average 8.2 kPa), and 86 (31.9%) predicted cirrhosis with LSM ranging from 12.8 to 75 kPa (average 24.3 kPa) (Table 2). The primary analysis, which used a binary FIB-4 score cutoff found that 212 (79%) persons had non-cirrhotic FIB-4 scores, while 58 (21%) had scores that predicted cirrhosis. A second analysis which identified indeterminate values found that 36 (13%) persons had non-cirrhotic FIB-4 scores, 176 (65%) persons had indeterminate FIB-4 scores, and 58 (22%) persons had cirrhotic FIB-4 scores (Table 2).

Those with non-cirrhotic VCTE values had an average FIB-4 score of 2.3, while those with cirrhosis by VCTE had an average FIB-4 score of 3.8 (P < .05). Conversely, in the primary analysis using a binary FIB-4 cutoff, those that were non-cirrhotic by FIB-4 had an average LSM of 10.8 kPa, while those that were cirrhotic by FIB-4 had an average LSM of 22.8 kPa (P < .05, Table 2).

A secondary analysis identified those with FIB-4 scores in an indeterminate range (n = 176); 48 (27.3%) persons with indeterminate FIB-4 scores had VCTE LSM that predicted cirrhosis (LSM range 12.8–72 kPA, average 19.25 kPa) and 128 (72.7%) persons with indeterminate FIB-4 scores had VCTE LSM that were non-cirrhotic (LSM range 4.2–12.5 kPa, average 8.23 kPa).
3.2. FIB-4 sensitivity and specificity

Of the 270 persons analyzed using a binary FIB-4 cutoff of 3.25 for liver cirrhosis, we found 36 true positives, 22 false positives, 162 true negatives, and 50 false negatives. Therefore, the sensitivity of the FIB-4 score in assessing liver cirrhosis was calculated at 42% and specificity was 88%. Additionally, we found the PPV of FIB-4 in predicting liver cirrhosis to be 62%, while the NPV was 76% (Table 3).

When FIB-4 scores in the indeterminate range (n = 176) were removed from the analysis, we found 36 true positives, 22 false positives, 34 true negatives, and 2 false negatives. Therefore, the sensitivity of the FIB-4 score in assessing liver cirrhosis increased to 95%, the specificity decreased to 61%, the PPV remained at 62%, and the NPV improved to 94% (Table 3).

3.3. Comorbidities

The average BMI across all subjects was 26.8 [range 16–40.6]. Most VCTE studies were done using the medium probe (n = 207, 77%) (Table 1). The average BMI value for studies completed with the medium probe was 25.4 [16–39], and 31.6 [23–40.6] for studies completed with the XL probe. Half of the subjects (n = 135) were identified as having an alcohol use disorder. Fourteen subjects (5%) were identified as having HIV co-infection (Table 1).

No significant difference in average LSM was found between subjects with a history of an alcohol use disorder and those without (14.6 kPa vs 12.2 kPa, P = .09 [Table 4]). Those with a BMI > 25 had a significantly higher average FIB-4 score compared with those that had a BMI ≤25 (3.2 vs 2.6, P = .04 [Table 4]). VCTE studies in those with HIV/HCV co-infection yielded a significantly lower average LSM when compared with measurements in those with HCV mono-infection (9.4 kPa vs 13.6 kPa, P < .01, [Table 4]).

4. Discussion

In accordance with the updated IDSA/AASLD HCV treatment guidelines, we used noninvasive liver markers to assess for probable liver cirrhosis in a US veteran population with chronic HCV infection. FIB-4 was compared with VCTE as the gold standard in ruling in and ruling out probable liver cirrhosis. Notably, while FIB-4 has traditionally been categorized as non-cirrhotic (values less than 1.45), indeterminate (values between 1.45–3.25), and cirrhotic (values greater than 3.25), current IDSA/AASLD guidelines make no mention of indeterminate values and use 3.25 as a binary cutoff to identify probable liver cirrhosis. We found that the majority of our FIB-4 scores were in the indeterminate range. When using a binary FIB-4 score in our analysis, we found that the sensitivity of FIB-4 was very
low at 42% and the PPV was only 62%. The sensitivity was driven down by the large number of false negatives (i.e., patients with FIB-4 scores less than 3.25 but VCTE values greater than 12.5 kPa). Therefore, we found that the use of FIB-4 alone to rule out liver cirrhosis was inadequate in our population, as evidenced by our high false negative rate. FIB-4 is relatively specific as few false positive studies were identified.

When excluding indeterminate FIB-4 scores from our analysis, we found a noticeable decrease in the number of false negative FIB-4 scores. Due to this, the sensitivity of FIB-4 in identifying probable liver cirrhosis improved greatly, along with an increase in NPV to 94.4%. Our results suggest that it is important to consider an indeterminate range when using FIB-4 scores to identify probable liver cirrhosis secondary to HCV infection, as opposed to a binary FIB-4 cutoff. In cases of indeterminate FIB-4 scores, a second staging method like VCTE can better inform the clinician whether probable liver cirrhosis is present.

Sources of variation in the FIB-4 score are well documented. In our analysis, the platelet count was significantly lower in subjects with FIB-4 scores in the cirrhotic range as compared to subjects with FIB-4 scores in the non-cirrhotic range (Table 2). While thrombocytopenia is expected in persons with liver cirrhosis, there are also extrahepatic conditions that could affect platelet counts and thus impact FIB-4 scores. Additionally, age is a significant component of FIB4, with some studies noting that this affects FIB-4 score interpretations in certain populations. For example, FIB-4 has been found to have a high false positive rate in elderly patients with Non-Alcoholic Fatty Liver Disease (NAFLD), which further limits its utility in ruling in liver cirrhosis in this patient demographic. Our average patient age in this study was 63 years, which means our results may not be applicable to younger populations.

The current literature has noted the diagnostic utility of VCTE and serologic markers in detecting significant liver fibrosis when used separately in HIV/HCV co-infected cohorts. Interestingly, in HIV/HCV co-infected populations, VCTE was found to be more reliable in detecting significant liver fibrosis when compared to serologic markers. To our knowledge, there is no current literature on the concordance of elastography and serologic markers when used to assess liver fibrosis in this co-infected population. In our study, HIV/HCV co-infected individuals were found to have significantly lower LSM as compared to LSM in HCV mono-infected patients. These unexpected results may be explained by the referral patterns seen in HIV/HCV co-infected individuals, as this population was and still is prioritized for Direct Acting Antiviral treatment. Additionally, the small number of HIV/HCV co-infected patients in our cohort was also a limiting factor and may have contributed to our unexpected findings.

The strengths of this study include the large number of HCV-infected subjects who underwent VCTE studies by a single physician operator. Additionally, the homogeneity of our population (mostly older males) may have allowed for better performance of the FIB-4 score. Laboratory data matching VCTE study date were easily accessible in CPRS, and therefore laboratory data quality and non-liver related conditions that could affect laboratory data were reviewed. While these results are promising, our study had limitations. These included the lack of self-reported alcohol use at the time of noninvasive staging regardless of ICD code designation, the retrospective methodology of this study, as well as the mostly male population.

The traditional gold standard to staging liver disease has been the liver biopsy. However, the management and treatment of patients with liver disease is increasingly reliant on noninvasive

### Table 2

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sample Count n (%)</th>
<th>Avg. LSM (kPa)</th>
<th>Avg. AST (IU/L)</th>
<th>Avg. ALT (IU/L)</th>
<th>Avg. Platelets (k/cm²)</th>
<th>Avg. Age (years)</th>
<th>Avg. BMI</th>
<th>Alternate Liver Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSM by VCTE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No cirrhosis (X&lt;12.5 kPa)</td>
<td>184 (68.1)</td>
<td>8.2 [4.1–12.5]</td>
<td>52</td>
<td>78</td>
<td>210</td>
<td>62</td>
<td>26.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Cirrhosis (X &gt;12.5 kPa)</td>
<td>86 (31.9)</td>
<td>24.3 [12.8–75]</td>
<td>54</td>
<td>78</td>
<td>174</td>
<td>63</td>
<td>27.2</td>
<td>3.8</td>
</tr>
<tr>
<td>FIB-4 Binary Cutoff</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No cirrhosis (X&lt;3.25)</td>
<td>212 (78.5)</td>
<td>2.1 [0.7–3.2]</td>
<td>50</td>
<td>56</td>
<td>214</td>
<td>62</td>
<td>27</td>
<td>10.8</td>
</tr>
<tr>
<td>Cirrhosis (X&gt;3.25)</td>
<td>58 (21.5)</td>
<td>5.3 [3.3–21.5]</td>
<td>97</td>
<td>83</td>
<td>140</td>
<td>63</td>
<td>26.4</td>
<td>22.8</td>
</tr>
<tr>
<td>Secondary Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indeterminate (1.45&lt;X&lt;3.25)</td>
<td>176 (65.2)</td>
<td>2.3 [1.5–3.2]</td>
<td>53</td>
<td>58</td>
<td>201</td>
<td>63</td>
<td>26.6</td>
<td>11.2</td>
</tr>
</tbody>
</table>

Average Liver Stiffness Measurement (LSM) across VCTE subcategories of “No cirrhosis” and “Cirrhosis,” as well as corresponding average AST, ALT, Platelets, Age, BMI, and FIB-4 Scores. Average FIB-4 Scores across FIB-4 subcategories of “No cirrhosis” and “Cirrhosis,” as well as corresponding average AST, ALT, Platelets, Age, BMI, and LSM values.

*Significantly different average AST, ALT, platelet, and FIB-4 values (P<.05) between "No cirrhosis" and "Cirrhosis" LSM subcategories by VCTE (Independent t-test). **Significantly different (P<.05) average AST, ALT, platelets, and LSM values between “No cirrhosis” and “Cirrhosis” FIB-4 subcategories (Independent t-test).

### Table 3

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Analysis: FIB-4 &gt;3.25 and VCTE &gt;12.5</td>
<td>41.9%</td>
<td>88%</td>
<td>62%</td>
<td>76.4%</td>
</tr>
<tr>
<td>Secondary Analysis: Indeterminate Values Removed</td>
<td>94.7%</td>
<td>60.7%</td>
<td>62.1%</td>
<td>94.4%</td>
</tr>
</tbody>
</table>

Sensitivity, specificity, positive predictive value, and negative predictive value of FIB-4 scores in identifying probable liver cirrhosis with VCTE as the gold standard. Separate analyses shown with indeterminate FIB-4 values included (primary) and excluded (secondary).
staging modalities. In our study, the FIB-4 score demonstrated good specificity with few false positive results noted. Therefore, the FIB-4 score performed well in ruling in probable liver cirrhosis, which could prevent the need for a liver biopsy or other unnecessary evaluation in patients with chronic HCV infection. However, caution is needed when using a binary FIB-4 score cutoff to screen for liver cirrhosis, as we found many false negatives in our analysis. Falsely negative FIB-4 scores could translate into harmful clinical outcomes. Namely, cases of probable liver cirrhosis would go undetected and opportunities for HCC surveillance would be missed. In addition, falsely negative FIB-4 scores may prompt clinicians to choose inappropriately short antiviral treatment durations. For those persons with a FIB-4 score in the indeterminate range, we recommend performing a VCTE study to further evaluate for the presence of liver cirrhosis. Our results support the expert recommendations by Tapper et al of using a combined approach of elastography and serologic testing, such as the FIB-4 index, to discriminate between patients at high risk for cirrhosis versus those at low risk. In cases where the results of this combined approach are discordant, liver biopsy can be considered.\(^{110}\) For patients with discordant elastography and serologic testing, the need for liver biopsy is substantially reduced.

**Table 4**

Comparison of the average FIB-4 score and liver stiffness measurement by HIV status, Body Mass Index (BMI) and alcohol use disorder (n = 270).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Average FIB-4 score</th>
<th>Average LSM (kPa)</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV positive (n = 14)</td>
<td>3.09</td>
<td>9.4</td>
<td>.007*</td>
</tr>
<tr>
<td>HIV negative (n = 256)</td>
<td>2.76</td>
<td>13.6</td>
<td></td>
</tr>
<tr>
<td>BMI ≤ 25 (n = 91)</td>
<td>3.22</td>
<td>12.2</td>
<td>.040*</td>
</tr>
<tr>
<td>BMI &gt; 25 (n = 179)</td>
<td>2.55</td>
<td>13.9</td>
<td></td>
</tr>
<tr>
<td>Alcohol use disorder present</td>
<td>3.06</td>
<td>14.6</td>
<td>No statistical significance</td>
</tr>
<tr>
<td>Alcohol use disorder absent</td>
<td>2.50</td>
<td>12.2</td>
<td></td>
</tr>
</tbody>
</table>

Average FIB-4 score and average LSM values as a function of HIV status, BMI, and alcohol use disorder status. 

*Significant differences \((P < .05)\) independent \(t\) test.

**Author contributions**

Conceptualization: Nabil Saleem, Emily Jeanne Cartwright.  
Data curation: Nabil Saleem, Emily Jeanne Cartwright.  
Formal analysis: Nabil Saleem, Emily Jeanne Cartwright.  
Investigation: Nabil Saleem, Emily Jeanne Cartwright.  
Methodology: Nabil Saleem, Emily Jeanne Cartwright.  
Project administration: Emily Jeanne Cartwright.  
Supervision: Emily Jeanne Cartwright.  
Validation: Emily Jeanne Cartwright.  
Visualization: Emily Jeanne Cartwright.  

**References**