Noninvasive Serum Fibrosis Markers for Screening and Staging Chronic Hepatitis C Virus Patients in a Large US Cohort

Scott D. Holmberg, Mei Lu, Loralee B. Rupp, Lois E. Lamerato, Anne C. Moorman, Vinutha Vijayadeva, Joseph A. Boscarino, Emily M. Henkle, and Stuart C. Gordon; for the Chronic Hepatitis Cohort Study (CHeCS) Investigators

1Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, Georgia; 2Data Coordinating Center and 3Center for Health Services Research, Henry Ford Health System, Detroit, Michigan; 4Kaiser Permanente Center for Health Research, Honolulu, Hawaii; 5Geisinger Health System, Danville, Pennsylvania; and 6Kaiser Permanente Center for Health Research–Northwest, Portland, Oregon

Background. Liver biopsy remains critical for staging liver disease in hepatitis C virus (HCV)–infected persons, but is a bottleneck to evaluation, follow-up, and treatment of HCV. Our analysis sought to validate APRI (aspartate aminotransferase [AST]–to-platelet ratio index) and FIB-4, an index from serum fibrosis markers (alanine aminotransferase [ALT], AST, and platelets plus patient age) to stage liver disease.

Methods. Biopsy results from HCV patients in the Chronic Hepatitis Cohort Study were mapped to an F0–F4 equivalent scale; APRI and FIB-4 scores at the time of biopsy were then mapped to the same scale.

Results. We identified 2372 liver biopsies from HCV-infected patients with contemporaneous laboratory values for imputing APRI and FIB-4. Fibrosis stage distributions by the equivalent biopsy scale were 267 (11%) F0; 555 (23%) F1; 648 (17%) F2; 394 (17%) F3; and 508 (21%) F4. Mean APRI and FIB-4 values significantly increased with successive fibrosis levels (P < .05). The areas under the receiver operating characteristic curve (AUROC) analysis distinguishing severe (F3–F4) from mild-to-moderate fibrosis (F0–F2) were 0.80 (95% confidence interval [CI], .78–.82) for APRI and 0.83 (95% CI, .81–.85) for FIB-4. There was a significant difference between the AUROCs of FIB-4 and APRI (P < .001); 88% of persons who had a FIB-4 score ≥2.0 were at stage F2 or higher.

Conclusions. In a large observational cohort, FIB-4 was good at differentiating 5 stages of chronic HCV infection. It can be useful in screening patients who need biopsy and therapy, for monitoring patients with less advanced disease, and for longitudinal studies.

Keywords. hepatitis C virus; chronic hepatitis; clinical staging.

Staging HCV infection is still mainly based on degree of histologic fibrosis in a liver biopsy sample, but there are many problems in relying on biopsy. Although percutaneous liver biopsy is usually a safe procedure, it is costly and does carry a small risk for complication [1]. There can easily be sampling errors, because only approximately 0.002% of the organ is biopsied, and inter- and intraobserver discrepancies of 10%–20% in assessing hepatic fibrosis have been reported [2, 3]. In addition, liver biopsy is performed or arranged for by a small number of specialists, creating a “bottleneck” in staging and treating patients infected with hepatitis C virus (HCV). The procedure is uncomfortable, if not painful, and some patients refuse the procedure and, consequently, evaluation for therapy. Further, as biopsy is usually performed once on a patient, the ability to monitor a patient’s liver fibrosis would benefit from an index based on serum fibrosis markers comparable to determining CD4+ cell counts as used for evaluating and monitoring patients with human immunodeficiency virus (HIV).
Thus, several indices constructed from noninvasive serum-based biomarkers of fibrosis—here called “serum fibrosis markers”—have been proposed and validated, usually within relatively small sets of treatment-naive patients with chronic hepatitis C [4]. Most attention has centered on the aspartate aminotransferase (AST)–to–platelet ratio index (APRI) [5–8] and the FIB-4 index [9–11], which is calculated from AST, alanine aminotransferase (ALT), platelet count, and patient age. More complicated indices using harder-to-obtain laboratory values [12–14] with or without transient elastography [15, 16] have also been proposed. However, APRI and FIB-4 have been of more interest to clinicians because they are simple to calculate and readily available from hospital or clinic laboratories during usual patient care. That is, these simple calculations based on serum result would be useful to screen patients with high values needing biopsy and clinical follow-up and to provide a system for categorizing stage of illness. It is critical to determine which HCV patients have advanced fibrosis to gauge the urgency of treatment as well as the need for upper endoscopy for varices, biannual ultrasounds for hepatocellular cancer screening, and closer clinical monitoring of cirrhotic patients.

The Chronic Hepatitis Cohort Study (CHeCS), a prospective, longitudinal, observational cohort study, was established to assess the clinical impact of chronic viral hepatitis B and C in the United States [17, 18]. CHeCS is a “dynamic” multicenter cohort study conducted at 4 large, integrated healthcare systems located in Detroit, Michigan; Danville, Pennsylvania; Portland, Oregon; and Honolulu, Hawaii, and represents a geographically, ethnically, and clinically diverse US-based cohort of, currently, about 3000 hepatitis B virus–infected and 12 000 HCV-infected patients. Because CHeCS is an observational study, the data collected from the electronic medical record are solely based on routine clinical care and thus representative of the uncontrolled healthcare environment of the “real world” clinical setting. The laboratory tests necessary for imputation of the serum fibrosis markers were not necessarily collected on the same date as the liver biopsy (but close in time). The goal of this analysis was to evaluate the capability of serum fibrosis markers, imputed from labs collected during the course of routine care and within 6 months of a biopsy, to accurately predict fibrosis level as interpreted by pathologists reading biopsies in an uncontrolled, real-world setting.

**METHODS**

The Chronic Hepatitis Cohort Study (CHeCS)
The patients included in this study are the chronic hepatitis C subpopulation of the CHeCS cohort, the recruitment and baseline characteristics of which have been described elsewhere [18]. In brief, the analysis included adults aged ≥18 years from the 4 participating healthcare organizations (Geisinger Health System, Danville, Pennsylvania; Henry Ford Health System, Detroit, Michigan; Kaiser Permanente Northwest, Portland, Oregon; Kaiser Permanente, Honolulu, Hawaii) with at least 1 admission or outpatient provider, laboratory, or emergency department encounter from 1 January 2006 through 31 December 2010.

The study underwent ethical review and was approved by the institutional review boards at each study site and the Centers for Disease Control and Prevention (Atlanta, Georgia). Trained medical abstractors conducted the chart reviews to confirm chronic infection status, as well as to collect biopsy results. The study was restricted to confirmed chronic hepatitis C patients who had the requisite serum fibrosis markers and biopsy fibrosis readings within 6 months of each other.

**Data Collection and Classification**
Patient data were collected and analyzed from electronic medical records including age (at time of liver biopsy); sex; race/ethnicity; annual income (derived from census tract data based on zip code or patient residence); serum ALT level and AST levels (elevated values were relative to the upper limit of normal value specific to each laboratory that performed the test); and platelet counts. The laboratory data were largely collected via electronic medical records; in addition, lab values from external laboratories were captured through the chart abstraction. It was not a requirement that all of the component lab values necessary for imputing the serum fibrosis markers be collected on the same day as each other; the serum fibrosis markers were imputed based on labs collected up to within 7 days of each other. Lab tests after liver transplantation were excluded from this analysis.

**Histologic Liver Assessment**
Liver biopsy samples were fixed in formalin and embedded in paraffin and were evaluated by pathologists for determination of fibrosis status. Fibrosis scores from different scoring systems (International Association for the Study of the Liver, Batts Ludwing, Metavir, Ishak, Knodell, Scheuer) were mapped to a F0–F4 equivalency scale. That is, fibrosis was ranked as follows: F0, no fibrosis; F1, portal fibrosis without septa; F2, portal fibrosis with few septa; F3, numerous septa without cirrhosis; and F4, cirrhosis. If patient had a liver transplant, the laboratory results and biopsies after the transplant were excluded from this analysis. If the patient had >1 biopsy, the most severe biopsy at the earliest date with available lab results was used for this analysis.

**Indices Bases on Serum Fibrosis Markers**
All patients’ laboratory data (ALT, AST, platelet count) were collected through electronic medical records. If multiple laboratory values were available, the results closest to the time of
biopsy were used. APRI, FIB-4, and, for purposes of comparison, AST/ALT ratio were calculated when the laboratory assessments were within 7 days of each other and within 6 months of the biopsy.

\[
\text{AST/ALT} = \frac{\text{AST}}{\text{ALT}}
\]

\[
\text{APRI} = \frac{\text{AST level} (\text{ULN}^+)}{\text{Platelet count} (10^9/L)} \times 100
\]

\[
\text{FIB} - 4 = \frac{\text{Age (years)} \times \text{AST} (U/L)}{\text{Platelet count} (10^9/L) \times [\text{ALT} (U/L)]^{1/2}}
\]

### Statistical Methods

Each serum fibrosis marker was evaluated for normality, but as they were not normally distributed, log transformation was used for the analysis. To determine the association of each index based on serum fibrosis markers with biopsy fibrosis staging, generalized estimating equation was used instead of the analysis of variance, as it provides a robust estimation with less restriction on the underlying distribution of data. The analysis tested for the mean differences among 5 biopsy fibrosis stages (the overall group effect), followed by pairwise comparisons between fibrosis levels if an overall group effect was detected at \( P < .05 \). The mean and its 95% confidence interval (CI) were calculated for each biopsy group.

In addition, the predictive ability of each index from serum fibrosis markers to differentiate advanced fibrosis (F3 and F4) from mild-to-moderate fibrosis (F0 through F2) was measured by the area under the receiver operating characteristic curve (AUROC) analysis. Nonparametric Mann-Whitney U test was used to assess statistically significant differences of AUROC between the 3 indices based on serum fibrosis markers. For each marker, the optimal cutoff point was identified to minimize misclassification with calculation of sensitivity and specificity. The same analysis was repeated using the indices when the laboratory assessments were within 7 days of each other and within 3 months of the biopsy.

### Results

A total of 10,473 patients had confirmed chronic HCV and, after excluding 41 patients with only biopsy after liver transplant, 4313 (41%) had unique fibrosis staging by liver biopsy. Of them, 2372 (55%) had calculable APRI, FIB-4, and AST/ALT scores within 6 months of the biopsy date. These patients were a mean age of 50 years at the time of biopsy and, like all hepatitis C patients in the CHeCS [18], were more likely to be male (61%), white (65%), and, for those in health plans, enrolled a mean of 7.4 years (88.7 months; Table 1).

### Overall Correlation of APRI and FIB-4 With Successive Stages of Liver Fibrosis

The fibrosis stage distributions by the equivalent scale were 267 (11%) F0; 555 (23%) F1; 648 (27%) F2; 394 (17%) F3; and 508 (21%) F4 (Table 2). Biomarker values were significantly associated with overall fibrosis stage levels \( P < .01 \). The mean and its 95% CIs were mutually exclusive of each other, indicating significant mean differences among biopsy fibrosis levels \( P < .05 \); Table 2).

The AST/ALT ratios, used by some clinicians and in other analyses [19], were calculated for purposes of comparison (Figure 1), but clearly performed less well than either APRI or FIB-4 (Figure 1).

### Ability of Indices From Serum Fibrosis Markers for Predicting Severe Fibrosis in Liver Histology (F3–F4)

The AUROCs in distinguishing severe fibrosis (F3–F4) from mild-to-moderate fibrosis (F0–F2) were 0.80 (95% CI, .78–.82)
for APRI, 0.83 (95% CI, .81–.85) for FIB-4, and 0.64 (95% CI, .61–.66) for AST/ALT ratio (P < .001; Figure 1). There was a significant difference between the AUROCs of FIB-4 and APRI, and between APRI and AST/ALT ratio (Figure 1). The optimal cutoff point for APRI was 0.81 (sensitivity 75%, specificity 74%), 1.81 for FIB-4 (sensitivity 74%, specificity 77%), and 0.82 for ALT/AST ratio (sensitivity 62%, specificity 60%). Of 981 patients with FIB-4 score ≥ 2.0, 862 (87.9%) had a biopsy reading of F2 or higher. Restricting to laboratory values obtained within 3 months of the biopsy, AUROCs in distinguishing severe fibrosis (F3–F4) from mild-to-moderate fibrosis were 0.81 (95% CI, .79–.83) for APRI, 0.84 (95% CI, .82–.86) for FIB-4, and 0.66 (95% CI, .63–.68) for AST/ALT ratio.

**DISCUSSION**

In a large observational real-world cohort of chronic hepatitis C patients, FIB-4 was superior to APRI and much superior to a simple AST/ALT ratio at distinguishing severe fibrosis from mild-to-moderate fibrosis. Both FIB-4 and APRI had excellent predictive ability when the serum fibrosis marker(s) could be collected up to within 6 months of the biopsy. FIB-4 scores were strongly associated with patient status within 5 stages of HCV infection determined by biopsy. To our knowledge, this is the largest such analysis of these serum fibrosis marker scores as derived from a US population of chronic hepatitis C patients.

There are several reasons why using FIB-4 would be helpful in guiding patient monitoring and care. Current guidelines for antiviral treatment for HCV recommend, among other things, liver biopsy confirmation of substantial fibrosis or cirrhosis [20, 21]. In limited studies to date, high FIB-4 scores (eg, ≥ 2.25) appear to discriminate between these severe stages (F3–F4) and low or moderate stages (F0–F2) of fibrosis [13, 22]. Use of FIB-4 may

---

**Table 2. Correlation of Hepatitis C Virus Disease Stage by Invasive (Liver Biopsy Staging) and Noninvasive (APRI and FIB-4 Scores), Chronic Hepatitis Cohort Study—2372 Biopsies, 2008–2011**

<table>
<thead>
<tr>
<th>Degree of Fibrosis (Stage)</th>
<th>Liver Biopsy Scoring System</th>
<th>Noninvasive Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IASL</td>
<td>Metavir</td>
</tr>
<tr>
<td>No fibrosis (F0) (n = 267)</td>
<td>No fibrosis (0)</td>
<td>F0</td>
</tr>
<tr>
<td>Fibrous portal expansion (F1) (n = 555)</td>
<td>Mild-portal fibrosis (1)</td>
<td>F1</td>
</tr>
<tr>
<td>Few bridges or septa (F2) (n = 648)</td>
<td>Moderate fibrosis (2)</td>
<td>F2</td>
</tr>
<tr>
<td>Numerous bridges or septa (F3) (n = 394)</td>
<td>Severe fibrosis (3)</td>
<td>F3</td>
</tr>
<tr>
<td>Cirrhosis (F4) (n = 508)</td>
<td>Cirrhosis (4)</td>
<td>F4</td>
</tr>
</tbody>
</table>

Abbreviations: APRI, aspartate aminotransferase–to-platelet ratio index; CI, confidence interval; FIB-4, a fibrosis index that combines 3 standard biochemical values (platelets, alanine aminotransferase, aspartate aminotransferase) plus patient age; IASL, International Association for the Study of the Liver.

* Score differences between stages, P < .05.
obviate the need for liver biopsy for uncomplicated earlier-stage HCV patients. Further, determining which of the 30%–40% of hepatitis C patients will progress to cirrhosis end-stage liver disease, hepatocellular carcinoma, and death has been problematic [23]: a noninvasive serum fibrosis marker score would avert this difficulty in monitoring patients’ disease progression.

Therapeutic decisions about when to start antiviral therapy or not are not the only reason that clinicians may want a noninvasive way to monitor and assess liver disease. It is critical to determine which HCV patients have advanced fibrosis to gauge the need for upper endoscopy for varices, biannual ultrasounds for hepatocellular carcinoma, and close clinical monitoring of cirrhotic patients.

There are other advantages to using FIB-4 or other serum fibrosis marker indices to initially stage and follow HCV patients. First, liver biopsy is usually performed or arranged for (to be done by radiologist) by a liver specialist, requiring the patient to seek care from such a specialist. As there are >3 million HCV-infected patients in the United States, but <2000 board-certified hepatologists, there is a scarcity of clinicians qualified to diagnose, follow, and treat HCV patients. Although liver biopsy is not required for treatment, in the CHeCS HCV-infected population, 38.4% had had a biopsy between 2001 and 2010 [18]. Requiring biopsy to justify antiviral therapy creates a bottleneck that may lead to many HCV-infected patients not seeking or receiving care, as in this population [17,18]. There is growing interest and attention from the perspective of healthcare advocates and hepatologists that hepatitis C care can and should be provided by internists, infectious disease specialists, family practitioners, and other clinicians [24]. Ease of monitoring would be especially helpful in systems such as Project ECHO in New Mexico, which has demonstrated the utility and effectiveness of guiding non-specialist clinicians by teleconference and other telecommunication in caring for HCV patients in remote, rural, or hard-to-access areas [25]. Still, even if non-specialists can manage uncomplicated HCV infection, it is important to note that management of late-stage, cirrhotic patients, especially those who may decompensate with antiviral therapy, should continue to be managed by hepatologists and others with experience in treating such patients.

Studies of the natural history, timing, and success of treatment of chronic HCV have been hampered by a lack of a relatively easy noninvasive staging system, such as CD4 cell count and viral load as used for HIV. Clinically, it is hard to monitor the progress of an individual patient without performing multiple biopsies. Thus, another advantage of using FIB-4 will be to allow longitudinal studies of the natural history of HCV and risk of and preventive factors for liver disease progression. Because liver biopsies are usually performed only once on a patient, understanding of the progression of HCV infection has been limited to studies of the few patients who have multiple biopsies [26] or by meta-analysis of several small studies [22,27]. Longitudinal analysis of the effects of antiviral drug therapy, alcohol use (or cessation), and other factors that may impact HCV disease progression is important, but requires a way of monitoring progression similar to that seen with HIV (CD4+ cell levels).

Transient elastography (FibroScan) may soon be approved for use in the United States, and this technology appears to be superior to FIB-4 or other serum fibrosis marker calculations for later-stage (F3–F4) hepatic fibrosis and cirrhosis, but also equally or less useful in the diagnosis of low-to-moderate liver fibrosis [4,28]. Besides its expense, the applicability (80%) of elastography is not as good as that for serum fibrosis markers, and unreliable results—that is, not meeting manufacturer’s recommendations—have been reported for 16% of tests [29]. Problems are caused by patient obesity, limited operator experience, or if a patient has eaten a meal within the previous 3 hours [4,29,30]. In any case, serum fibrosis markers will for the near future remain more readily available, reliable, and less expensive to the widening group of physicians who are treating chronic hepatitis C.

Limitations to this analysis include variability in these serum fibrosis markers at various stages of liver disease (fibrosis). In terms of assessing liver disease severity, it has not been demonstrated that assessment of structure (biopsy) is more reliable than indices derived from liver injury (ALT, AST) and hematologic (platelet) tests. However, even assuming that liver histology is the gold standard, it is subject to inter- and intraobserver discrepancies of 10%–20% in those reading biopsy specimens [2,3]. Thus, we did not—and could not—rely on central reading of >2000 biopsies at the 4 sites; we wanted to investigate performance of noninvasive serum fibrosis markers and biopsy as performed in a wide range of real-world settings and situations. Nonetheless, biopsies were somewhat overrepresented in men and white persons compared to HCV prevalence in these groups in the general population [31], and so these factors must be considered when generalizing from these data.

Although assigning multiple fibrosis staging systems to a single category (F0–F4) may result in misclassification, presumably equally numbers of specimens were incorrectly categorized to a higher or lower stage. However, such variability may have limited clinical applicability. Based on our analysis, a FIB-4 score of 1.81 provides the best sensitivity and specificity for distinguishing stages F3 and F4 from lesser stages of liver fibrosis. As a simpler guide, a threshold FIB-4 score of 2.0 or greater would identify 88% of those at F2 or higher stage of liver fibrosis, who are appropriate for further evaluation, including biopsy, and treatment.

In summary, this analysis suggests that use of FIB-4 will facilitate screening, identification, and treatment of HCV patients
needing liver biopsy and antiviral therapy, be accessible to non-hepatologist clinicians who do or wish to care for patients with chronic HCV infection, and provide a reasonable staging system for the analysis of HCV infection and the factors that accelerate (eg, alcohol use) and stop or retard (eg, antiviral therapy) disease progression. Accordingly, the CHeCS Investigators are currently analyzing several outcomes—such as mortality, hospitalization, and efficacy of antiviral drug therapy—stratified by patients’ FIB-4 levels.

Notes

Financial support. CHeCS is funded by the CDC Foundation, which received unrestricted grants from Abbott Laboratories; Genentech, a member of the Roche Group; Janssen Pharmaceutical Companies of Johnson & Johnson; and Vertex Pharmaceuticals.

Potential conflicts of interest. S. C. G. receives grant/research support from Abbott Pharmaceuticals, Bristol-Myers Squibb, Exalenz BioScience, Gilead Pharmaceuticals, GlaxoSmithKline, GlobalImmune, Intercept Pharmaceuticals, Merck, Roche Pharmaceuticals, Tibotec, Vertex Pharmaceuticals, and Zymogenetics; serves as a consultant for Achillion, Bristol-Myers Squibb, CVS Caremark, Gilead Pharmaceuticals, Merck, Salix Pharmaceuticals, Johnson & Johnson, and Vertex; and serves on the data monitoring board for Tibotec. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

APPENDIX

The CHeCS Investigators include the following investigators and sites: Scott D. Holmberg, Eyasu H. Teshale, Philip R. Spradling, and Anne C. Moorman, Division of Viral Hepatitis, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia; Stuart C. Gordon, David R. Nerenz, Mei Lu, Lois Lamerato, Loralee B. Rupp, Nonna Akkerman, Nancy J. Oja-Tebbe, Chad M. Cogan, and Dana Larkin, Henry Ford Health System, Detroit, Michigan; Joseph A. Boscarino, Zahra S. Daar, Robert E. Smith, Patrick J. Curry, Brandon D. Geise, and Joe B. Leader; Geisinger Health System, Danville, Pennsylvania; Cynthia C. Nakasato, Vinutha Vijayadeva, Kelly E. Sylva, John V. Parker, and Mark M. Schmidt, Kaiser Permanente Hawaii, Honolulu, Hawaii; Emily M. Henkle, Tracy L. Dodge, Erin M. Keast, and Lois Drew, Kaiser Permanente Northwest, Portland, Oregon.