Research Article

JOURNAL OF HEPATOLOGY

Viral Hepatitis

Evaluation of APRI and FIB-4 scoring systems for non-invasive assessment of hepatic fibrosis in chronic hepatitis B patients☆

W. Ray Kim[1,*], Thomas Berg[2], Tarik Asselah[3], Robert Flisiak[4], Scott Fung[5], Stuart C. Gordon[6], Harry L.A. Janssen[5], Pietro Lampertico[7], Daryl Lau[8], Jeffrey D. Bornstein[9], Raul E. Aguilar Schall[9], Phillip Dinh[9], Leland J. Yee[9], Eduardo B. Martins[9], Seng Gee Lim[10], Rohit Loomba[11], Jörg Petersen[12], Maria Buti[13], Patrick Marcellin[3]

[1]*Stanford University School of Medicine, Stanford, CA, USA;* [2]*Universitätsklinikum Leipzig, Leipzig, Germany;* [3]*Hôpital Beaujon, Paris, France;* [4]*Medical University of Bialystok, Bialystok, Poland;* [5]*University of Toronto, Toronto, Ontario, Canada;* [6]*Henry Ford Hospital, Detroit, MI, USA;* [7]*University of Milan, Milan, Italy;* [8]*Harvard University, Boston, MA, USA;* [9]*Gilead Sciences, Inc., Foster City, CA, USA;* [10]*Yoo Loo Lin School of Medicine, Singapore, Singapore;* [11]*University of California San Diego, San Diego, CA, USA;* [12]*IFI Institute for Interdisciplinary Medicine, Asklepios Klinik St. Georg, Hamburg, Germany;* [13]*Hospital General Universitari Vall d'Hebron, Barcelona, Spain*

Background & Aims: While the gold standard in the assessment of liver fibrosis remains liver biopsy, non-invasive methods have been increasingly used for chronic hepatitis B (CHB). This study aimed to evaluate the performance of two commonly used non-invasive scoring systems (aspartate aminotransferase-to-platelet ratio index (APRI) and fibrosis index based on four factors (FIB-4)) to predict fibrosis stage in CHB patients.

Methods: Demographic, histologic and clinical laboratory data from two trials investigating tenofovir disoproxil fumarate in CHB were analyzed. Predicted fibrosis stage, based on established scales and cut-off values for APRI and FIB-4 scores, was compared with Ishak scores obtained from liver biopsy at baseline and at 240 week follow-up.

Results: In the 575 patients with a baseline liver biopsy, APRI and FIB-4 scores correlated with Ishak stage (p <0.01); however extensive overlap in the distribution of both scores across Ishak stages prevented accurate determination of fibrosis. The majority (81–89%) of patients with advanced fibrosis or cirrhosis were missed by the scores. Similarly, 71% patients without fibrosis were misclassified as having clinically significant fibrosis. APRI and FIB-4 scores at week 240 tended to be low and underestimate fibrosis stage in the patients with liver biopsies after 240 weeks of therapy. APRI or FIB-4 reduction did not correlate with fibrosis regression after 240 weeks of antiviral therapy.

Conclusions: APRI and FIB-4 scores are not suitable for use in clinical practice in CHB patients for assessment of hepatic fibrosis according to Ishak stage, especially in gauging improvements in liver fibrosis following therapy.

Introduction

Histologic assessment of the liver to determine the degree of fibrosis is essential for evaluating patients with chronic liver disease. For patients with chronic hepatitis B (CHB), most guidelines recommend histologic assessment to determine treatment candidacy, in particular when hepatitis B virus (HBV) DNA and alanine aminotransferase (ALT) levels are near the threshold for starting therapy [1–3]. Robust data from large-scale clinical trials indicate that long-term treatment with potent, effective anti-HBV therapy can lead to significant regression of fibrosis [4,5] and reversal of cirrhosis in a substantial proportion of treated patients [5].

Regression of hepatic fibrosis is one of the new frontiers in hepatology. Clinically, follow-up evaluation of fibrosis has a significant role in prognosticating patients with varying degrees of fibrosis, consideration of treatment discontinuation, and determination of candidacy for surveillance for hepatocellular carcinoma (HCC).

Percutaneous liver biopsy has been the gold standard for assessment of fibrosis; however, limitations of this procedure include cost, risk of serious complications, sampling errors and inter- and intra-observer variations [6–9]. Recently, indirect assessments of liver fibrosis by 'non-invasive' means have been developed. In addition to physical methods, scoring systems based on laboratory tests could be an alternative method for

Keywords: APRI; Chronic hepatitis B; FIB-4; Fibrosis.
Received 21 May 2015; received in revised form 20 October 2015; accepted 17 November 2015; available online 2 December 2015
☆ Guest editor: Didier Samuel
* Corresponding author. Address: 300 Pasteur Drive, Stanford University, Stanford, CA 94305, USA. Tel.: +1 650 725 6511.
E-mail address: wrkim@stanford.edu (W.R. Kim).
Abbreviations: CHB, chronic hepatitis B; HBV, hepatitis B virus; ALT, alanine aminotransferase; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; AST, aspartate aminotransferase; APRI, AST-to-platelet ratio index; FIB-4, fibrosis index based on four factors; HIV, human immunodeficiency virus; ROC, receiver operating characteristics; TDF, tenofovir disoproxil fumarate; HBeAg, hepatitis B 'e' antigen; ULN, upper limit of normal; PPV, positive predictive value; LOX2, lysyl oxidase homolog 2; HBsAg, hepatitis B surface antigen.

Research Article

determining the extent of hepatic fibrosis. Two commonly used scoring systems for chronic hepatitis C virus (HCV) are the 'aspartate aminotransferase (AST)-to-platelet ratio index' (APRI) [10] and the 'fibrosis index based on four factors' (FIB-4), originally developed in patients co-infected with HCV and human immunodeficiency virus (HIV), taking account of AST, ALT, platelet count and patient age [11]. These scoring systems have successfully predicted hepatic fibrosis in large cohorts of patients infected with HCV [12,13]. A number of studies have also described the test performance characteristics of APRI and FIB-4 scores in patients with CHB [14–16]. These studies suggest that APRI and FIB-4 are suitable markers for detecting liver fibrosis stage with a moderate sensitivity and accuracy. Currently, given that the components of APRI and FIB-4 scores are readily available, they have been used in clinical practice as well as in epidemiological research. For example, the World Health Organization (WHO) CHB guidelines recommend APRI to determine fibrosis stage in resource-limited countries [3].

Published studies about APRI and FIB-4 scores have focused on metrics such as sensitivity, specificity and the area under the receiver operating characteristics (ROC), which address the scores' ability to discriminate between patients with early *vs.* advanced fibrosis. However, for the scores to be useful, the assessment must go beyond these parameters and include predictive values and calibration (ability for the scores to correctly predict fibrosis stage). The main utility of these scores in practice would include identification of candidates for antiviral therapy and surveillance for HCC. High positive and negative predictive values would be essential to avoid missed opportunities to improve long-term outcomes. Lastly and importantly, no information is available about applying these scores to evaluate regression of fibrosis as a result of long-term antiviral suppression of HBV.

The current analysis evaluated the APRI and FIB-4 scoring systems in a large cohort of well-characterized CHB patients, by assessing the association between the predicted hepatic fibrosis stage and the histology results from liver biopsies prior to treatment initiation. In addition, we evaluated the utility of the scores in follow-up assessment of fibrosis regression or progression following long-term therapy.

Materials and methods

Study design and patients

This analysis used data obtained in two phase 3 double-blind clinical trials of tenofovir disoproxil fumarate (TDF) for CHB (NCT00117676 and NCT00116805, also referred to as GS-US-174-0102 and GS-US-174-0103, respectively) [17]. Demographic, histologic, and clinical laboratory data were extracted at baseline and at follow-up. The unique feature of the trials that allowed undertaking this analysis included prolonged duration of the trial with a planned on-treatment follow-up of 384 weeks and available serial histologic data at baseline, week 48, and week 240.

The details of the trials have been reported elsewhere [17]. Briefly, in the double-blind phase, hepatitis B 'e' antigen (HBeAg)-positive and HBeAg-negative CHB patients were randomized to receive either TDF or adefovir dipivoxil for 48 weeks before entering an open-label phase of once-daily TDF treatment. Efficacy and safety data have been published previously, after 48 weeks and after 240 weeks of therapy [5,17].

Liver biopsy was a pre-specified procedure and samples were taken at baseline (within 6 months before screening) and between weeks 44 and 48. In addition, a non-mandatory biopsy was taken between treatment weeks 228 and 240 (year 5) [5]. One independent central pathologist, blinded with regard to the timing of biopsy and treatment assignment, examined all biopsy slides prospectively. Fibrosis was staged with the modified Ishak histological activity index scale ranging from 0–6. Baseline data of trial participants who had complete biopsy and laboratory data *vs.* those who did not were compared using the Wilcoxon rank-sum test (continuous variables), and the Fisher exact test (categorical variables) to confirm included patients were representative of the entire study population.

APRI and FIB-4 scores

APRI and FIB-4 scores were calculated at baseline and week 240 for all patients using clinical laboratory data based on the following formulas:

$$\text{APRI} = ([\text{AST}/\text{ULN}^*]/\text{platelet count}) \times 100$$

$$\text{FIB-4} = (\text{age} \times \text{AST})/(\text{platelet count} \times \sqrt{\text{ALT}})$$

Note: *ULN = upper limit of normal; 34 U/L for females, 36 U/L for males.

These laboratory results were also available at protocol-specified follow-up intervals during treatment, allowing calculation of APRI and FIB-4 at weeks 24, 48, 72, 96, 120, 144, 168, 192, 216 and 240, although corresponding biopsy samples were not available at these time points.

The APRI and FIB-4 scores have more than one cut-off point for specific fibrosis stages, aimed to maximize the sensitivity and specificity of diagnosis. For interpreting APRI, two different scales have previously been proposed [10]. The first scale aims to identify patients with cirrhosis (defined as Ishak stage 5–6); an APRI score >2.00 is the cut-off for cirrhosis whereas a score <1.00 is used to predict Ishak stage of 0–4. The second scale detects clinically significant fibrosis (Ishak stage 3–6); an APRI score >1.50 is the cut-off for significant fibrosis, whereas a score <0.50 predicts an Ishak stage of 0–2 [10]. Similarly, two different scales for interpreting FIB-4 scores have been proposed [11]. The first scale identifies patients with advanced fibrosis (Ishak stage 4–6); a FIB-4 score >3.25 is the cut-off for advanced fibrosis and a score <1.45 classifies patients as Ishak stage 0–3. The second scale for FIB-4 detects clinically significant fibrosis (Ishak stage 2–6); a FIB-4 score >1.00 indicates clinically significant fibrosis, whereas a cut-off of <0.60 predicts absence of fibrosis (Ishak 0–1). Of note, both APRI and FIB-4 have gaps between sets of cut-off values, creating an unclassifiable zone.

The current analysis applied both scales and proposed cut-off values for APRI and FIB-4 scores to predict the Ishak fibrosis stage. The predicted Ishak fibrosis stage was compared with the observed Ishak fibrosis stage for each patient who underwent liver biopsy at baseline and at week 240. The distributions of APRI and FIB-4 scores for each observed Ishak fibrosis stage were examined at baseline and week 240.

The predictive performance of both sets of cut-offs for APRI and FIB-4 for Ishak fibrosis stage against liver biopsy was expressed using positive predictive values (PPVs). Because of the unclassifiable zone between the APRI and FIB-4 cut-off thresholds, three sets of PPVs were calculated to compare the results of applying each of the two different cut-off scales for both APRI and FIB-4: (1) the lower bound, (2) the upper bound, (3) the total number of patients correctly categorized by the set of cut-off values.

We also considered whether new cut-off thresholds for APRI and FIB-4 scores, optimized for CHB patients, could be determined on the basis of the current data set. ROC curve analysis was performed to determine the different fibrosis levels of interest for APRI and FIB-4 respectively. Three sets of cut-offs for cirrhosis (Ishak: 5–6) and advanced fibrosis (Ishak: 3–6) were calculated: (1) sensitivity $\geqslant$90%, (2) specificity $\geqslant$90% or (3) an optimized balance between sensitivity and specificity (Youden's index) [18].

The distributions of calculated APRI and FIB-4 scores according to the observed Ishak stages were compared using a Cuzick test of trend (significance level $p = 0.05$) [19]. The association between change in Ishak stage and fibrosis score was analyzed using the non-parametric Kruskal-Wallis test (significance level $p = 0.05$) by unit change in Ishak stage. Given the small count of patients increasing their Ishak stage, these were grouped into a single category.

Results

Baseline data

Complete baseline data were available for 575/641 CHB patients; 66 patients were excluded (seven patients without biopsy data and 59 missing $\geqslant$1 laboratory component of APRI or FIB-4. The median age was 40 years; (75% male, 59% HBeAg-negative)) (Table 1). Patients had elevated serum AST and ALT and median HBV DNA was 7.9 $\log_{10}$ copies/ml, reflecting the trial enrolment criteria. The majority of patients (97%) had Ishak fibrosis stage $\geqslant$2 (24% had cirrhosis [Ishak 5 or 6]). The median APRI score