Full Title:

BIOMARKERS OF HEPATIC FIBROSIS IN CHRONIC HEPATITIS C: A COMPARISON OF TEN BIOMARKERS EMPLOYING TWO DIFFERENT ASSAYS FOR HYALURONIC ACID

Authors:

Sudeep Tanwar[1] M.D, Paul M.Trembling[1] M.D., Brian J. Hogan[1] M.D., Elizabeth L. Ellis[2] M.D., Julie Parkes[2] M.D. Scott Harris[2] PhD., Paul Grant[3] PhD, Eleni Nastouli[3] M.D., Mathias Ocker[4] M.D., Klaus Wehr[5] M.D., Christoph Herold[5] M.D., Daniel Neureiter[6] M.D., Detlef Schuppan[7-10] M.D., William M. Rosenberg[1] M.D.

Affiliations:

1. Institute for Liver & Digestive Health, University College London, London, United Kingdom.
2. Public Health Sciences and Medical Statistics, University of Southampton, Southampton, United Kingdom
3. Department of Clinical Microbiology and Virology, University College London Hospitals NHS Foundation Trust, London, United Kingdom.
4. Bayer Pharma AG, Clinical Sciences - Experimental Medicine Oncology, Berlin, Germany.
5. Department of Medicine 1, University of Erlangen, Erlangen, Germany.
6. Institute of Pathology, University of Salzburg, Salzburg, Austria.
7. Institute of Molecular and Translational Medicine, University of Mainz, Germany.
8. Department of Medicine, University of Mainz, Mainz, Germany.
9. Research Center for Immunology (FZI), University of Mainz, Germany.

10. Division of Gastroenterology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA.

Corresponding Author:

Dr Sudeep Tanwar,

UCL Institute for Liver and Digestive Health

Division of Medicine,

University College London

Royal Free Campus, Rowland Hill Street

Hampstead

LONDON NW3 2PF

Direct line +44 (0)207 794 0500 x 31157 Fax No 0207 317-7746

e-mail: Sudeep.Tanwar@nhs.net

Author Contributions: ST, PT, BH, JP, EN, MO, KW, CH, DN, DS and WMR contributed to the study design and manuscript drafting. EE, PG, EN, MO, KW, CH, DS were responsible for patient recruitment, sample collection and processing. DN scored the biopsy specimens used in the study. ST, JP, and SH performed the statistical analyses.

Conflicts of Interest: William Rosenberg is an inventor of the ELF test and has received support for research and speaking from Siemens Healthcare Diagnostics Inc. The proprietary assays developed for the ELF test by Siemens Healthcare Diagnostics Inc. (Tarrytown, New York, United States of America) were used and analyses were performed on an Immuno-1 auto-analyser (Siemens Healthcare Diagnostics Inc., Tarrytown, New York, United States of America).

All the other authors have no conflicts of interest to declare.

Abbreviations: AUROC, area under receiver operating characteristic curve; ELF, Enhanced Liver Fibrosis; HA, hyularonic acid; CHC, chronic hepatitis C; PEG-INFα2b, pegylated interferon alfa-2b; PIIINP, terminal peptide of pro-collagen III; TIMP-1, tissue inhibitor of matrix metaloproteinase-1; SVR, sustained virological response; SD, standard deviation; ordAUROC, AUROC corrected by Obuchowski measure

Acknowledgements: The PROFI-C study was endorsed by the German Hepatitis Network (HepNet). Prof. Otto Dietze co-scored the liver biopsy specimens. The Fibrometer scores were calculated by Prof. Paul Calès.

Number of Tables: 6; Number of Figures: 2

ABSTRACT

Background: advancing fibrosis is regarded as the most important factor when stratifying patients with CHC for retreatment.

Goals: (1) to compare the performance of 10 biomarkers of fibrosis, including patented tests, amongst patients with CHC and treatment failure; (2) to assess the impact on biomarker performance of using 2 different assays of HA.

Study: for 80 patients, liver histology (Metavir) was compared to biomarker scores using sera obtained within 6 months of liver biopsy (indirect biomarkers -AST:ALT ratio, APRI, Forns index, FIB-4, Fibrometer-V3G, direct biomarkers - ELF, Fibrospect-II, Hyularonic acid-HA, Fibrometer-V2G, Hepascore). Direct biomarker scores were calculated using 2 validated assays for HA (ELISA-Siemens, radiometric-Pharmacia).

Results: using the ELISA assay for HA to calculate the direct panels, all 10 of the biomarkers exhibited comparable overall discriminatory performance (unweighted Obuchowski-measure, ordROC 0.92-0.94, $p\text{-value}>0.05$) except AST:ALT ratio and APRI (ordROC 0.86-0.88, $p\text{-value}<0.05$). For the detection of moderate (F2-4) and advanced (F3-4) fibrosis, the AUROC of Fibrometer-2G were significantly higher than AST:ALT ratio and APRI but none of the other biomarkers. Good correlation was observed between the two HA assays (Intra-class correlation coefficient=0.873) with the ELISA assay exhibiting superior diagnostic performance (ordROC 0.92 Vs.0.88, $p\text{-value}=0.003$). Importantly, the performance of many of the direct biomarkers at their diagnostic thresholds was heavily influenced by the choice of HA assay.

Conclusions: Whilst many biomarkers exhibited good diagnostic performance for the detection of advancing fibrosis our results indicate that diagnostic performance may be significantly affected by the selection of individual component assays.

Abstract Word Count: 250

Keywords: Chronic Hepatitis C, Retreatment, Biomarkers, Fibrosis, Hyularonic acid

INTRODUCTION

Amongst patients with chronic hepatitis C (CHC) and prior treatment failure, the development of advancing fibrosis is currently regarded as the most important factor when stratifying patients for retreatment.[1] Patients with CHC and prior treatment failure with advanced or rapidly advancing fibrosis are candidates for early retreatment as they are at immediate risk of the complications of chronic liver disease (CLD) including portal hypertension, liver failure and hepatocellular carcinoma. The reference standard for assessing hepatic fibrosis in CHC remains the histological staging of a liver biopsy specimen. The limitations of liver biopsy including sampling error[2], inter- and intra-observer variability[3], and procedural complications[4] As a result, non-invasive methods have been developed for the cross-sectional staging of liver fibrosis. Hitherto, biomarkers of liver fibrosis have been widely studied for their ability to discriminate between different stages of liver fibrosis.[5] Biomarkers can be categorized in several ways including into indirect biomarkers (combinations of parameters which are related to liver function including ALT) and direct biomarkers (measuring parameters related to the processes involved in liver matrix turnover. Direct biomarkers may also be combined with indirect biomarkers). Further distinctions can be made into those markers deemed specific for single disease such as CHC (for example Fibrometer V2G) or those markers which have been validated for use in a variety of liver diseases such as Hepascore or ELF.

Hitherto, non-invasive biomarkers have been subjected to numerous derivation and validation studies which have collectively suggested that many of these tests have reliable performance for the detection of severe fibrosis in CHC.[6] Due to factors such as spectrum bias[7], any comparison of the performance of biomarkers derived

from patient populations with differing characteristics can be subject to type I error. Increasingly the use of biomarkers to assess fibrosis continues to move from a research application to become part of clinical practice. Furthermore, for the majority of biomarker panels, the algorithms combining component tests are published and widely available. It is therefore essential to understand whether biomarker performance is affected by the use of constituent component assays other than those used during derivation studies. This impact on biomarker discriminatory performance may be experienced both overall and at individual diagnostic thresholds.

This study had the following aims: (1) to validate and compare the performance of biomarkers of fibrosis, including patented assays, in patients with specifically with CHC and prior treatment failure (2) to determine whether the diagnostic performance of direct panels that contain HA as a constituent component are influenced by the choice of HA assay.

METHODS

Patient Population

Patients in this study were enrolled in the PROFI-C trial which was an investigator-initiated, prospective, randomized trial involving 18 centers in Germany and Austria investigating the effect of high dose silymarin plus pegylated interferon alpha 2b in non-responders or relapsers to standard treatment for CHC. Ethical approval was granted by the local ethics committees of the participating centers with initial ethical approval granted by Clinical **Ethics Committee** at the **University Hospital Erlangen**. One hundred and eight patients participated in the PROFI-C study.[8] Participants were male and female patients aged between 18 and 65 years with evidence of CHC (COBAS Amplicor HCV Monitor, Roche Molecular Diagnostics, Mannheim, Germany) after failure therapy with either interferon or pegylated interferon and ribavirin. Patients were also required to have histologically proven chronic hepatitis on a liver biopsy specimen (as interpreted by two histopathologists with consensus) within 6 months prior to entry into the study. Written consent was obtained from all patients before admission to the study.

Exclusion criteria included acute hepatitis, therapy with steroids or immunosuppressive drugs in the previous three months, Child-Pugh stage B or C cirrhosis, thrombocytopenia ($<100 \times 10^9$/L), leucopenia ($<3 \times 10^9$/L), other chronic liver diseases, autoimmune diseases, HIV infection, alcohol abuse (defined as the consumption of >40g per day in males and >20g per day in females), active drug abuse, pregnancy or psychiatric diseases including depression.

Histological assessment

Liver biopsies were fixed in formalin and embedded in paraffin. Hematoxylin-Eosin staining was used for grading of inflammation and the Chromotrope-aniline blue staining for staging the amount of liver fibrosis. [9, 10] All specimens were graded and staged according to the 5 stage Metavir Score. [11] Histological assessment was performed by 2 independent pathologists (D.N. and O.D.) who were blinded to the clinical data and randomization status of the patients in the study. Interobserver variability was determined by the Kappa statistic (Kappa=0.624). All liver biopsy specimens that were discordantly staged were re-reviewed by both pathologists with a final score determined after further discussion.

Sample Collection and Calculation of Biomarkers **for CHC**

Only patients recruited into PROFI-C who underwent liver biopsy prior to therapy and had stored sera were evaluated in this study (n=80). Patient samples were tested for hematological and biochemical parameters. Serum samples were stored at -70ºC prior to transfer to the central laboratory, where serum samples were analyzed for levels of HA, TIMP-1 and PIIINP using the proprietary assays developed for the ELF test by Siemens Healthcare Diagnostics Inc (Tarrytown, New York, USA). The assays are magnetic particle separation immunoassays and were performed on the ADVIA Centaur® immunoassay system (Siemens Medical Solutions Diagnostics Inc, Tarrytown, New York, USA). Hepatitis C virus RNA was quantified and genotyping performed on all samples. A full description of the virologic analysis is described in the supplementary data.

Biomarkers Evaluated in this study

Ten biomarkers of fibrosis (table 1) were evaluated in this study which can be characterized into indirect (AST to ALT ratio[12], AST to Platelet Ratio Index (APRI)[13], Forns Index[14], FIB-4[15], Fibrometer V3G[16]), and direct biomarkers (Hepascore[17], Fibrometer V2G[18], ELF [19], Fibrospect II[20], HA[21]). A full description of how these marker scores were calculated is described within the supplementary data.

Exploring the Impact of using an alternative assay for HA on the performance of the direct markers

Serum levels of HA were also measured locally in the PROFI-C study with a radiometric assay (Pharmacia AB, Uppsala, Sweden) which has been validated for the detection of fibrosis in CHC both as a single marker[21] and as the constituent component of biomarker panels.[22] This allowed us to explore whether the performance of both the direct markers, as well as the HA tests individually would be affected by the choice of assay.

Statistical Analyses

Statistical analyses were performed using SPSS for Windows (version 20, SPSS Inc, Chicago, IL) and R for Windows (version 2.15.1, The R Foundation for Statistical Computing). Patient demographic and clinical laboratory characteristics were descriptively summarized and reported as mean ± standard deviation (SD) and range. All tests were two-sided and statistical significance assessed at the 0.05 threshold. The diagnostic performance of the biomarkers as compared to liver biopsy was assessed using receiver operating characteristic (ROC) curves. The AUROC

and 95% confidence intervals of AUROC were calculated. Good performance for a test within our studied cohort was defined as an AUROC > 0.8.[23] The Obuchowski[24] method of correcting for spectrum effect was applied in a similar fashion to previously published literature [Supplementary data]. As the severity of histological liver fibrosis in patients with CHC with prior treatment failure has not been well characterized, the Obuchowski measure presented in this study has not been weighted according to the prevalence of fibrosis stages in a reference population. AUROC were compared using the method of Delong.[25] Sensitivity, specificity and predictive values were calculated at thresholds derived from ROC curves. The thresholds evaluated included those previously proposed for the respective biomarkers markers for the detection of moderate and advanced fibrosis which were compared with both the Q-point (where sensitivity and specificity are equal) and the Youden cut-off (highest sum of sensitivity and specificity minus 1) . Logistic regression was used to determine whether combinations of biomarker panels were more effective than individual panels in discriminating between patients with and without moderate fibrosis, advanced fibrosis and cirrhosis.

RESULTS

Patient characteristics and baseline histology

The baseline characteristics of the 80 patients included in this study are displayed in table 2. The majority of the patient population was male and had a mean age of 48.5 years. The predominant HCV genotype was genotype 1.

Comparison of all 10 markers (direct panels calculated with Siemens HA assay) (Table 3)

Ability to Discriminate Moderate Fibrosis (F0-1 Vs. F2-4)

Whereas all the direct panels and HA as an individual assay were able to discriminate between patients with and without moderate fibrosis with an AUROC of >0.8, the only indirect panel that achieved this level of performance was Fibrometer V3G (AUROC 0.86). Of all the markers tested, Fibrometer V2G generated the highest AUROC (0.88) for the detection of moderate fibrosis which was significantly higher *(p=value<0.05)* than all of the indirect panels other than Fibrometer V3G, or any of the other direct markers *(p-value>0.05)* [supplementary data].

Ability to Discriminate Advanced Fibrosis (F0-2 Vs. F3-4)

Other than Fibrospect II, all the direct panels and HA alone achieved AUROC of >0.8 in their ability to discriminate between patients with and without advanced fibrosis; the only indirect panels that did not achieve AUROC > 0.8 was AST:ALT ratio (AUROC 0.65) and APRI (AUROC 0.71). For the detection of advanced fibrosis, Fibrometer V2G generated the highest AUROC (0.84) which was significantly higher *(p-value<0.05)* than AST:ALT ratio and APRI but none of the other markers tested *(p-value>0.05)* [table 3 and supplementary data].

Ability to Discriminate Cirrhosis (F0-3 Vs. F4)

All the markers tested achieved AUROC >0.8 in their ability to discriminate between patients with and without cirrhosis apart from AST:ALT ratio and APRI. Forns index generated the highest AUROC (0.92), [supplementary data].

Overall Performance (Obuchowski Measure) (Table 3, Figures 1 and 2)

The highest unweighted Obuchowski measure (ordROC) was attained by Fibrometer V2G (0.94) and Fibrometer V3G (0.94) which were significantly higher *(p-value<0.05)* than those attained by AST:ALT ratio (0.86) and APRI (0.88) but none of the other markers tested. Of the 2 other direct markers tested, ELF generated the highest ordROC (0.93) which was the best performing non disease-specific marker of fibrosis together with Hepascore (0.93)

Performance of the markers at their published thresholds for the detection of moderate fibrosis (F2-4) – prevalence 63% [Table 5]

Diagnostic thresholds have been described for all of the markers in this context other than AST:ALT ratio.

Indirect Markers

Overall, the sensitivity and specificity of the indirect markers at their published thresholds for detecting moderate fibrosis were comparable to those observed in their original publications other than Fibrometer V3G (more sensitive (100%), less specific (30%)). The highest positive likelihood ratio for the detection of moderate fibrosis (2.6) was generated by the proposed threshold of APRI (PPV of 82% in our population).

Direct Biomarkers

The sensitivity and specificity of ELF, Hepascore and Fibrospect II at their published thresholds for the detection of moderate fibrosis in our study were comparable to those observed in their original publications. This was not the case for the proposed diagnostic threshold of Fibrometer 2G which was markedly more sensitive (100%) but less specific (30%). The highest positive predictive value (88%) and likelihood ratio (4.1) for the detection of moderate fibrosis were obtained using the published thresholds of ELF.

Performance of the markers at their published thresholds for the detection of advanced fibrosis (F3-4) – prevalence 40% [Table 6]

Diagnostic thresholds have not been proposed for a diagnosis of advanced fibrosis for Fibrospect II or any of the indirect markers other than FIB-4.

Indirect Markers

The performance attributes of the proposed threshold of FIB-4 (1.45) for the detection of advanced fibrosis appeared consistent with its performance in our study population (Q-point 1.44). This FIB-4 threshold generated a positive likelihood ratio of 2.3 (PPV 61%), negative likelihood ratio of 0.29 (NPV 84%) and diagnostic odds ratio of 7.9 for the detection of advanced fibrosis.

Direct Biomarkers

Whilst the performance of direct biomarkers at their proposed thresholds for the detection of advanced fibrosis was comparable to that seen in their original publications, the Q-point and Youden index of ELF and HA in our study were the most similar to their proposed thresholds. The lowest negative likelihood ratio for the

exclusion of advanced fibrosis was attained by Fibrometer 2G (0.06) which generated a NPV of 96%. The highest positive likelihood ratio for the detection of advanced fibrosis was generated by HA itself (3.9) which resulted in a PPV of 72%. At their proposed thresholds for the detection of advanced fibrosis, Fibrometer 2G generated the highest diagnostic odds ratio (32.3).

Exploring whether combinations of biomarkers are more effective using logistic regression

Logistic regression was used to explore the use of combinations of biomarkers. The results presented below outline which biomarker combinations produced the highest numerical AUROC. As the sample size employed in our study was not powered to investigate this, these increases in AUROC were not statistically significant.

Moderate fibrosis (F2-4)

Stepwise logistic regression identified that both Fibrometer V2G *(p-value<0.001, OR 4.98, 95% CI 13.1-19029)* and ELF *(p-value=0.008, OR 3.07, 95% CI 1.16-8.09)* were significantly associated with a diagnosis of at least moderate fibrosis after accounting for the remaining variables as potential confounders. A combination of Fibrometer V2G and ELF had an AUROC of 0.90 (95% CI 0.84-0.97).

Advanced Fibrosis (F3-4)

Stepwise logistic regression identified that both Forns index *(p-value<0.001, OR 2.16, 95% CI 1.32-3.55)* and ELF *(p-value=0.001, OR 2.72, 95% CI 1.39-5.34)* exhibited statistical independence for a diagnosis of advanced fibrosis. A combination of Forns and ELF had an AUROC of 0.87 (95% CI 0.78-0.96)

Cirrhosis (F4)

Stepwise logistic regression identified that the Forns test was the only biomarker significantly associated with a diagnosis of cirrhosis after accounting for the remaining markers as potential confounders *(p-value<0.001, OR 3.019, 95% CI 1.50-6.05).*

Effect of using a different assay for HA to calculate direct Panels (table 3)

Relationship between the 2 HA assays (figure 3)

Both assays exhibited a high degree of correlation (Intra-class correlation ICC=0.873, 95% CI 0.802-0.919, *p<0.0001*). The relationship between the 2 assays is presently graphically within the supplementary data and is described by the equation:

$HA_{(Pharmacia)} = 33.67 + [0.49 \times HA_{(Siemens)}]$

Effect of the Pharmacia HA on the ability of the direct markers to discriminate between different degrees of fibrosis (Table 3)

The AUROC of HA as assessed by the Siemens assay was significantly higher than that attained using the Pharmacia assay for the detection of moderate fibrosis (AUROC 0.79 Vs 0.68, *p-value=0.005*) and advanced fibrosis (AUROC 0.80 Vs.0.72, *p-value=0.042*) but not for the detection of cirrhosis (AUROC 0.89 Vs. 0.85, *p-value>0.05*). Furthermore, the overall discriminatory power of the Siemens assay was significantly higher than that of the Pharmacia assay (ordROC 0.92 Vs. 0.88, *p-value=0.003*). Calculation of the direct markers with the Pharmacia assay resulted in

numerically lower AUROCs than when calculated with the Siemens assay. This was most marked in the context of Fibrospect II in its ability to discriminate between moderate fibrosis (AUROC 0.66 Vs. 0.79) and advanced fibrosis (AUROC 0.83 Vs. 0.70).

Effect of the Pharmacia HA on the performance of the direct markers at their published diagnostic thresholds

Whereas Fibrometer 2G was largely unaffected, the use of the Pharmacia assay resulted in a reduction in the performance of the other direct markers at their published diagnostic thresholds.

The most marked impact was seen when the ELF test was calculated with the 2 different assays for HA. When calculated with the ELISA HA assay, the ELF test at a threshold of 9.13 had a positive likelihood ratio of 3.9 (PPV 88%) and negative likelihood ratio of 0.36 (NPV 62%) for the detection of moderate fibrosis. However when calculated with the radiometric assay the ELF test at a threshold of 9.13 had a positive and negative likelihood ratio of both 1. This resulted in an unchanged *a priori* and *a posteriori* probability of moderate fibrosis regardless of the ELF score obtained in our study (figure 2). This resulted in all patients without moderate fibrosis being incorrectly classified as having moderate fibrosis.

DISCUSSION

In this study we have compared the performance of 10 biomarkers of fibrosis in a population of patients with CHC and prior treatment failure. We have evaluated the diagnostic performance of these biomarkers by studying their ability to discriminate between moderate fibrosis, advanced fibrosis, cirrhosis and all fibrosis stages (ordROC) and by validating their published diagnostic thresholds. Within our population, all the direct markers tested had good diagnostic performance for detecting moderate fibrosis. However, the only indirect marker that achieved this level of performance for the detection of moderate fibrosis was Fibrometer V3G. Aside from APRI and AST:ALT ratio, the comparisons between the other 8 markers in their ability to discriminate between advanced fibrosis and cirrhosis were more favorable. Both the Forns index and FIB-4 exhibited excellent performance for discriminating between patients with and without a histologic diagnosis of cirrhosis. This is biologically plausible given that these indirect markers of fibrosis incorporate the platelet count which is classically affected by cirrhosis and portal hypertension. With regard to overall fibrosis stratification, both of the virus specific patented panels, Fibrometer V2G and V3G, exhibited the highest level of performance. Of the other 8 panels which have been proposed to be used in all etiologies of liver disease, the best performing tests were ELF and Hepascore.

Using logistic regression we found that indirect and direct markers demonstrated statistical independence for the detection of mild to advanced fibrosis. Whilst these combinations of indirect and direct markers resulted in a numerical improvement in performance, this advantage must be balanced against the need to perform more tests with the consequent increase in costs. Given the relatively modest sample size

employed in our study these findings will need to be validated in a larger cohort with sufficient statistical power.

In our study we have observed that markers with only 3 components (ELF and Fibrospect II) have comparable performance to those with as many as 8 components (Fibrometer V2G and V3G). Furthermore some of the more complex tests are dependent on the inclusion of demographic data that requires the collection, transmission and entry of clinical data with attendant costs. By contrast we have also witnessed that more complex tests are more forgiving of a variation in the performance of one its constituent components. In the context of biomarkers with constituent components that are relatively inexpensive and freely available this certainly is an advantage.

Within our population of patients with CHC and prior treatment failure, many of the tests performed differently at their diagnostic thresholds than reported within their reference studies in CHC. This may relate to variation in spectrum bias and to sample size; with a limited number of patients the AUROC curve may demonstrate peaks within a plateau that can hinder accurate threshold determination. We also note previous work which demonstrated that Fibrometer exhibited consistent performance after retesting both overall and at individual diagnostic thresholds.[26]

In our study we also explored the influence of assay selection for HA. The use of 2 closely correlated (ICC=0.873) assays for HA to calculate the direct markers resulted in changes both in overall diagnostic performance and at the diagnostic thresholds of the biomarkers. Despite the close correlation of these assays, the AUROC of HA as measured by the radiometric assay was significantly lower than that measured by the ELISA assay in the detection of moderate and advanced fibrosis. Furthermore,

when the direct markers were calculated with HA as measured using the radiometric assay, a modest reduction in the AUROC was observed in most instances. This was not the case, however, for Fibrospect II which suffered a marked reduction ability to discriminate moderate fibrosis (AUROC 0.79 Vs.0.66) and advanced fibrosis (AUROC 0.83 Vs.0.70). With regard to the performance of the biomarkers at their diagnostic thresholds, the use of a different assay for HA had the most marked effect of the performance of the ELF. Despite generating similar AUROC, the use of the Pharmacia assay rendered the ELF test much less useful as a diagnostic tool with both a positive or negative result at its proposed threshold producing an unchanged *a priori* and *a posteriori* probability of moderate fibrosis in our cohort. These variations in test performance resulting from the choice of assay, have predominantly affected the markers with the fewest number of analytes. Whereas both ELF and Fibrospect II have only 3 constituent components the direct biomarkers with up to 8 components are more forgiving of the use of an inferior performing component. This emphasizes the need to use the specified individual component assays that have been validated for a particular biomarker both in research studies and in clinical practice. Our observation that the diagnostic performance of Fibrometer was resistant to the choice of HA assay has also been made by previous investigators.[27, 28]

Regardless of treatment status, patents with CHC and a diagnosis of cirrhosis should embark on a surveillance program for the early detection of hepatocellular carcinoma and endoscopic features of portal hypertension. Our results have shown that all 10 of the markers studied in our population have good performance (AUROC >0.8) for the detection of cirrhosis with some having excellent performance (AUROC>0.9). Whilst this may be in part due to the low prevalence of cirrhosis in our study, this level of

performance from even the more basic indirect tests is encouraging particularly as not all centers will have access to some of the expensive proprietary panels that we have investigated.

In summary, we have compared the performance of 10 biomarkers in a population of patients with CHC and prior treatment failure. We have observed that many of the markers have good diagnostic performance for their ability to discriminate between moderate and advanced fibrosis stages. Overall, the best performing markers were the virus specific panels, Fibrometer V2G and V3G. With regard to biomarkers developed for use in all etiologies of liver disease the best performing panels were ELF and Hepascore. Importantly, we have witnessed that the performance of the markers at their diagnostic thresholds can be variable and influenced by the choice of their component assays. This emphasizes the need to use component assays that have been validated for a particular biomarker.

REFERENCES

1. Tillmann HL. Hepatitis C infection and presence of advanced fibrosis: wait or treat? Why wait? There is no time to lose, is there? *J Hepatol*. 2013;58:412-414.
2. Bedossa P, Dargere D and Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology*. 2003;38:1449-1457.
3. Regev A, Berho M, Jeffers LJ, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol*. 2002;97:2614-2618.
4. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. *Hepatology*. 2009;49:1017-1044.
5. Schuppan D and Afdhal NH. Liver cirrhosis. *Lancet*. 2008;371:838-851.
6. Chou R and Wasson N. Blood tests to diagnose fibrosis or cirrhosis in patients with chronic hepatitis C virus infection: a systematic review. *Ann Intern Med*. 2013;158:807-820.
7. Ransohoff DF and Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med*. 1978;299:926-930.
8. Ocker M, Neureier D, Schuppan D, et al. The combination of silymarin and pegylated interferon alpha2b attenuates the progression of liver fibrosis in prior nonresponders and relapsers with chronic hepatitis C infection - the PROFI C trial. *Z Gastroenterol*. 2008;46:P4_07.
9. Churg J and Prado A. A rapid Mallory trichrome stain (chromotrope-aniline blue). *AMA Arch Pathol*. 1956;62:505-506.
10. Barikbin R, Neureiter D, Wirth J, et al. Induction of heme oxygenase 1 prevents progression of liver fibrosis in Mdr2 knockout mice. *Hepatology*. 2012;55:553-562.
11. Bedossa P and Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology*. 1996;24:289-293.
12. Sheth SG, Flamm SL, Gordon FD, et al. AST/ALT ratio predicts cirrhosis in patients with chronic hepatitis C virus infection. *Am J Gastroenterol*. 1998;93:44-48.
13. Snyder N, Gajula L, Xiao SY, et al. APRI: an easy and validated predictor of hepatic fibrosis in chronic hepatitis C. *J Clin Gastroenterol*. 2006;40:535-542.
14. Forns X, Ampurdanes S, Llovet JM, et al. Identification of chronic hepatitis C patients without hepatic fibrosis by a simple predictive model. *Hepatology*. 2002;36:986-992.
15. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology*. 2007;46:32-36.
16. . *Report to the World Health Organisation, 1957-1960.* Sydney, N.S.W.: V.C.N. Blight, Govt. Printer; 1961.
17. Adams LA, Bulsara M, Rossi E, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. *Clin Chem*. 2005;51:1867-1873.
18. Cales P, Halfon P, Batisse D, et al. Comparison of liver fibrosis blood tests developed for HCV with new specific tests in HIV/HCV co-infection. *J Hepatol*. 2010;53:238-244.
19. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology*. 2004;127:1704-1713.
20. Patel K, Gordon SC, Jacobson I, et al. Evaluation of a panel of non-invasive serum markers to differentiate mild from moderate-to-advanced liver fibrosis in chronic hepatitis C patients. *J Hepatol*. 2004;41:935-942.
21. McHutchison JG, Blatt LM, de Medina M, et al. Measurement of serum hyaluronic acid in patients with chronic hepatitis C and its relationship to liver histology. Consensus Interferon Study Group. *J Gastroenterol Hepatol*. 2000;15:945-951.
22. Leroy V, Hilleret MN, Sturm N, et al. Prospective comparison of six non-invasive scores for the diagnosis of liver fibrosis in chronic hepatitis C. *J Hepatol*. 2007;46:775-782.
23. Weinstein MC and Fineberg HV. Clinical Decision Analysis (Saunders). 1980.
24. Obuchowski NA. An ROC-type measure of diagnostic accuracy when the gold standard is continuous-scale. *Stat Med*. 2006;25:481-493.

25. DeLong ER, DeLong DM and Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.
26. Leroy V, Halfon P, Bacq Y, et al. Diagnostic accuracy, reproducibility and robustness of fibrosis blood tests in chronic hepatitis C: a meta-analysis with individual data. *Clin Biochem*. 2008;41:1368-1376.
27. Veillon P, Gallois Y, Moal V, et al. Assessment of new hyaluronic acid assays and their impact on FibroMeter scores. *Clin Chim Acta*. 2011;412:347-352.
28. Cales P, Veillon P, Konate A, et al. Reproducibility of blood tests of liver fibrosis in clinical practice. *Clin Biochem*. 2008;41:10-18.

FIGURE 1

Panel A: Distribution of fibrosis stages in the study population. Panels B-F: Boxplots of indirect biomarkers with respect to fibrosis stage in study population (Panel B: AST to ALT ratio, Panel C: APRI, Panel D: Forns Index, Panel E: FIB-4 Index, Panel F: Fibrometer V3G).

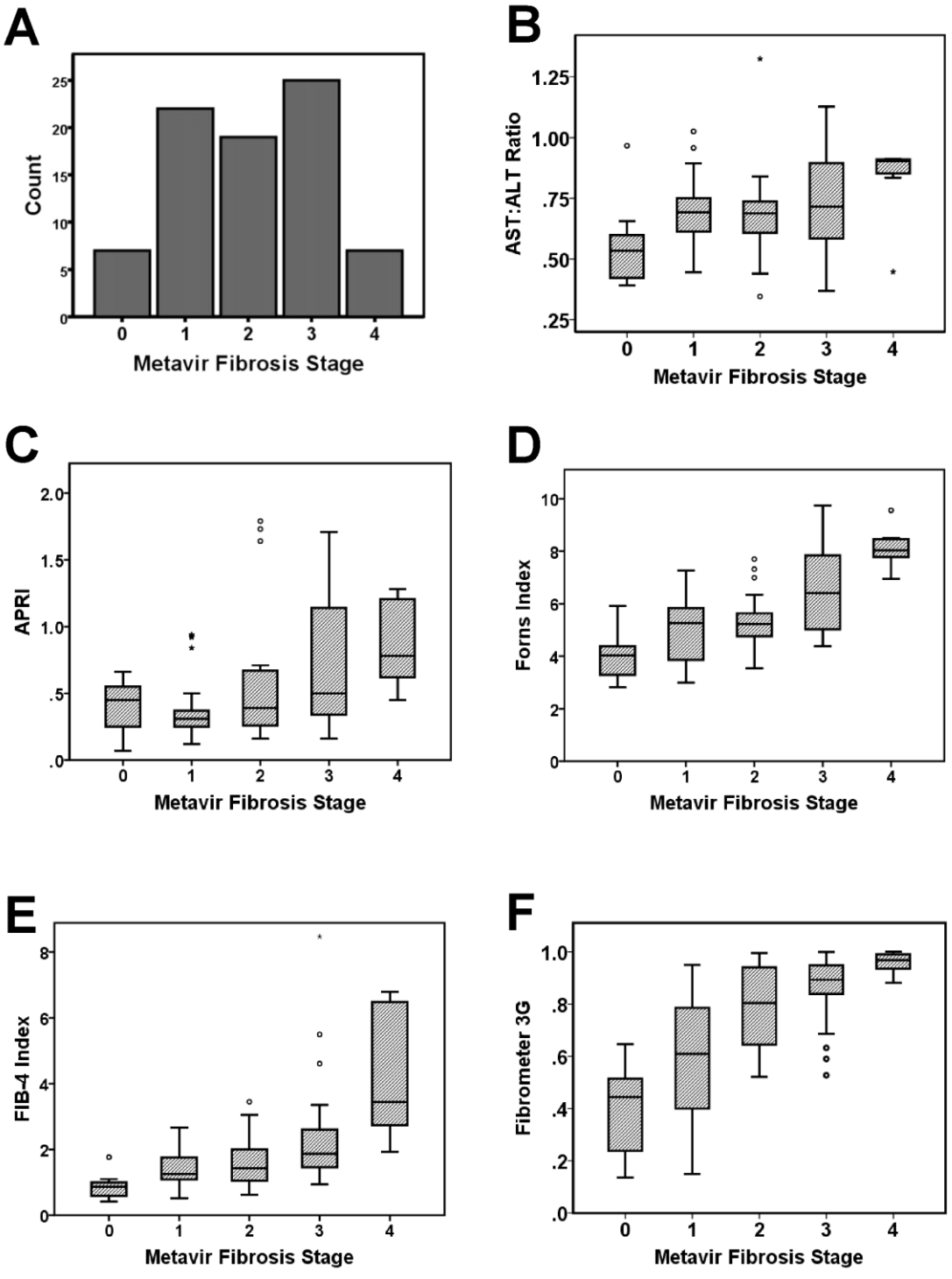

FIGURE 2

Panels A-E: Boxplots of direct biomarkers calculated with two assays for hyaluronic acid with respect to fibrosis stage in the study population (Panel A: Hyaluronic Acid, Panel B: Hepascore, Panel C: Fibrometer V2G, Panel D: ELF, Panel E: Fibrospect II).

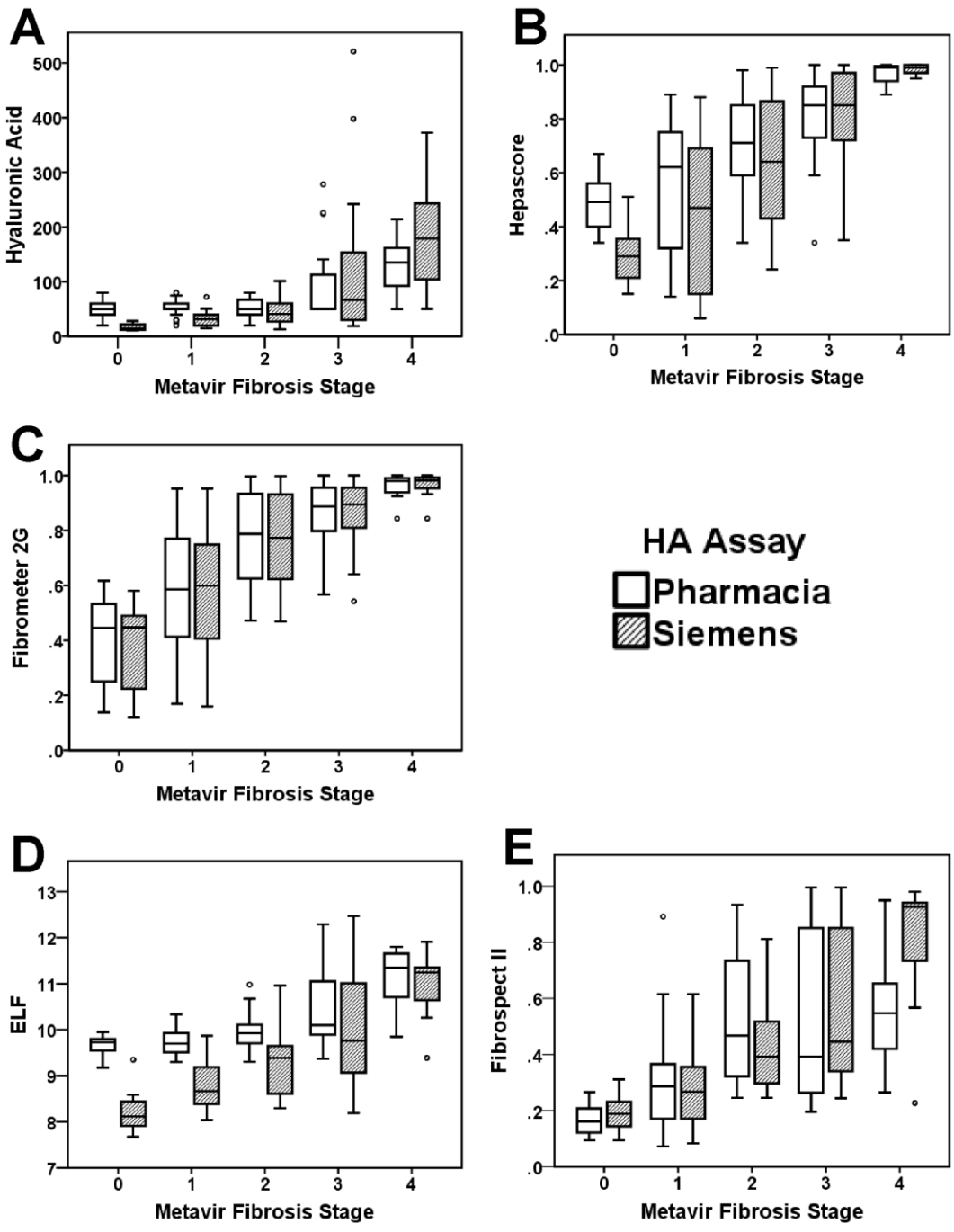

TABLE 1. The constituent components used to calculate the indirect, direct and hybrid markers used in the study.

Marker Panel	Type	Constituent Components													
		Liver Matrix				Biochemistry						Haematology		Demographics	
		HA	TIMP1	PIIINP	A2M	AST	ALT	GGT	Bil	Chol	Urea	PLT	PT	Sex	Age
AST:ALT ratio	Indirect					■	■								
APRI	Indirect					■						■			
FORNS	Indirect						■	■		■		■			
FIB4	Indirect					■	■					■			■
FIBROMETER 3G	Indirect				■			■			■	■	■	■	■
HEPASCORE	Direct	■			■			■	■					■	■
FIBROMETER 2G	Direct	■			■						■	■	■	■	■
HA	Direct	■													
ELF	Direct	■	■	■											

FIBROSPECT II	**Direct**	[redacted]													

TABLE 2. Baseline Demographics of the study cohort

Variable			n=80
Demographics	Age		48.9 ± 9.8
	Male n,%		40 (57%)
	BMI		24.3 ± 4.2
Virology	Genotype	1 (n,%)	67 (84%)
		2 or 3 (n,%)	7 (9%)
		4 (n,%)	6 (7%)
	Log Viral Load (IU/ml) (median, range)		6.21 (3.84 - 8.20)
Haematology	Hb (g/dl)		15.2 ± 1.4
	WCC (10^9/L)		6.4 ± 1.7
	PLT (g/L)		206 ± 61
	PI (%)		95.9 ± 9.6
Biochemistry	ALT (IU/L)		117 ± 97
	AST (IU/L)		81.6 ± 74.5
	GGT (IU/L)		102.5 ± 92.1
	A2M (g/L)		4.28 ± 1.21
	Bilirubin (µmol/L)		12.7 ± 7.5
	Albumin (g/L)		47.7 ± 7.6
	AFP (ng/ml)		8.2 ± 11.7
Fibrosis Stage	Metavir F0 (n,%)		7 (9%)
	Metavir F1 (n,%)		22 (28%)
	Metavir F2 (n,%)		19 (24%)
	Metavir F3 (n,%)		25 (31%)
	Metavir F4 (n,%)		7 (9%)

Values are presented as mean ± standard deviation unless otherwise specified

TABLE 3. Performance of the 10 biomarkers with respect to discriminating moderate fibrosis (F2-F4), advanced fibrosis (F3-F4) and cirrhosis (F4) and the overall diagnostic accuracy (all the direct biomarkers have been calculated using the Siemens assay for HA).

Marker	F0-1 (n=30) Vs. F2-4 (n=50)				F0-2 (n=48) Vs. F3-4 (n=32)				F0-3 (n=73) Vs. F4 (n=7)				Unweighted Obuchowski Measure	
	AUC	95% CI	P-value	Std. Error	AUC	95% CI	P-value	Std. Error	AUC	95% CI	P-value	Std. Error	ordAUROC	Std. Error
AST:ALT ratio	0.61	0.48-0.73	0.116	0.06	0.65	0.52-0.78	0.027	0.06	0.76	0.55-0.97	0.024	0.11	0.86	0.02
APRI	0.71	0.59-0.82	0.002	0.06	0.71	0.60-0.83	0.001	0.06	0.78	0.65-0.91	0.015	0.07	0.88	0.02
FORNS	0.78	0.67-0.88	<0.001	0.05	0.82	0.72-0.91	<0.001	0.05	0.92	0.86-0.98	<0.001	0.03	0.92	0.01
FIB4	0.77	0.67-0.88	<0.001	0.05	0.81	0.71-0.90	<0.001	0.05	0.90	0.82-0.98	<0.001	0.04	0.92	0.01
FIBROMETER 3G	0.86	0.78-0.94	<0.001	0.04	0.81	0.72-0.90	<0.001	0.05	0.86	0.77-0.98	0.002	0.05	0.94	0.01
HA (Siemens)	0.80	0.71-0.90	<0.001	0.05	0.80	0.70-0.90	<0.001	0.05	0.88	0.79-0.98	0.001	0.05	0.92	0.01
HEPASCORE	0.85	0.76-0.93	<0.001	0.04	0.83	0.74-0.92	<0.001	0.05	0.86	0.70-1.00	0.002	0.08	0.93	0.01
FIBROMETER 2G	0.88	0.80-0.95	<0.001	0.04	0.84	0.75-0.93	<0.001	0.04	0.88	0.77-1.00	0.001	0.05	0.94	0.01
ELF	0.84	0.73-0.92	<0.001	0.04	0.82	0.72-0.92	<0.001	0.05	0.89	0.79-1.00	0.001	0.10	0.93	0.02
FIBROSPECT II	0.84	0.76-0.93	<0.001	0.05	0.79	0.70-0.89	<0.001	0.05	0.83	0.62-1.00	0.004	0.11	0.92	0.01

TABLE 4. Performance of the direct biomarkers with respect to discriminating moderate fibrosis (F2-F4), advanced fibrosis (F3-F4), cirrhosis (F4) and the overall diagnostic accuracy (unweighted Obuchowski measure) when calculated using 2 assays for HA (Siemens and Pharmacia).

Marker	HA assay used	F0-1 (n=30) Vs. F2-4 (n=50)				F0-2 (n=30) Vs. F3-4 (n=50)				F0-3 (n=73) Vs. F4 (n=7)				Unweighted Obuchowski Measure	
		AUC	95% CI	P-value	Std. Error	AUC	95% CI	P-value	Std. Error	AUC	95% CI	P-value	Std. Error	ordAUROC	Std. Error
HA	Siemens	0.80	0.71-0.90	<0.001	0.05	0.80	0.70-0.90	0.001	0.05	0.88	0.79-0.98	0.001	0.05	0.92	0.01
	Pharmacia	0.69	0.57-0.80	0.006	0.06	0.72	0.56-0.84	<0.001	0.06	0.85	0.69-1.00	0.002	0.08	0.88	0.01
HEPASCORE	Siemens	0.85	0.76-0.93	<0.001	0.04	0.83	0.74-0.92	<0.001	0.04	0.86	0.70-1.00	0.002	0.08	0.93	0.01
	Pharmacia	0.81	0.72-0.91	<0.001	0.05	0.79	0.69-0.90	<0.001	0.05	0.84	0.65-1.00	0.003	0.10	0.92	0.02
FIBROMETER 2G	Siemens	0.88	0.80-0.95	<0.001	0.04	0.83	0.74-0.91	<0.001	0.04	0.88	0.77-1.00	0.001	0.05	0.94	0.01
	Pharmacia	0.87	0.79-0.95	<0.001	0.04	0.81	0.72-0.91	<0.001	0.04	0.86	0.75-0.97	0.002	0.05	0.94	0.01
ELF	Siemens	0.84	0.73-0.92	<0.001	0.04	0.82	0.72-0.92	<0.001	0.04	0.89	0.79-1.00	0.001	0.10	0.93	0.01
	Pharmacia	0.79	0.69-0.89	<0.001	0.05	0.79	0.69-0.90	<0.001	0.05	0.87	0.73-1.00	0.001	0.05	0.91	0.01
FIBROSPECT II	Siemens	0.84	0.76-0.93	<0.001	0.05	0.79	0.70-0.89	<0.001	0.05	0.83	0.62-1.00	0.004	0.11	0.92	0.02
	Pharmacia	0.81	0.71-0.91	<0.001	0.05	0.66	0.55-0.78	0.013	0.05	0.70	0.54-0.85	0.090	0.08	0.90	0.02

TABLE 5. Performance of the indirect and direct biomarkers evaluated in this study using the thresholds described in their original publications in detection of moderate fibrosis (F0-1 Vs F2-4, n=30 Vs n=50, prevalence 63%).

Marker	HA Assay used	Original publication test threshold and performance			Performance of original threshold in current Study									Current Study Thresholds	
		Threshold	Sens	Spec	Sens	Spec	number –ve Vs. +ve	PPV	NPV	LR+	LR-	DOR	Patients Correctly Classified	Q-point	Youden
AST:ALT	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.69	0.77
APRI	-	0.5	83.2	54	52	80	46 Vs. 34	82%	50%	2.6	0.60	4.3	62%	0.38	0.38
FORNS	-	5	88	71	78	57	28 Vs. 52	76%	60%	1.8	0.39	4.7	70%	5.4	4.5
FIB4	-	1.0	69.4	58.4	90	30	14 Vs. 66	69%	64%	1.3	0.30	4.4	68%	1.3	1.4
FIBROMETER 3G	-	0.440	81.3	74.1	100	30	9 Vs. 71	71%	100%	1.4	0	∞	74%	0.76	0.59
HA	Siemens	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	34	52
	Pharmacia				N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	53	53
HEPASCORE	Siemens	0.5	63	89	82	63	27 Vs. 53	79%	67%	2.2	0.29	7.8	75%	0.62	0.58
	Pharmacia				90	43	17 Vs. 63	73%	72%	1.6	0.23	6.8	73%	0.67	0.67
FIBROMETER 2G	Siemens	0.419	80	76	100	30	8 Vs. 72	71%	100%	1.4	0	∞	74%	0.70	0.69
	Pharmacia				100	27	8 Vs. 72	70%	100%	1.4	0	∞	73%	0.71	0.71
ELF	Siemens	9.13	73	64	70	83	38 Vs. 42	88%	62%	4.1	0.36	10.6	75%	8.99	9.20
	Pharmacia				100	0	0 Vs. 80	63%	37%	1.0	1.0	N/A	63%	9.88	10.02
FIBROSPECT II	Siemens	0.36	77	73	66	82	43 Vs. 37	86%	59%	3.9	0.41	9.5	72%	0.33	0.31
	Pharmacia				62	80	41 Vs. 39	84%	55%	3.1	0.48	6.5	69%	0.33	0.39

TABLE 6. Performance of the indirect and direct biomarkers evaluated in this study using the thresholds described in their original publications in detection of advanced fibrosis (F0-2 Vs F3-4, n=48 Vs n=32, prevalence 40%).

Marker	HA Assay used	Original Publication test threshold and performance			Performance of original threshold in current Study									Current Study Thresholds	
		Threshold	Sens	Spec	Sens	Spec	number –ve Vs. +ve	PPV	NPV	LR+	LR-	DOR	Patients Correctly Classified	Q-point	Youden
AST:ALT	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.71	0.78
APRI	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.47	0.40
FORNS	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	5.64	5.64
FIB4	-	1.45	74	80	81	65	37 Vs. 43	61%	84%	2.3	0.29	7.9	71%	1.68	1.44
FIBROMETER 3G	-	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.86	0.84
HA	Siemens	60	88	59	59	85	66 Vs. 14	72%	76%	3.9	0.48	8.2	75%	47	57
	Pharmacia				47	91	64 Vs. 16	78%	72%	5.2	0.58	9.0	73%	55	68
HEPASCORE	Siemens	0.5	88	74	90	50	27 Vs. 53	55%	88%	1.8	0.2	9	66%	0.73	0.72
	Pharmacia				91	31	17 Vs. 63	47%	84%	1.3	0.29	4.5	55%	0.74	0.81
FIBROMETER 2G	Siemens	0.628	84	79	97	50	25 Vs. 55	57%	96%	1.94	0.06	32.3	69%	0.82	0.81
	Pharmacia				91	50	27 Vs. 53	55%	89%	1.82	0.18	10.1	66%	0.84	0.73
ELF	Siemens	9.59	85	63	65	82	50 Vs. 30	71%	78%	3.19	0.23	13.8	75%	9.32	9.77
	Pharmacia				94	29	16 Vs. 64	47%	88%	1.32	0.20	6.4	55%	9.94	10.07
FIBROSPECT II	Siemens	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.36	0.42
	Pharmacia				N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.35	0.34