Published in final edited form as:

Clin Gastroenterol Hepatol. 2014 December ; 12(12): 2113–2120.e3. doi:10.1016/j.cgh.2014.04.037.

Multiplex Protein Analysis to determine Fibrosis Stage and Progression in Patients with Chronic Hepatitis C

Keyur Patel^{1,2}, Katja S. Remlinger³, Terence G. Walker⁴, Peter Leitner⁵, Joseph E. Lucas⁶, Stephen D. Gardner³, John G. McHutchison^{1,7}, Will Irving^{8,*}, and Indra Neil Guha^{8,*}

¹Duke Clinical Research Institute, Durham, North Carolina

²Liver Cell Biology, Centenary Institute, University of Sydney, Sydney, NSW, Australia

³GlaxoSmithKline, RTP, North Carolina

⁴OpGen, Gaithersburg, MD

⁵BioAgilytix Labs, RTP, NC

⁶Institute for Genome Sciences and Policy, Duke University, Durham, North Carolina

⁷Gilead Sciences, Foster City, CA

⁸NIHR Biomedical Research Unit in Gastrointestinal and Liver Diseases at Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, UK

Abstract

Backgrounds & Aims—Non-invasive tests cannot differentiate between adjacent stages of fibrosis, which limits assessment of disease progression and regression during therapy. We investigated whether levels of cytokines and extracellular matrix proteins in serum and biopsy samples can be used to determine actual stage of liver fibrosis in patients with chronic hepatitis C (CHC), and in prognosis.

Methods—We collected data from 383 treatment-naïve patients with CHC from the Duke Hepatology Clinical Research Database and Biorepository, from 2006 through 2009, for use in the training set. Serum samples were obtained from 100 individuals without CHC (controls). We selected 37 serum biomarkers for customized array analysis, using the SearchLight multiplex sandwich ELISA. Data from 434 treatment-naïve patients with CHC, obtained from the Trent HCV cohort, were used in the validation analysis. Multivariable modeling, marker selection, and

© 2014 The American Gastroenterological Association. Published by Elsevier Inc. All rights reserved.

Address for Correspondence: Keyur Patel, MD, GI/Hepatology Research, Duke Clinical Research Institute, PO Box 17969, Durham, NC 27715, Tel: 919 668 7802, Fax: 919 668 7164, keyur.patel@duke.edu.

*Co-Senior Authors

Disclosures: KSR and SDG are employees and stockholders of GSK; TGW and PL are former employees of GSK; JGM is an employee and stockholder of Gilead Sciences. KP, JEL, WI, and ING have nothing to disclose

Author Contributions: KP, KSR, TGW, PL SDG, JEL, JGW and WI, ING designed, conducted the study, and assisted in writing of this manuscript. KSR and JEL conducted the statistical analysis. Each author has approved the final submitted draft.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

validation included RandomForest and Obuchowski measures, with independent comparison to FibroSURE.

Results—Four serum markers (levels of hyaluronic acid, VCAM1, A2M, and RBP4) and age associated with fibrosis stage (F0–1, F2–3, or F4); these had Obuchowski measures of 0.85–0.89, with misclassification rates of 38% and 29% in training and validation sets, compared to 50% for the FibroSURE test. In the training set, area under the curve (AUC) values for the multiplex markers were similar to those from the FibroSURE test: stages F0 vs F1 (0.51 vs 0.53), F1 vs F2 (0.60 vs 0.59), F2 vs F3 (0.69 vs 0.72), and F3 vs F4 (0.51 vs 0.52). AUC values were similar in the validation cohort. In longitudinal analyses of 133 paired biopsies, 9 markers (level of alanine aminotransferase, γ -glutamyltranspeptidase, hyaluronic acid, ICAM1, interleukin (IL)4, CXCL10, CXCL9, and VCAM1) were associated with change in the histologic activity index (*P* values ranging from .000 to .049) and 4 (GMCSF, IL12, IL2, and MMP13) were associated with a change in fibrosis stage (*P* values ranging from .001 to .042).

Conclusions—We identified serum biomarkers that can be measured by multiplex ELISA to determine levels of fibrosis in patients with CHC, although misclassification is frequent and results are comparable with those from the FibroSURE test. Changes in protein levels in biopsy samples associated with progression of fibrosis in patients.

Keywords

ALT; vascular cell adhesion molecule; alpha-2-macroglobulin; retinol binding protein 4; fibrogenesis

INTRODUCTION

Chronic hepatitis C (CHC) infection is characterized by varying degrees of inflammation and hepatic injury. Even in the era of direct acting antivirals, current standard-of-care includes Interferon (IFN)-based therapy that is often poorly tolerated, and not suitable for all patients¹. Limitations of liver biopsy for accurate disease staging are established^{2, 3}. The semi-quantitative histopathology grading systems in CHC do not reflect actual amount of extracellular matrix, but were initially developed to standardize and improve observer variability, and to determine thresholds for initiation of IFN-based therapy in CHC patients^{4, 5}. Despite limitations of biopsy as an imperfect reference, numerous non-invasive tests have been developed as an alternative to liver biopsy for assessing fibrosis stage prior to treatment initiation, and to follow disease progression in viral hepatitis⁶. These certainly appear to have clinical utility for detection of advanced stage chronic liver disease, and providing prognostic information. However, due to inherent variability in the reference standard, these non-invasive tests appear to have limited predictive utility in following longitudinal changes in fibrosis, along with higher misclassification rates for intermediate stages^{7, 8}.

The evolution of antiviral treatment regimens, with increased tolerability and efficacy, also now raises the question of whether the differentiation of significant fibrosis (METAVIR F2) retains clinical significance. The stratification of fibrosis into mild disease (F0 and F1), moderate-advanced disease (F2 and F3) and cirrhosis (F4) may have greater clinical utility

than the previously accepted broad dichotomous classification. Furthermore, the reliable non-invasive differentiation of adjacent stages of fibrosis would undoubtedly enhance the development of novel anti-fibrotic agents, although achievement of this goal continues to be hampered by the inherent limitations of biopsy.

The dynamic nature of fibrogenesis results in expression and secretion of various extracellular matrix (ECM) components, profibrogenic cytokines, growth factors, and enzymes involved in matrix synthesis, deposition and degradation. Modulation of the underlying process responsible for injury in viral hepatitis can result in regression of fibrosis, and even cirrhosis^{9–11}. However, following changes in fibrosis in relation to the natural history of underlying disease process (or in response to therapy) requires repeated assessment of hepatic injury through frequent liver biopsy sampling that is not clinically feasible. None of the currently available non-invasive approaches for fibrosis staging were designed to detect changes in matrix turnover. Recent developments in high throughput protein detection methods now allow us to explore a wide repertoire of potential biomarkers that are involved in various phases of the fibrogenesis cascade. Our hypothesis was that differing panels of markers would be required to (1) reliably differentiate adjacent fibrosis stages, and (2) accurately classify patients as mild/moderate/ severe fibrosis stage categories

Our specific aims in the first part of our study was to assess cross-sectional fibrosis in two large independent cohorts of CHC to determine if using multiplex detection methods for candidate markers in serum could differentiate a) three clinically relevant categories of fibrosis (mild-moderate-severe) and b) adjacent stages of fibrosis. Our aim in the second part of the study was to evaluate the utility of candidate marker panels for assessment of longitudinal changes in fibrogenesis in paired biopsies.

PATIENTS AND METHODS

Patients with chronic hepatitis C infection were identified from the Duke Hepatology Clinical Research Database and Biorepository. Three hundred and eighty three treatment naïve CHC patients (Duke HCV) were included in the study as a training cohort. Of these, 133 patients also had paired baseline and post-treatment biopsy and serum samples available for evaluation. The validation cohort included 434 CHC treatment naïve patients from the established prospective Trent HCV cohort¹². Serum from 100 normal controls was obtained from Bioreclamation Inc. (Jericho, NY). Samples for this study were identified from 2006 to 2009. All patients had provided prior consent and the study protocol approved by the Duke University and UK local research ethics committees.

Histological staging and grading

Liver biopsies had been independently evaluated by tertiary center expert hepatopathologists, and only patients with liver biopsies that were deemed of sufficient quality (based on both ≥ 5 evaluable portal tracts and $>1\text{cm}$) were included in the study. These were biorepository samples collected when biopsy specimens of at least 1cm and/or 5 portal tracts along with experienced histology assessment were adequate for development of activity grading in hepatitis C¹³, FibroTest¹⁴ and European Liver Fibrosis Group marker

panel¹⁵. Fibrosis was staged according to the METAVIR system and Histologic Activity Index (HAI) scores were categorized as mild (HAI 0–5), moderate (6–9) or severe (>10)⁴.

Biomarker Selection and Detection

Candidate biomarkers selected for evaluation in this study were in part based on expert opinions from an International Fibrosis Group Meeting¹⁶, and availability on the SearchLight assay platform (Supplementary Table 1). Selected biomarkers were assayed by laboratory staff without access to clinical or demographic data using a multiplex ultrasensitive SearchLight Chemiluminescent Protein Array platform (Pierce Biotechnology, c/o Thermo Fisher Scientific Inc., Rockford, IL) (see supplementary data for methodology). All serum samples were also independently evaluated for the fibrosis marker panel HCV FibroSURE (LabCorp, Burlington, NC).

Statistical Analysis

Several different analyses were performed in order to identify biomarkers with the potential to differentiate between fibrosis stages. Using ANOVA as an initial univariate analysis we compared each analyte in CHC patients to normal controls. Paired samples were available for a subset of the Duke cohort (n=133), and repeat measures analysis was used for this data. SAS 9.2 (SAS Institute, Cary, NC) was used for analysis. Our main analysis focused on multivariate modeling to identify panels of biomarkers from the original full set of 37 candidate biomarkers that could differentiate between mild (F0/F1) and moderate-to-severe (F2/F3/F4) fibrosis stages, mild-moderate-and-severe (F0/1-F2/3-F4) stages, as well as adjacent fibrosis stages (F0-F1-F2-F3-F4). A wrapper variable selection approach using randomForest (R software, www.R-project.org, randomForest Package) as the learning machine was used to build classification models¹⁷. Areas under the ROC curve (AUROC) or penalized Obuchowski measures (OB) were used as the optimization criteria to select the number of biomarkers p needed in the model¹⁸ (see supplementary data for additional details).

RESULTS

Baseline Demographics

Training cohort patients (n=383) were predominantly Caucasian (81%) and male (68%), with a median (25th–75th percentiles) age of 44(40–48) years, biopsy length 13(11–15) mm, and minimal stage (F0-F1) fibrosis (234/383, 61%). The baseline distribution of fibrosis stages for the training cohort was F0 (14%), F1 (47%), F2 (17%), F3 (12%) and F4 (10%). The validation cohort (n=434) included a comparable proportion of males (295/434, 68%) of median age 40 (34–48) years, with mostly minimal-mild stage disease: F0 (32%), F1 (34%), F2 (13%), F3 (4%) and F4 (17%).

Training Cohort Biomarker Selection

Thirty four biomarkers were evaluated in samples from the training cohort and healthy controls using the SearchLight multiplex platform. An additional three biomarkers were measured using standard ELISA methods (HA, CK18, RBP). Of these 37 biomarkers, 36 were differentially expressed in CHC patients compared to normal controls (supplementary

Table 1). Eighteen biomarkers were noted to have a significant linear trend with METAVIR fibrosis stage (Table 1).

Prognostic Classification of Fibrosis Stage

Classification of biopsies as mild (F0-1), moderate-advanced (F2-3) and severe (F4) disease stages may provide clinically useful prognostic information. Five markers (HA, VCAM1, A2M, Age, RBP4) were able to differentiate these three disease severity levels in the training cohort with a penalized Obuchowski (OB) measure of 0.85. The overall misclassification rate was 38% (146/383, F0-1=27/234, F2-3=88/112 and F4=31/37)(Figure 1a). The validation group had a penalized OB of 0.89, with an overall misclassification rate of 29% (125/434; F0-1=24/286, F2-3=52/74, and F4=49/74)(Figure 1b). Misclassification rates were highest for stage F2-3 disease (training =88/112 and validation=52/74) and the majority of these patients were misclassified as having stage F0-1 disease (81/88 and 47/52 respectively). There was no significant improvement with inclusion of up to nine additional markers (OB =0.845-0.85). In the training cohort, three stage classification performance for FibroSURE was OB=0.82 with a misclassification rate of 50%, even after excluding n=70 with FibroSURE index scores of 0.32–0.48 (corresponding to F1-F2).

Adjacent Fibrosis Stage Differentiation

RandomForest prediction modeling and application of standardized penalized Obuchowski (OB) measures for differentiating between any of the fibrosis stages F0-F4 indicated a six variable (GGT, PAI-1, PDGF-BB, Total Bilirubin, Gender, RBP4) OB= 0.89 and 0.84 for training and validation groups respectively, with a comparative FibroSURE OB=0.87. Additional random Forest models allowed for SearchLight marker selection for differentiating between F0-F1, F1-F2, F2-F3, and F3-4 in both the training and validation cohorts (Table 2).

Stage F0 and F1—In the training cohort, eleven variables (TIMP2, VCAM1, Age, PDGF-BB, M30, ALT, HA, TBIL, E-Selectin, A2M and TIMP-1) allowed for differentiation between stages F0 (n=54) and F1 (n=180) with AUROC of 0.51 (CI: 0.42–0.60). This was comparable to FibroSURE (AUROC 0.53, CI: 0.44–0.62), and there were no significant differences in the performance of these two marker panels. Similar poor performance characteristics were noted for the eleven marker panel in the validation cohort (F0=139, F1=147; AUROC= 0.54, CI: 0.475-0.61).

Stage F1 and F2—Eight markers (GGT, PDGF-BB, A2M, HA, MMP2, PAI-1, RBP4, ALT) were selected in the training cohort as having optimal performance for differentiating stage F1 (n=180) from stage F2 (n=65) with an AUROC of 0.605 (CI: 0.52–0.69), compared to an AUROC of 0.59 (CI: 0.51–0.66) for FibroSURE. No significant difference could be detected between the two different marker panels (p=0.64). Similar modest performance characteristics were seen in the validation cohort for these eight markers (F1=147, F2=57; AUROC 0.67; CI: 0.58–0.75), and there was no significant improvement in predictive performance with additional markers.

Stage F2 and F3—Eleven markers (VCAM1, Total bilirubin, HA, TIMP2, TIMP1, CK18, Haptoglobin, MMP2, RBP4, E-selectin, A2M) were selected in the training cohort for optimal differentiation of stages F2 (n=65) and F3 (n=47) with an AUROC of 0.69 (CI: 0.59–0.80) that was not significantly lower than FibroSURE (AUROC 0.72; CI: 0.62–0.81; p=0.59). Performance for these eleven markers in the validation cohort (F2=57, F3=17; AUROC 0.56; CI: 0.385–0.74) was not significantly lower compared to the training cohort (p= 0.24) and FibroSURE (p=0.13).

Stage F3 and F4—Only two variables (HA and PAI-1) were able to differentiate stages F3 (n=47) from F4 (n=37) in the training cohort with an AUROC of 0.51 (CI: 0.38–0.64). This predictive performance was comparable and not significantly different from FibroSURE (AUROC = 0.52; CI: 0.39–0.65; p-value=0.90). Performance in the validation cohort (F3=17, F4=74) for including two (HA or PAI-1) and upto six additional markers was marginally higher than the training cohort with an AUROC range of 0.66 to 0.70 (p-value=0.03 to 0.10)

Markers Associated with Longitudinal Assessment of Fibrosis Stage and Activity

One hundred and thirty three CHC patients in the training cohort had paired liver biopsies before and at least six months after antiviral therapy. Most patients had mild stage F1 fibrosis at baseline (69/133, 52%) and following treatment (78/133, 59%). There were marginal differences in the distribution of mild (0–5), moderate (6–9) or severe HAI (10–18) scores at baseline (28/129 (22%), 40/129 (31%) and 61/129 (47%); missing n=4) compared to follow-up biopsy scores (45/130 (35%), 43/130 (33%) and 42/130 (32%); missing n=3) (p=0.05).

The majority of patients (80/133, 60%) did not have any change (increase or decrease of one stage) in their fibrosis following treatment. As expected in a post-therapy CHC cohort, there was a decline in median HAI necroinflammatory score on follow-up biopsy (9 vs. 8; p = 0.001). Sixteen markers had significant differential expression in paired biopsies (Table 3). Change in fibrosis stage was not associated with change in FibroSURE index (p= 0.62). However, there were significant trend test associations between HAI change (HAI $\pm$ 2) and increase in both FibroSURE (p=0.006) and ActiTest (p=0.0003) (Supplementary Figure 1). Nine markers (ALT, GGT, HA, ICAM1, IL4, CXCL10, CXCL9, VCAM1) were associated with HAI change, and four (GMCSF, IL12, IL2, MMP13) with change in fibrosis stage (Table 4).

DISCUSSION

This large study included greater than 900 patient samples in training and validation cohorts, and is the first to prospectively evaluate the potential diagnostic utility of a varying combination of novel candidate serum biomarkers in differentiating between adjacent fibrosis stages in CHC infection. The role of these candidate biomarkers delineating adjacent METAVIR fibrosis stages appears limited, with AUROC values of 0.51–0.69. FibroSURE prediction of adjacent stages in our training cohort was also modest with AUROCs of 0.53–0.72. Performance measures for differentiating the more clinically relevant mild, moderate and severe disease stages were more acceptable with a weighted OB measure of 0.85–0.89.

However, selected markers were still associated with overall misclassification rates of 29–38%, especially for stages F2-3 where most patients were misclassified as having mild (F0-1) disease stage. In comparison, FibroSURE had a higher overall misclassification rate of 50%, with higher error rates of 59% for F2-3 and F4 compared to 45% for F0-1.

The diagnostic limitations of biopsy in a heterogeneous disease state due to sampling and observer errors are well-documented³. The semi-quantitative staging systems do not accurately measure extracellular matrix deposition that is non-linear and varies with advancing disease¹⁹. A significant barrier to the clinical development of anti-fibrotic therapy is the lack of alternatives to liver biopsy that allow the dynamic processes of fibrogenesis and fibrolysis to be measured accurately and at an early stage in relation to efficacy¹⁶. We selected serum biomarkers that would reflect a spectrum of histologic injury in CHC infection, for example including markers of hepatocyte injury and inflammation, fibrogenic signaling, immune activation, angiogenesis, and matrix remodeling. As such, our study is the first to try to develop and validate both direct and indirect candidate biomarkers to specifically differentiate adjacent stage disease. The poor performance characteristics for selected markers in this study likely relate to significant ECM and activity overlap between semi-quantitative measures of adjacent fibrosis stage. Biopsy quality, sampling error, and poor observer agreement for intermediate disease stages are additional factors limiting the accuracy of a single histological fibrosis stage as an outcome variable, thus further reducing the predictive performance of potential biomarkers⁷. Even with ideal performance parameters of 0.9 for biopsy sensitivity and specificity, and disease prevalence of 40%, a perfect marker would not exceed an AUROC of 0.9 for stage F2²⁰. This is likely to be much lower for differentiating adjacent stage disease. A prior study evaluated digitized virtual biopsy images, and noted that specimens 15mm in length had optimal adjacent stage AUROCs of 0.82 and 0.86 for F1 v F2 and F2 v F3 respectively³. In a pooled analysis of > 3500 CHC patients with FibroSURE and biopsy, observed AUROC for F1 and F2 was 0.66, with an adjusted value for 20% expected false positive or negative error rate for a 15mm biopsy increasing the relative AUROC to 0.80. Likewise, unadjusted observed AUROC for stages F0 v F1 (0.64), and F3 v F4 (0.69) were higher than observed in our training cohort with a smaller sample size for candidate markers or FibroSURE. However, F2 v F3 (0.66) was lower in the pooled analysis compared to our observed FibroSURE AUROC of 0.72²¹. Although the modest biopsy size and sampling error may in part explain lower performance characteristics in our study for our candidate markers and FibroSURE, pathologist experience and poor agreement in scoring of intermediate stage disease are of greater relevance²². The misclassification of biopsy, with greater than 25% false negative/positive rates even for delineating broad overlapping fibrosis stages, is further exacerbated for intermediate stage disease. Biopsy has a U-shape performance for intermediate stages, and relatively to biomarkers, biopsy is probably worse for intermediate stages, due to two reasons: an observer variability that is worse between F2 and F1²² and to the small difference in area of fibrosis between F2 and F1. This diagnostic inaccuracy represents an inherent limitation to biomarker studies such as ours that continue to depend upon biopsy staging as a reference standard. Laparoscopic and multiple sampling approaches may yield a larger specimen size and reduce sampling error, but are associated with increased risk and

cost²³, and in a prior virtual biopsy study, biopsy length >25mm did not significantly increase accuracy for METAVIR stage³.

Statistical approaches such as Bayesian latent class models for assessing accuracy of diagnostic tests with imperfect reference standards have confirmed the utility of FibroSURE but require further validation in our cohort^{24, 25}.

A recent study assessing multiple virtual biopsy sections and FibroSURE in a large chronic liver disease cohort indicated that weighted AUROC for F2 v F1 in CHC patients by FibroSURE was 0.505 and similar to a 30mm virtual biopsy⁸. For F1 v F2, our study indicated AUROC of 0.59–0.67 for training and validation cohorts respectively and 0.59 for FibroSURE. This suggests that even an optimal 30 mm biopsy has a low accuracy for discriminating F2 from F1 compared to a “true” reference, and therefore cannot be used to identify better biomarkers than the first generation identified based on smaller sample size¹⁴. We also evaluated for markers predictive of change in fibrosis on paired biopsy assessments in the training cohort. However, most patients did not have a significant change in fibrosis in this cohort. Although changes in histology was accompanied by markers such as GMCSF, IL-12, IL-2 and MMP-13, due to the low accuracy of even a 30 mm biopsy, where only two stages can be identified, our study highlights the constraints of using histology as the reference standard to detect meaningful directional change.

The pathogenesis of hepatic fibrosis is complex and extends beyond pathways related to hepatic stellate cell (HSC) activation that result in initiation, perpetuation and resolution of hepatic injury. Other pathways of fibrogenesis include non-HSC derived ECM, chemokine and adipokine signaling, innate immunity, ECM stiffness, genetic regulation of HSC and myofibroblasts, along with resolution through apoptosis and senescence²⁶. The chronic inflammatory nature of HCV infection is likely to involve a variable combination of these pathogenic factors depending on disease severity, immune response and other host-viral interactions. Translating this complex and dynamic process into developing biomarkers of clinical relevance is difficult. In our study, certain adipose derived (adiponectin and leptin), pro-inflammatory (IL1, IL2, IL6, IL12) or anti-inflammatory (IL4, IL10) cytokines showed no significant association with fibrosis stage in our CHC training cohort. However, other adipose derived factors (PAI-1), cell adhesion molecules (ICAM1, E-Selectin, VCAM1), and apoptosis markers (CK18 M30) were differentially expressed in these patients. Despite the limitations of evaluating complex biomarkers in this cross-sectional study, there were interesting differences in marker selection depending on disease severity. Certain markers such as VCAM1, E-Selectin and TIMP1 and -2 were selected in a predictive model for minimal and moderate stages but not for F3 v F4, perhaps reflecting differences in immune activation and healing response compared to more advanced stages. Hyaluronic acid was consistently selected across all adjacent stage models, reflecting its long established role as a direct marker of fibrosis, and inclusion in several predictive serum marker algorithms⁶. Indirect markers including A2M, ALT, bilirubin, and GGT, which comprise several proprietary marker algorithms such as FibroSURE, were also selected in the multiplex array, indicating their diagnostic utility across the spectrum of fibrosis. Certainly a broader range of both direct and indirect markers were selected when considering adjacent stage

differentiation for moderate-advanced disease (F2-3). This indicates more complex inflammatory signaling, fibrogenesis and matrix remodeling with disease progression.

Recent advances in imaging methods for assessing liver stiffness have provided an alternative to serum-based tests for estimating fibrosis stage, and certainly appear to complement the diagnostic role of serum biomarkers⁶. However, variable stiffness thresholds depending on etiology of chronic disease, and significant overlap for the stiffness parameter across non-linear categorical measures of fibrosis, also appear to limit their utility in differentiating adjacent stage disease. Biopsy image analysis and measurement of collagen proportionate area could provide a more objective outcome measure of fibrosis for non-invasive tests, but are not included in routine histologic assessment²⁷, and are limited by sampling variability³. However, given the limitations of biopsy for accurate classification of individual stages, we evaluated the diagnostic potential of our biomarker array to differentiate mild (F0/1), moderate-advanced (F2-F3) and severe (F4) disease stages. This extends the role of biopsy (and non-invasive tests) to provide prognostic information, that may be perhaps more clinically relevant for treatment decisions in the current era of direct acting antiviral therapies in CHC infection^{28, 29}. As expected, by removing the criteria of predicting the more stringent adjacent disease stages, both FibroSURE and the selected multiplex variables (HA, VCAM1, A2M, RBP4) and age were able to differentiate overlapping disease stages with improved performance characteristics. Further validation will be required to assess the potential role of these multiplex markers in providing a prognostic group classification in clinical practice, perhaps in combination with elastography to reduce misclassification rates.

In summary, although multiplex ELISA platforms now allow for rapid targeted protein quantitation, diagnostic performance characteristics of our customized array of biomarkers for adjacent fibrosis stage differentiation, or for assessment of changes in histology in CHC infection was relatively poor, and comparable to simple marker panels such as FibroSURE. Although assessment of disease progression to advanced stages is clinically relevant, there is unlikely to be any utility of assessing complex and varying combination of multiplex markers in the routine clinical setting for differentiating adjacent stage disease. Due to the imperfect reference standard it is possible that our multiplex panel or other biomarkers could have better diagnostic accuracy than expected. Validated biomarkers such as FibroSURE have improved performance when the size of biopsy is optimal and demonstrate better prognostic assessment compared to histology. Future biomarkers should use optimal reference standards, include quantitative histology assessment, and consider new analytic methods without a gold standard. However, in clinical practice with evolving therapeutic options in CHC infection, differentiation of mild-moderate-advanced stages is still relevant, but serum based tests alone are still likely to be associated with significant misclassification rates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: Supported by GlaxoSmithKline, RTP, North Carolina, the National Institute for Health Research (NIHR) Biomedical Research Unit in Gastrointestinal and Liver Diseases at Nottingham University Hospitals NHS Trust and the University of Nottingham, and in part by Duke University's CTSA grant 1 UL1 RR024128-01 from NCRR/NIH

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

1. Talal AH, Lafleur J, Hoop R, et al. Absolute and relative contraindications to pegylated-interferon or ribavirin in the US general patient population with chronic hepatitis C: results from a US database of over 45 000 HCV-infected, evaluated patients. *Aliment Pharmacol Ther.* 2013; 37(4): 473–81. [PubMed: 23289640]
2. Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. *Hepatology.* 2009; 49(3):1017–44. [PubMed: 19243014]
3. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology.* 2003; 38(6):1449–57. [PubMed: 14647056]
4. The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology.* 1994; 20(1 Pt 1):15–20. [PubMed: 8020885]
5. Consensus statement. *J Hepatol*; EASL International Consensus Conference on hepatitis C; Paris. 26–27 February 1999; 1999. p. 3-8.
6. Castera L. Noninvasive methods to assess liver disease in patients with hepatitis B or C. *Gastroenterology.* 2012; 142(6):1293–302. [PubMed: 22537436]
7. Cales P, de Ledinghen V, Halfon P, et al. Evaluating the accuracy and increasing the reliable diagnosis rate of blood tests for liver fibrosis in chronic hepatitis C. *Liver Int.* 2008; 28(10):1352–62. [PubMed: 18492022]
8. Poynard T, Lenaour G, Vaillant JC, et al. Liver biopsy analysis has a low level of performance for diagnosis of intermediate stages of fibrosis. *Clin Gastroenterol Hepatol.* 2012; 10(6):657–63. [PubMed: 22343514]
9. Ellis EL, Mann DA. Clinical evidence for the regression of liver fibrosis. *J Hepatol.* 2012; 56(5): 1171–80. [PubMed: 22245903]
10. Marcellin P, Gane E, Buti M, et al. Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. *Lancet.* 2012; 381(9865):468–75. [PubMed: 23234725]
11. Poynard T, McHutchison J, Manns M, et al. Impact of pegylated interferon alfa-2b and ribavirin on liver fibrosis in patients with chronic hepatitis C. *Gastroenterology.* 2002; 122(5):1303–13. [PubMed: 11984517]
12. Mohsen AH. The epidemiology of hepatitis C in a UK health regional population of 5.12 million. *Gut.* 2001; 48(5):707–13. [PubMed: 11302973]
13. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology.* 1996; 24(2):289–93. [PubMed: 8690394]
14. Imbert-Bismut F, Ratziu V, Pieroni L, et al. Biochemical markers of liver fibrosis in patients with hepatitis C virus infection: a prospective study. *Lancet.* 2001; 357(9262):1069–75. [PubMed: 11297957]
15. Rosenberg WM, Voelker M, Thiel R, et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology.* 2004; 127(6):1704–13. [PubMed: 15578508]
16. McHutchison J, Poynard T, Afdhal N. Fibrosis as an end point for clinical trials in liver disease: a report of the international fibrosis group. *Clin Gastroenterol Hepatol.* 2006; 4(10):1214–20. [PubMed: 16979947]

17. Svetnik V, Liaw A, Tong C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. 2003; 43(6):1947–58. [PubMed: 14632445]
18. Lambert J, Halfon P, Penaranda G, et al. How to measure the diagnostic accuracy of noninvasive liver fibrosis indices: the area under the ROC curve revisited. *Clin Chem*. 2008; 54(8):1372–8. [PubMed: 18539647]
19. Standish RA, Cholongitas E, Dhillon A, et al. An appraisal of the histopathological assessment of liver fibrosis. *Gut*. 2006; 55(4):569–78. [PubMed: 16531536]
20. Mehta SH, Lau B, Afdhal NH, et al. Exceeding the limits of liver histology markers. *J Hepatol*. 2009; 50(1):36–41. [PubMed: 19012989]
21. Poynard T, Morra R, Halfon P, et al. Meta-analyses of FibroTest diagnostic value in chronic liver disease. *BMC Gastroenterol*. 2007; 7:40. [PubMed: 17937811]
22. Rousselet MC, Michalak S, Dupre F, et al. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology*. 2005; 41(2):257–64. [PubMed: 15660389]
23. 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(10):2614–8. [PubMed: 12385448]
24. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995; 141(3):263–72. [PubMed: 7840100]
25. Poynard T, de Ledinghen V, Zarski JP, et al. Relative performances of FibroTest, Fibroscan, and biopsy for the assessment of the stage of liver fibrosis in patients with chronic hepatitis C a step toward the truth in the absence of a gold standard. *J Hepatol*. 2012; 56(3):541–8. [PubMed: 21889468]
26. Lee UE, Friedman SL. Mechanisms of hepatic fibrogenesis. *Best Pract Res Clin Gastroenterol*. 2011; 25(2):195–206. [PubMed: 21497738]
27. Goodman ZD, Becker RL Jr, Pockros PJ, et al. Progression of fibrosis in advanced chronic hepatitis C: evaluation by morphometric image analysis. *Hepatology*. 2007; 45(4):886–94. [PubMed: 17393526]
28. Parkes J, Roderick P, Harris S, et al. Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut*. 2010; 59(9):1245–51. [PubMed: 20675693]
29. Ngo Y, Munteanu M, Messous D, et al. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. *Clin Chem*. 2006; 52(10):1887–96. [PubMed: 16931569]

FIGURE 1. Prognostic Classification of Fibrosis Stage in (a) Training cohort (b) Validation Cohort

Fig 1(A) Overall 146 (38%) patients were misclassified in the training cohort using five selected markers. For F0-1, 25/27 (93%) patients were classified as moderate, and 2/27 (7%) as severe disease; F2-3, 81/88 (92%) classified as mild and 7/88 (8%) as severe disease; and for F4, 13/31 (42%) patients were classified as mild and 18/31 (58%) as moderate disease. Fig 1(B) 125 (29%) patients were misclassified in the validation cohort. For F0-1, 20/24 (83%) were classified as moderate and 4/24 (17%) as severe disease; F2-3 47/52 (90%) were classified as mild and 5/52 (10%) as severe disease; and for F4, 15/49 (31%) were predicted to have mild, and 34/49 (69%) moderate stage disease.

Table 1

Biomarkers Associated with METAVIR Fibrosis Stage in Training Cohort

INCREASED	p-value	DECREASED	p-value
A2M (g/L)	< 0.0001	HAPT (g/L)	0.0173
ALT (IU/L)	0.0033	MMP1	0.0209
E-Selectin (pg/mL)	< 0.0001	PAI-1 total (pg/mL)	0.0393
GGT (IU/L)	< 0.0001	PDGF-BB (pg/mL)	0.0301
HA (ng/mL)	< 0.0001	RBP4 (ng/mL)	< 0.0001
HGH pg/ml	0.0011		
ICAM1 (pg/mL)	0.0003		
CK18(M30) (U/L)	0.0009		
MMP2 (pg/mL)	< 0.0001		
TBIL (umol/L)	0.0001		
TIMP1 (pg/mL)	0.01		
TIMP2 (pg/mL)	0.0012		
VCAM1 (pg/mL)	< 0.0001		

Tukey's trend test was used to test for linear trend between METAVIR stages. Biomarkers in the table above have a significant linear increasing or decreasing trend between F0 to F4 (p-value < 0.05).

HA, Hyaluronic acid; A2M, Alpha-2 macroglobulin; VCAM1, vascular cell adhesion molecule-1; GGT, Gamma glutamyltransferase; RBP4, Retinol binding protein-4; TIMP2, Tissue inhibitor of metalloproteinase-2; MMP, Matrix metalloproteinase; IL8, Interleukin 8; PDGF-BB, Human Platelet Derived Growth Factor-BB; ALT, Alanine aminotransferase; TBili, Total Bilirubin; CK18 (M30), Cytokeratin-18 fragment M30; PAI-1, Plasminogen activator inhibitor; HGH, Human growth hormone; ICAM1, Intercellular adhesion molecule-1; HAPT, Haptoglobin;

Table 2

Biomarker Performance for Adjacent Stage Differentiation

	F0 v. F1 ^a			F1 v. F2 ^a			F2 v. F3 ^b			F3 v. F4 ^b			F0F1 vs. F2F3F4 ^b		
Selected markers	TIMP2, VCAM1, Age, PDGF-BB, CK18, ALT, HA, TBIL, E-Selectin, A2M, TIMP1			GGT, PDGF-BB, A2M, HA, MMP2, PAI-1, RBP4, ALT			VCAM1, Total bilirubin, HA, TIMP2, TIMP1, CK18, Haptoglobin, MMP2, RBP4, E-selectin, A2M			HA, PAI-1			HA, A2M, VCAM1, GGT, RBP4, TIMP2, E-Selectin, MMP-2		
TRAINING															
AUROC (95% CI)	0.51 (0.42–0.60)			0.605 (0.52–0.69)			0.69 (0.59–0.80)			0.51 (0.38–0.64)			0.75 (0.70–0.80)		
Sensitivity	0.21	0.32	0.37	0.29	0.52	0.63	0.81	0.70	0.62	0.81	0.70	0.62	0.80	0.70	0.60
Specificity	0.81	0.70	0.61	0.80	0.70	0.60	0.48	0.55	0.68	0.28	0.40	0.43	0.52	0.67	0.78
VALIDATION															
AUROC (95% CI)	0.54 (0.475–0.61)			0.67 (0.58–0.75)			0.56 (0.385–0.74)			0.66 (0.54–0.79)			0.82 (0.775–0.86)		
Sensitivity	0.28	0.37	0.48	0.39	0.61	0.74	0.82	0.71	0.65	0.81	0.70	0.61	0.80	0.70	0.61
Specificity	0.81	0.70	0.60	0.80	0.70	0.60	0.16	0.39	0.39	0.35	0.53	0.65	0.65	0.78	0.87
FibroSURE (TRAINING)															
AUROC (95% CI)	0.53 (0.44–0.62)			0.59 (0.51–0.66)			0.72 (0.62–0.81)			0.52 (0.39–0.65)			0.70 (0.64–0.75)		
Sensitivity	0.16	0.42	0.48	0.28	0.37	0.48	0.81	0.70	0.62	0.81	0.70	0.65	0.80	0.70	0.60
Specificity	0.81	0.70	0.63	0.80	0.70	0.61	0.49	0.65	0.68	0.19	0.36	0.40	0.49	0.58	0.69

^aSensitivity assessed for varying specificity thresholds of 0.8–0.6 for earlier stage disease.

^bSpecificity assessed for varying sensitivity thresholds of 0.8–0.6 for late stage disease and F0-1 v. F2-4

Table 3

Biomarkers with Differential Expression in Paired Biopsy Cohort

Biomarker	MFC	MFC (95% CI)	p-Value
ALT	-2.1	(-2.45, -1.8)	<0.0001
CK18	-1.36	(-1.49, -1.23)	<0.0001
IL4	-1.27	(-1.43, -1.13)	<0.0001
CXCL9	-1.21	(-1.32, -1.11)	<0.0001
VCAM 1	-1.16	(-1.25, -1.08)	<0.0001
IL8	-1.15	(-1.23, -1.07)	0.0002
IFNγ	-1.16	(-1.26, -1.07)	0.0004
E-Selectin	-1.14	(-1.23, -1.06)	0.0005
IL10	-1.18	(-1.3, -1.07)	0.0009
CXCL10	-1.15	(-1.26, -1.05)	0.0029
MMP1	-1.08	(-1.13, -1.03)	0.0027
ICAM1	-1.12	(-1.22, -1.04)	0.0053
TNFα	-1.16	(-1.29, -1.04)	0.0074
PDGF-BB	-1.08	(-1.13, -1.02)	0.0069
RBP4	1.11	(1.03, 1.2)	0.0083
CRP	-1.27	(-1.56, -1.04)	0.0204

MFC - Mean Fold Change Post Treatment vs. Pre Treatment

Table 4

Markers associated with Change in HAI or Fibrosis Stage in Paired Biopsies from Training cohort.

Type	Biomarker	Williams Statistic	p-value	Marker Direction
HAI Change	ALT (IU/L)	3.5762	0.0003	Increase
HAI Change	GGT (IU/L)	3.4724	0.0004	Increase
HAI Change	HA (ng/ ml)	1.7823	0.0470	Increase
HAI Change	ICAM1	2.3291	0.0124	Increase
HAI Change	IL4	1.7597	0.0493	Increase
HAI Change	CXCL10	2.5376	0.0071	Increase
HAI Change	CXCL9	2.8528	0.0028	Increase
HAI Change	PAI-1	1.9754	0.0302	Decrease
HAI Change	VCAM1 (pg/ ml)	2.9859	0.0019	Increase
Fibrosis Change	GMCSF (pg/ml)	1.8325	0.0419	Decrease
Fibrosis Change	IL12p70 (pg/ml)	3.0888	0.0013	Decrease
Fibrosis Change	IL2 (pg/ml)	2.3697	0.0111	Decrease
Fibrosis Change	MMP13 (pg/ml)	2.0962	0.0225	Decrease

HAI change indicates ± 2 Histologic necroinflammatory activity score;

Fibrosis change indicates ± 1 METAVIR fibrosis stage