Title: Development and Validation of Hepamet Fibrosis Scoring System—a Simple, Non-invasive Test to Identify Patients With Nonalcoholic Fatty liver Disease With Advanced Fibrosis

Short title: Hepamet Fibrosis Score detects fibrosis in NAFLD

Authors: Javier Ampuero[1], Raluca Pais[5], Rocío Aller[2], Rocío Gallego-Durán[1], Javier Crespo[3], Carmelo García-Monzón[4], Jerome Boursier[6], Eduardo Vilar[7], Salvatore Petta[8], Zheng Ming-Hua[9], Desamparados Escudero[10], Jose Luis Calleja[11], Patricia Aspichueta[12], Moisés Diago[13], Jose Miguel Rosales[14], Joan Caballería[15], Judith Gómez-Camarero[16], Oreste Lo Iacono[17], Salvador Benlloch[18], Agustín Albillos[19], Juan Turnes[20], Jesus M Banales[21], Vlad Ratziu[5], Manuel Romero-Gómez[1]; on behalf of the HEPAmet Registry.

List of collaborators of HEPAmet Registry: Salvador Agustin[22], Francisco Jorquera[23], Ruben Frances[24], Javier Garcia-Samaniego[25], Javier Salmeron[26], Conrado Fernandez-Rodriguez[27], Pamela Estevez[28], Raul Andrade[29], German Soriano[30], Miguel Fernandez-Bermejo[31], María Teresa Arias Loste[3], Rebeca Sigüenza[2], Aurora Giannetti[8], Elvira del Pozo Maroto[4].

Affiliations:

[1] Hospital Universitario Virgen del Rocio, Sevilla. Instituto de Biomedicina de Sevilla. CIBERehd, Spain.

[2] Hospital Clínico Universitario de Valladolid. Centro de Investigación de Endocrinología y Nutrición. Universidad de Valladolid, Spain.

[3] Hospital Universitario Marqués de Valdecilla, Santander, Spain.

[4] Liver Research Unit Hospital Universitario Santa Cristina Instituto de Investigación Sanitaria Princesa Madrid, Spain.

[5] Université Pierre et Marie Curie Hopital Pitie Salpetriere, Paris, France.

[6] Hepatology Department, University Hospital, Angers, France.

[7] Division of Gastroenterology and Hepatology, Indiana University School of Medicine, IN, USA.

This is the author's manuscript of the article published in final edited form as:
Ampuero, J., Pais, R., Aller, R., Gallego-Durán, R., Crespo, J., García-Monzón, C., … del Pozo Maroto, E. (2019). Development and Validation of Hepamet Fibrosis Scoring System—a Simple, Non-invasive Test to Identify Patients With Nonalcoholic Fatty liver Disease With Advanced Fibrosis. Clinical Gastroenterology and Hepatology. https://doi.org/10.1016/j.cgh.2019.05.051

[8] Section of Gastroenterology and Hepatology Di.Bi.M.I.S., University of Palermo, Palermo, Italia.

[9] NAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of Wenzhou Medical University, China. Institute of Hepatology, Wenzhou Medical University, China.

[10] Hospital Clínico de Valencia, Spain.

[11] Hospital Universitario Puerta de Hierro, Madrid, Spain.

[12] Biocruces Research Institute, Barakaldo. Department of Physiology, Faculty of Medicine and Nursing, University of Basque Country UPV/EHU, Leioa, Spain.

[13] Hospital General Universitario de Valencia, Spain.

[14] Agencia Sanitaria Costa del Sol, Marbella, Spain.

[15] Liver Unit. Hospital Clínic. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBPAS). CIBERehd. Barcelona, Spain.

[16] Hospital Universitario de Burgos, Spain.

[17] Hospital Universitario Tajo, Aranjuez, Spain.

[18] Hospital Universitari i Politecnic La Fe, Valencia. CIBERehd, Spain.

[19] Hospital Universitario Ramón y Cajal, Madrid, Spain.

[20] Complejo Hospitalario de Pontevedra, Spain.

[21]Department of Liver and Gastrointestinal Diseases, Biodonostia Research Institute – Donostia University Hospital –, University of the Basque Country (UPV/EHU), Ikerbasque, CIBERehd, San Sebastian, Spain.

[22] Hospital Vall d'Hebrón, Barcelona, Spain

[23] Hospital Universitario de León, Spain.

[24] Hospital General Universitario de Alicante. Universidad Miguel Hernández. CIBERehd, Spain.

[25] Hospital Universitario La Paz. CIBERehd. IdiPAZ. Madrid, Spain.

[26] Hospital Universitario San Cecilio, Granada, Spain.

[27] Hospital Universitario Fundación de Alcorcón, Universidad Rey Juan Carlos, Spain.

[28] Complejo Hospitalario Universitario de Vigo, Spain.

[29] Unidad de Gestión Clínica de Enfermedades Digestivas, Instituto de Investigación Biomédica de Málaga-IBIMA, Hospital Universitario Virgen de la Victoria, Universidad de Málaga, CIBERehd, Málaga, Spain.

[30] Hospital de la Santa Creu i San Pau, Barcelona, Spain.

[31] Hospital San Pedro de Alcantara, Caceres, Spain.

Correspondence to:

Javier Ampuero MD, PhD

Digestive Disease Department and CIBERehd, Virgen del Rocio University Hospital,

Avenida Manuel Siurot s/n, 41013 Sevilla, Spain

E-mail: jampuero-ibis@us.es

Phone: (+34) 955-015761 Fax: (+34) 955-015899

Prof. Manuel Romero-Gómez, MD, PhD

Digestive Disease Department and CIBERehd, Virgen del Rocio University Hospital,

Avenida Manuel Siurot s/n, 41013 Sevilla, Spain

E-mail: mromerogomez@us.es

Phone: (+34) 955-015761 Fax: (+34) 955-015899

ELECTRONIC WORD COUNT: 3,250 words

TABLES: 3

FIGURES: 3

SUPPLEMENTARY MATERIAL: 5 Tables and 6 Figures.

AUTHOR STATEMENT:

Guarantor of the article: JA, MRG

Study design: JA, MRG

Drafting the manuscript: JA, MRG

Statistical analyses and interpretation: JA, MRG

Data acquisition and critical review of the manuscript: All authors.

All authors approved the final version of the article, including the authorship list.

EXTERNAL FINANCIAL SUPPORT

This project has been partially funded by the "Consejería de Salud de la Junta de Andalucía" (PI-0075-2014) and "Spanish Ministry of Economy, Innovation and Competition, *Instituto de Salud Carlos III"* (PI16/01842).

*The founders has not had any role in the design, analysis, writing or interpretation of this project.

POTENTIAL CONFLICT OF INTEREST

None.

ABBREVIATIONS (in alphabetical order)

AUROC: Area under the receiver operating characteristic curve

BMI: Body mass index

CI: Confidence interval

DM: Diabetes mellitus

DOR: Diagnostic odds ratio

HFS: Hepamet Fibrosis Score

HOMA: Homeostatic model assessment

IDI: Integrated discrimination improvement

NAFLD: Non-alcoholic fatty liver disease

NASH: Non-alcoholic steatohepatitis

NFS: NAFLD fibrosis score

NPV: Negative predictive value

NRI: Net reclassification improvement

OR: Odds ratio

PPV: Positive predictive value

Abstract:

Background & Aims: Fibrosis affects prognoses for patients with nonalcoholic fatty liver disease (NAFLD). Several non-invasive scoring systems have aimed to identify patients at risk for advanced fibrosis, but inconclusive results and variations in features of patients (diabetes, obesity and older age) reduce their diagnostic accuracy. We sought to develop a scoring system based on serum markers to identify patients with NAFLD at risk for advanced fibrosis.

Methods: We collected data from 2452 patients with NAFLD at medical centers in Italy, France, Cuba, and China. We developed the Hepamet fibrosis scoring system using demographic, anthropometric, and laboratory test data, collected at time of liver biopsy, from a training cohort of patients from Spain (n=768) and validated the system using patients from Cuba (n=344), Italy (n=288), France (n=830), and China (n=232). Hepamet fibrosis score (HFS) were compared with those of previously developed fibrosis scoring systems (the NAFLD fibrosis score [NFS] and FIB-4). The diagnostic accuracy of the Hepamet fibrosis scoring system was assessed based on area under the receiver operating characteristic (AUROC) curve, sensitivity, specificity, diagnostic odds ratio, and positive and negative predictive values and likelihood ratios.

Results: Variables used to determine HFS were patient sex, age, homeostatic model assessment score, presence of diabetes, levels of aspartate aminotransferase, and albumin, and platelet counts; these were independently associated with advanced fibrosis. HFS discriminated between patients with and without advanced fibrosis with an AUROC curve value of 0.85 whereas NFS or FIB-4 did so with AUROC values of 0.80 (P=.0001). In the validation set, cut-off HFS of 0.12 and 0.47 identified patients with and without advanced fibrosis with 97.2% specificity, 74% sensitivity, a 92% negative predictive value, a 76.3% positive predictive value, a 13.22 positive likelihood ratio, and a 0.31 negative likelihood ratio. HFS were not affected by patient age, body mass index, hypertransaminasemia, or diabetes. The Hepamet fibrosis scoring system had the greatest net benefit in identifying patients who should undergo liver biopsy analysis and led to significant improvements in reclassification, reducing the number of patients with undetermined results to 20% from 30% for the FIB-4 and NFS systems ($P<.05$).

Conclusions: Using clinical and laboratory data from patients with NAFLD, we developed and validated the Hepamet fibrosis scoring system, which identified patients with advanced fibrosis with greater accuracy than the FIB-4 and NFS systems. the Hepamet system provides a greater net benefit for the decision-making process to identify patients who should undergo liver biopsy analysis.

KEY WORDS: HOMA, steatosis, prognostic factor, diagnostic tool, cirrhosis

Need to Know

Background: Non-invasive scoring systems are needed to detect and monitor liver fibrosis in patients with NAFLD because the reliability of liver biopsy analysis is limited. Previously developed systems (the NFS and FIB-4 systems) have limited accuracy in identifying patients with advanced fibrosis. Their scores are affected by patient body mass index and age, requiring adjusted cut-off values to increase their specificity.

Findings: We developed a scoring system, called the Hepamet fibrosis scoring system, based on clinical and laboratory test results. This system identified patients with NAFLD who had advanced fibrosis with a high level of specificity, and did not require adjustment of cut-off scores to increase its accuracy or the number of patients correctly classified. Hepamet fibrosis scores identified patients with advanced fibrosis with higher levels of accuracy than the NFS and FIB-4 systems in an independent validation cohort.

Implications for patient care: The Hepamet fibrosis scoring system can be used in primary care to identify patients with fatty liver disease at highest risk for advanced fibrosis and reduce unnecessary referrals and in specialized units to increase detection of advanced fibrosis.

INTRODUCTION

The burden of non-alcoholic fatty liver disease (NAFLD) has been dramatically growing in parallel with obesity, diabetes, and metabolic syndrome outbreaks[1]. NAFLD has become the most common cause of chronic liver disease, representing a risk factor for cirrhosis, hepatocellular carcinoma, and liver transplantation[2], as well as for extra-hepatic manifestations such as cardiovascular[34] and kidney disease[5], and extrahepatic malignancies[6]. Fibrosis has been identified as the major determinant of the long-term prognosis of NAFLD patients[7]. In the current scenario, the correct identification of patients at risk of progression is a critical step in the management of NAFLD[8]. No symptoms and normal transaminase levels are common features of NAFLD. Thus, we need to develop tools able to detect this silent entity. Liver biopsy has been considered the gold standard for the diagnosis of NAFLD, although it is sometimes imperfect due to sample-to-sample variability and interpretation, and some additional concerns such as the cost and potential complications. Several algorithms based on serological biomarkers have been developed to identify patients at risk of advanced fibrosis. Both NAFLD fibrosis score (NFS)[9] and FIB-4 index[10] are the serological non-invasive methods most widely used to exclude the presence of advanced fibrosis. However, they have shown some limits such as the influence of baseline variables included in the formula to calculate the score (i.e., age[11] in FIB-4 and obesity in NFS[12]). Moreover, non-interpretable results (so-called grey zone) could reach up to 30% of patients[13] in these tests.

The identification of NAFLD patients at risk of liver fibrosis progression is a critical unmet need representing a timely challenge for clinicians. In this study, we developed a serum-based non-invasive score to improve the prediction of advanced fibrosis and further diagnostic decision-making process in patients with NAFLD.

METHODS

Selection of Patients

An international multicenter cross-sectional study was designed including 2,452 consecutive biopsy-proven NAFLD patients. The research was initially conducted with patients from the Spanish HEPAmet Registry. This registry is governed by the Spanish Association for the Study of the Liver (AEEH) and the Network of Biomedical Research Centre for the Study of the Liver and Digestive Diseases (CIBERehd). Monitoring is a fundamental element of the database, ensuring the accuracy of data and minimization of bias. The study was later externally validated in biopsy-proven NAFLD patients from geographically separate tertiary international medical centers from Italy, France (two independent hospitals), Cuba, and China.

Patients underwent a liver biopsy according to the routine decisions in the clinical practice. The inclusion criterium was biopsy-proven NAFLD, irrespective of the existence of NASH or fibrosis stage. Exclusion criteria were significant alcohol intake (>30 g daily for men and >20g daily for women) and evidence of concomitant liver disease (i.e., viral or autoimmune hepatitis, HIV, drug-induced fatty liver, hemochromatosis or Wilson's disease). The study was performed in agreement with the Declaration of Helsinki and with local and national laws and approved by the Ethics and Clinical Research Committee of every center. All patients were informed of the nature of the study and gave their written consent to participate.

Clinical assessment

Demographic characteristics, anthropometric measures, and laboratory tests (ALT, AST, GGT, triglycerides, cholesterol, HDL-c, LDL-c, fasting glucose, HbA1c, insulin, creatinine, albumin) were recorded at the same time of liver biopsy. A fasting blood sample was taken for routine biochemical analyses. HOMA was calculated based on insulin and glucose (fasting

insulin x fasting glucose / 405). Furthermore, NAFLD Fibrosis Score[9] and FIB-4[10,14] were computed.

Histological assessment

The diagnosis of NAFLD was based on histological criteria. All liver biopsies were assessed by experienced hepato-pathologists, who were blinded regarding patient's evaluation and clinical data. Samples of <15 mm length or <10 portal tracts were considered not suitable for diagnosis and fibrosis staging and were excluded. To define steatohepatitis, we used SAF scoring system[15] combining steatosis, inflammatory activity, and fibrosis. Several histological aspects were measured. First, steatosis was rated as 1 (5%-33%), 2 (33%-66%) and 3 (>66%). Second, activity grade is the addition of hepatocyte ballooning (0–2) and lobular inflammation (0–2). Lastly, liver fibrosis was taken into account the fibrosis shown in zone 3 perisinusoidal: F0 (none portal fibrosis), F1 (some-most portal fibrosis), F2 (few bridging fibrosis), F3 (much-bridging fibrosis), F4 (cirrhosis). We defined advanced fibrosis (F0-2 *vs.* F3-F4) for statistical purposes.

Objectives

We aimed to develop a serological non-invasive score (based on standard variables) to predict fibrosis in patients with NAFLD, for the following purposes: a) to improve the advanced fibrosis screening compared to the most used non-invasive methods (NFS and FIB-4); b) to assess the effectiveness of the score to predict advanced fibrosis in presence of baseline conditions that could bias the results (age, BMI, diabetes, and hypertransaminasemia); c) to assess the health outcomes of the implementation of the score on the diagnostic decision-making process.

Statistical analyses

Variables used for the Hepamet Fibrosis Score were measured at enrolment. To develop and validate our model, we drew two independent cohorts of 758 subjects for model development (Spanish cohort) and 1,694 individuals for model validation [French No.1 (n=444), French No.2 (n=386), Italian (n=288), Cuban (n=344), and Chinese (n=232) cohorts]. Data were reported as the mean ± standard deviation for normal and median (interquartile range) for non-normal continuous variables, while frequency was used for discrete variables. In the univariable comparisons, we used the Student t-test and ANOVA with Bonferroni adjustments for continuous samples and chi-square test or Fisher's exact test for qualitative ones. Non-parametric alternatives (Mann–Whitney U and Kruskal-Wallis tests) were used for non-normal distributions. Independent variables with significance $p \leq 0.10$ were introduced in a first multivariable analysis (backward Wald logistic regression analysis) to identify factors independently related to advanced fibrosis. To improve the prediction, a second multivariable analysis was performed after the transformation of the continuous variables into qualitative and ordinal ones according to the thresholds corresponding to a fourth and a two-times higher prevalence for advanced fibrosis (**Supplementary Figure 1**). Odds ratios (OR) and their 95% confidence intervals were estimated. Values were considered to be statistically significant when $p<0.05$. Akaike's information criterion, which is an estimator of the relative quality of statistical models for a given set of data, was additionally computed to select the most robust predictors.

The calibration of the Hepamet Fibrosis Score was assessed using a calibration belt[16]. It creates a confidence band for the calibration curve based on a function that relates expected to observed probabilities of advanced fibrosis across classes of risk. The calibration belt identifies significant deviations from the ideal calibration, as well as the direction of the variation. The area under the ROC curve (AUROC) was computed to corroborate the results observed in the derivation and validation sets, determine the diagnostic accuracy of the predictive models and select different thresholds for predicting advanced fibrosis. Youden

Index (sensitivity + specificity – 1)[17] was calculated to identify the optimal lower cut-off, and the higher cut-off was determined to show 97% of specificity. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), percent correctly classified, likelihood ratios and diagnostic odds ratio (OR) were computed for the selected cutoffs, as well as the post-tests probabilities. We presented a decision curve analysis to evaluate (net benefit) whether the application of the prediction model does more good (identification of advanced fibrosis) than harm (unnecessary biopsy). The selected probability thresholds represented the level of diagnostic certainty above which the patient would choose to be biopsied. The highest curve at any given threshold probability is the optimal decision-making strategy to maximize the net benefit[18]. Also, we calculated the net reclassification index (NRI) and the integrated discrimination index (IDI) to address the risk refinement and the incremental prognostic impact of the Hepamet Fibrosis Score[19].

The method used for missing data was complete-case analysis since statistical packages excluded individuals with any missing value. STATA (12.0, STATA Corporation, College Station, TX, USA) statistical package was used in all analyses and GraphPad Prism (version 6.0; GraphPad Software, Inc., La Jolla, CA) for graphics.

RESULTS

Patients' characteristics

Table 1 shows the baseline features of the estimation and validation cohorts (the individual sets can be seen in **Table S1**). Out of the overall cohort, 54.5% of patients were males with a mean age of 51.9±13.1 years old. The overall prevalence of significant and advanced fibrosis and cirrhosis was 37.7% (925/2452), 20.6% (506/2452) and 5.7% (140/2452), respectively. Briefly, patients included in the estimation cohort were older and showed lower levels of transaminases, HOMA, and triglycerides than the validation cohort. In addition, the training

set showed a higher prevalence of obesity and a lower rate of diabetes. Regarding liver damage, the percentage of significant and advanced fibrosis, as well as cirrhosis was lower in the estimation (22%, 12.1%, and 2.9%, respectively) than the validation population (44.7%, 24.4%, and 7%, respectively).

Development of Hepamet Fibrosis Score

The first step to develop our model was to perform the univariable analysis in the estimation cohort. We found the following variables associated with advanced fibrosis: age (p=0.0001), female sex (p=0.001), diabetes (p=0.0001), ALT (p=0.002), AST (p=0.0001), albumin (p=0.0001), HOMA (p=0.0001), total cholesterol (p=0.017), and platelets (p=0.0001). The first multivariable analysis (including quantitative variables) showed age [OR 1.05 (95%CI 1.03-1.08); p=0.0001], female sex [OR 2.08 (95%CI 1.18-3.66); p=0.011], diabetes [OR 1.66 (95%CI 0.92-3.00); p=0.093], HOMA [OR 1.16 (95%CI 1.10-1.23); p=0.0001], AST [OR 1.02 (95%CI 1.01-1.03); p=0.0001], albumin [OR 2.54 (95%CI 1.30-4.98); p=0.006], and platelets [OR 0.99 (95%CI 0.987-0.995); p=0.0001] independently associated with advanced fibrosis (**Table S2**).

The second multivariable analysis, after transforming the quantitative into categorical variables, found the following variables associated with advanced fibrosis in the estimation cohort: female sex [OR 2.40 (95%CI 1.33-4.33); p=0.004], age 45-64 years old [OR 2.68 (95%CI 1.06-6.77); p=0.037], age ≥65 years old [OR 5.58 (95%CI 2.09-14.92); p=0.001], HOMA ≥4 [OR 4.47 (95%CI 1.49-13.42); p=0.008], diabetes [OR 8.88 (95%CI 3.10-25.44); p=0.0001], AST 35-69 IU/L [OR 2.45 (95%CI 1.37-4.38); p=0.002], AST ≥70 IU/L [OR 8.38 (95%CI 3.72-18.91); p=0.0001], albumin <4 g/dL [OR 2.45 (95%CI 1.14-5.29); p=0.022], platelets 155-220 x10^9/L [OR 2.42 (95%CI 1.35-4.34); p=0.003], and platelets <155 x10^9/L [OR 9.33 (95%CI 4.01-21.67); p=0.0001] (**Table 2)**. The discrimination ability of the second multivariable analysis was higher than the first one (**Figure S2**).

Therefore, the individual risk score for advanced fibrosis was calculated using the following formula derived from the multivariable analysis:

$1 / (1 + e^{(5.390 - 0.986 \times \text{Age [45-64 years old]} - 1.719 \times \text{Age [}\geq\text{65 years old]} + 0.875 \times \text{Male sex} - 0.896 \times \text{AST [35-69 IU/L]} - 2.126 \times \text{AST [}\geq\text{70 IU/L]} - 0.027 \times \text{Albumin [4-4.49 g/dL]} - 0.897 \times \text{Albumin [<4 g/dL]} - 0.899 \times \text{HOMA [2-3.99 with no DM]} - 1.497 \times \text{HOMA [}\geq\text{4 with no DM]} - 2.184 \times \text{Diabetes Mellitus} - 0.882 \times \text{platelets x 1.000/microL [155-219]} - 2.233 \times \text{platelets x 1.000/microL [<155]})}$

A freely online application to estimate the predicted advanced fibrosis rate is available at the following website: https://www.hepamet-fibrosis-score.eu/.

Calibration and discrimination ability of Hepamet Fibrosis Score

Figure S3 shows the observed and predicted probability of advanced fibrosis by Hepamet Fibrosis Score in the estimation and validation sets. Predicted and observed probabilities of advanced fibrosis were similar in the derivation (p=0.351) and validation cohorts (p=0.815).

We show the discrimination ability of the different scores for the estimation and validation cohorts in **Table 3a** and cohort-by-cohort in **Table S3.** Hepamet Fibrosis Score was significantly superior to NFS and FIB-4 both in the estimation cohort and the validation set (**Figure S4**). Also, Hepamet Fibrosis Score revealed the smallest Akaike's information criterion value (HFS: AIC 1837 *vs.* FIB-4: AIC 2023 *vs.* NFS: AIC 2052).

Validation of Hepamet Fibrosis Score

The Hepamet Fibrosis Score cut-offs were 0.12 and 0.47 for advanced fibrosis in the estimation cohort. The performance of the model was evaluated using the same cut-offs in the validation cohort, demonstrating comparable results for advanced fibrosis (**Table 3b**). Besides, we show the sensitivity-specificity plot for the estimation and validation cohorts in **Figure S5**. **Table S4** provides the diagnostic performance of Hepamet Fibrosis Score, NFS, and FIB-4 for the diagnosis of advanced fibrosis in the overall cohort. The prevalence of advanced fibrosis was significantly decreased with the lower cut-off of HFS (8%) in

comparison with NFS (10.7%; p=0.012) and FIB-4 (10.3%, p=0.027). Regarding the higher cut-off, HFS showed a greater prevalence of advanced fibrosis (76.3%) than NFS (55.6%; p<0.0001) and similar than FIB-4 (74.1%; p=0.603). The modifying probability plot for positive and negative likelihood ratio, depending on the cut-off of HFS, is shown in **Figure S6**. According to the number of patients with non-interpretable results, the "grey zone" was lower when using Hepamet Fibrosis Score (21%) than FIB-4 (26%; p<0.05) and NFS (30.8%; p<0.05).

Influence of baseline variables on the Hepamet Fibrosis Score

Hepamet Fibrosis Score showed a significantly higher diagnostic OR for the lower cut-off (<0.12) than age-adjusted FIB-4 and NFS to rule out advanced fibrosis, irrespective of the presence or absence of diabetes (**Figure 1a**) and hypertransaminasemia (**Figure 1b**), as well as BMI (**Figure 1c**) and age groups (**Figure 1d**). On the other hand, the higher cut-off of HFS (>0.47) was superior to NFS >0.675 to rule in advanced fibrosis in all scenarios. Comparing with FIB-4 >2.67, HFS >0.47 showed the greater difference in the diagnostic OR for the groups with *a priori* low risk of liver damage (lack of diabetes, ALT<40, lean and younger patients), while it was slightly better in high-risk patients (**Figures 2a, 2b, 2c,** and **2d**).

Clinical usefulness of Hepamet Fibrosis Score: A decision curve analysis

A decision curve analysis was added to analyze the clinical utility of Hepamet Fibrosis Score guiding to perform a liver biopsy compared with NFS and FIB-4. The decision curve analysis indicated that, from a threshold probability of >10%, we could obtain more net benefit guided by Hepamet Fibrosis Score than the reference strategies (NFS and FIB-4) and to biopsy all or no patients. Particularly, we could obtain a net benefit of 10.4%, 6%, 3.1% and 1.1% at threshold probabilities of 20%, 40%, 60% and 80% (**Figure 3**). Although the percentages could seem low, it must be interpreted in the context of the prevalence. The maximum possible value of the net benefit that can be achieved in this study corresponds to the

prevalence of advanced fibrosis (20.6%). For example, a net benefit of 10.4% achieved at 20% threshold probability represents until 50% (0.104/0.206*100%) of the maximal benefit.

Hepamet Fibrosis Score led to significant improvements in reclassification, compared to NFS [NRI 31.7% (95%CI 15.1–48.2)] and FIB-4 [NRI 25.3% (95%CI 16–33.7)]. These results indicate that Hepamet Fibrosis Score correctly reclassified subjects with and without advanced fibrosis. Also, Hepamet Fibrosis Score improved the IDI significantly in comparison with NFS [IDI 0.1170 (95%CI 0.1077–0.1263)] and FIB-4 [IDI 0.07 (95%CI 0.0624–0.0776)] (**Supplementary Table 5**).

DISCUSSION

In the current study, including a large international cohort of biopsy-proven NAFLD patients, we demonstrated that Hepamet Fibrosis Score (HFS) (including age, sex, diabetes, HOMA, AST, albumin, and platelets) determine liver fibrosis staging better than NFS and FIB-4. This new score showed greater clinical utility to guide the decision to make diagnostic liver biopsies in patients with NAFLD, representing a user-friendly tool that emerges as an accurate non-invasive method beyond transaminases to screen and manage a silent disease.

Several serum-based methods have been developed to detect individuals at risk of advanced fibrosis in NAFLD[20]. NFS and FIB-4 (initially designed for hepatitis C[21]) are the most used scores, showing AUROCs around 0.80 for advanced fibrosis[22]. Hepamet Fibrosis Score improved the diagnostic accuracy significantly for advanced fibrosis in comparison with them. Two major strengths must be highlighted in its development: the wide external international validation and the statistical approach. Firstly, Hepamet Fibrosis Score has been calculated with almost 2,500 patients from five countries (Spain, France, Italy, Cuba, and China), including various ethnicities (Caucasian, Latin, and Asian populations) and different rates of baseline features (diabetes, obesity, the prevalence of fibrosis). Given that HFS scored

similarly between these cohorts, the final results must be considered robust. Secondly, we selected a multivariable analysis to develop the score using categorical variables. This approach showed better diagnostic accuracy because of the effect of capping age, platelets, albumin, and AST levels. For example, older age was associated with advanced fibrosis in our study, but its impact caused more false than true positive cases over than 65 years old, similarly to other studies[11]. Also, HOMA was combined with diabetes in the same variable to improve reliability and because HOMA is not a useful marker for insulin resistance in diabetes (i.e., it is modified by insulin sensitizers or exogenous insulin). Thus, HOMA does not need to be calculated in diabetic patients. On the other hand, HFS <0.12 showed the lowest negative and HFS $\geq$0.47 the highest positive likelihood ratio for advanced fibrosis. Consequently, the post-test probabilities using Hepamet Fibrosis Score were significantly better than NFS and FIB-4.

Current biochemical non-invasive methods show some major drawbacks. On the one hand, there are a high proportion of patients allocated to the "grey zone" in NFS and FIB-4[23]. By contrast, patients assigned to undetermined results were significantly lower for Hepamet Fibrosis Score than FIB-4 and NFS. On the other hand, many baseline factors can influence the diagnostic performance of serum-based scores. First, both NFS and FIB-4 require age-adjusted cut-offs to improve the diagnostic accuracy (particularly, specificity) for advanced fibrosis in patients older than 65 years old[11]. By contrast, Hepamet Fibrosis Score did not require to be adjusted for age. Second, it has been estimated that up to two-thirds of cirrhotic patients showed normal levels of transaminases, which represent the main alert of underlying liver disease in clinical practice[24]. HFS showed the highest diagnostic effectiveness of the three scores in the population without hypertransaminasemia, so it could be useful covering the gap of early identification of at-risk NAFLD patients. Third, non-invasive scores have moderate success in predicting fibrosis in obese patients[12]. HFS had the highest diagnostic OR to rule out advanced fibrosis across all the BMI groups, while the higher cut-off was

significantly superior in lean patients compared with FIB-4 and NFS. Notably, the percentage of false positives rose dramatically with the BMI for NFS. Fourth, diabetes influences the accuracy of the prediction of the non-invasive scores[25]. In our study, HFS showed the highest diagnostic effectiveness of the scores in patients without diabetes, while it was slightly better than FIB-4 for patients with this entity.

Adding decision curve analysis to statistical approaches based on metrics could help for clinical decision making[26]. In our study, this statistical approach weighed the true and false positive results of Hepamet Fibrosis Score (detecting advanced fibrosis *vs.* unnecessary biopsy) and demonstrated a greater net benefit leading the decision of performing a liver biopsy, compared to NFS and FIB-4. No previous calculation of net benefit has been found in the literature of non-invasive methods in NAFLD. Also, the NRI suggested that Hepamet Fibrosis Score was able to improve the correct classification of patients. This point is relevant because EASL guidelines recommend the use of non-invasive scores to help in decision making[27]. The usefulness of Hepamet Fibrosis Score on detection of NAFLD-fibrosis in general population by primary care and other non-hepatologist physicians should be addressed in future studies, as well as its combination with transient elastography in order to maximize the accuracy of the prediction of liver fibrosis.

In summary, in this large international study, Hepamet Fibrosis Score demonstrated to be more accurate to stage liver fibrosis in NAFLD, with better calibration and net benefit, than NFS and FIB-4. Future studies analyzing the impact of HFS on clinical outcomes in NAFLD and a potential combination of Hepamet Fibrosis Score with imaging biomarkers to improve the continuum of care of the patients with NAFLD are warranted.

TABLE LEGENDS

Table 1. Baseline characteristics of the estimation and validation cohorts.

Table 2. Variables associated with advanced fibrosis in the estimation cohort. **BMI, ALT and total cholesterol were included in the multivariable analysis, but they were not significant.*

Table 3. A) Discrimination ability of the Hepamet Fibrosis Score, compared with NAFLD Fibrosis Score and FIB-4, in both estimation and validation cohorts. B) Operating characteristics for the two selected cut-offs of the Hepamet Fibrosis Score, regarding advanced fibrosis in both estimation and validation cohorts. **Age-adjusted cut-off for subjects older than 65 years old were used for NFS and FIB-4.*

Supplementary Table 1. Baseline characteristics of the individual cohorts.

Supplementary Table 2. Univariable and multivariable analyses (including quantitative variables) regarding advanced fibrosis in the estimation cohort.

Supplementary Table 3. Discrimination ability of the Hepamet Fibrosis Score, compared with NAFLD Fibrosis Score and FIB-4, cohort by cohort.

Supplementary Table 4. Operating characteristics for the two selected cut-offs of the Hepamet Fibrosis Score, compared with NAFLD Fibrosis Score and FIB-4, regarding advanced fibrosis in the overall cohort.

Supplementary Table 5. Net reclassification index and integrated discrimination improvement between Hepamet Fibrosis Score and the other models.

FIGURE LEGENDS

Figure 1. Unadjusted diagnostic OR for advanced fibrosis for the lower cut-offs for Hepamet Fibrosis Score, NAFLD Fibrosis Score, and FIB-4, depending on: A) BMI; B) Age; C) Hypertransaminasemia; D) Diabetes mellitus. **Age-adjusted cut-off for subjects older than 65 years old were used for NFS and FIB-4.*

Figure 2. Unadjusted diagnostic OR for advanced fibrosis for the higher cut-offs for Hepamet Fibrosis Score, NAFLD Fibrosis Score, and FIB-4, depending on: A) BMI; B) Age; C) Hypertransaminasemia; D) Diabetes mellitus.

Figure 3. Decision curve analysis showing the highest net benefit of the strategy based on Hepamet Fibrosis Score.

Supplementary Figure 1. Transformation of the continuous into qualitative variables.

Supplementary Figure 2. Accuracy of the Hepamet Fibrosis Score, comparing the first and second multivariable analyses, in predicting advanced fibrosis in the estimation cohort.

Supplementary Figure 3. Calibration belt for the Hepamet Fibrosis Score. A) Estimation cohort. B) Validation cohort.

Supplementary Figure 4. Accuracy of the Hepamet Fibrosis Score, compared with NAFLD Fibrosis Score and FIB-4, in predicting advanced fibrosis in the estimation cohort.

Supplementary Figure 5. Plot of sensitivity *versus* specificity for Hepamet Fibrosis Score. A) Estimation cohort. B) Validation cohort.

Supplementary Figure 6. Plot showing post-test probability depending on the prevalence, and positive and negative likelihood ratios. A) HFS cut-off 0.12. B) HFS cut-off 0.47.

REFERENCES

1. Bellentani S. The epidemiology of non-alcoholic fatty liver disease. Liver Int 2017;37:81-84.

2. Younossi ZM, Blissett D, Blissett R, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology 2016;64:1577-1586.

3. Ampuero J, Romero-Gómez M. Influence of non-alcoholic fatty liver disease on cardiovascular disease. Gastroenterol Hepatol 2012;35:585-593.

4. Ampuero J, Gallego-Durán R, Romero-Gómez M. Association of NAFLD with subclinical atherosclerosis and coronary-artery disease: Meta-analysis. Rev Esp Enfermedades Dig 2015;107:10-16.

5. Musso G, Cassader M, Cohney S, et al. Fatty Liver and Chronic Kidney Disease: Novel Mechanistic Insights and Therapeutic Opportunities. Diabetes Care 2016;39:1830-1845.

6. Kim G-A, Lee HC, Choe J, et al. Association between non-alcoholic fatty liver disease and cancer incidence rate. J Hepatol 2018;68:140-146.

7. Hagström H, Nasr P, Ekstedt M, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017;67:1265-1273.

8. Ampuero J, Aller R, Gallego-Durán R, et al. The effects of metabolic status on non-alcoholic fatty liver disease-related outcomes, beyond the presence of obesity. Aliment Pharmacol Ther 2018;48:1260-1270.

9. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: A noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007;45:846-854.

10. McPherson S, Stewart SF, Henderson E, et al. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut 2010;59:1265-1269.

11. McPherson S, Hardy T, Dufour J-F, et al. Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. Am J Gastroenterol 2017;112.

12. Ooi GJ, Burton PR, Doyle L, et al. Modified thresholds for fibrosis risk scores in nonalcoholic fatty liver disease are necessary in the obese. Obes Surg 2017;27:115-125.

13. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. J Hepatol 2018;68:305-315.

14. Shah AG, Lydecker A, Murray K, et al. Comparison of Noninvasive Markers of Fibrosis in Patients

With Nonalcoholic Fatty Liver Disease. Clin Gastroenterol Hepatol 2009;7:1104-1112.

15. Bedossa P, FLIP Pathology Consortium. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology 2014;60:565-575.

16. Finazzi S, Poole D, Luciani D, et al. Calibration Belt for Quality-of-Care Assessment Based on Dichotomous Outcomes. PLoS One 2011;6:e16110.

17. Hilden J, Glasziou P. Regret graphs, diagnostic uncertainty and Youden's Index. Stat Med 1996;15:969-986.

18. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. Med Decis Mak 2006;26:565-574.

19. Pencina MJ, D'Agostino RB, D'Agostino RB, et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157-172; discussion 207-12.

20. Xiao G, Zhu S, Xiao X, et al. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. Hepatology 2017;66:1486-1501.

21. Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317-1325.

22. Petta S, Wong VW-S, Cammà C, et al. Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver fibrosis in patients with NAFLD. Aliment Pharmacol Ther 2017;46:617-627.

23. Castera L. Diagnosis of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: Non-invasive tests are enough. Liver Int 2018;38:67-70.

24. Ampuero J, Romero-Gómez M. Editorial: looking for patients at risk of cirrhosis in the general population-many needles in a haystack. Aliment Pharmacol Ther 2018;47:692-694.

25. Bertot LC, Jeffrey GP, de Boer B, et al. Diabetes impacts prediction of cirrhosis and prognosis by non-invasive fibrosis models in non-alcoholic fatty liver disease. Liver Int 2018;38:1793-1802.

26. Vickers AJ. Decision Analysis for the Evaluation of Diagnostic Tests, Prediction Models, and Molecular Markers. Am Stat 2008;62:314-320.

27. European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL–EASD–EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. J Hepatol 2016;64:1388-1402.

Figure 1a

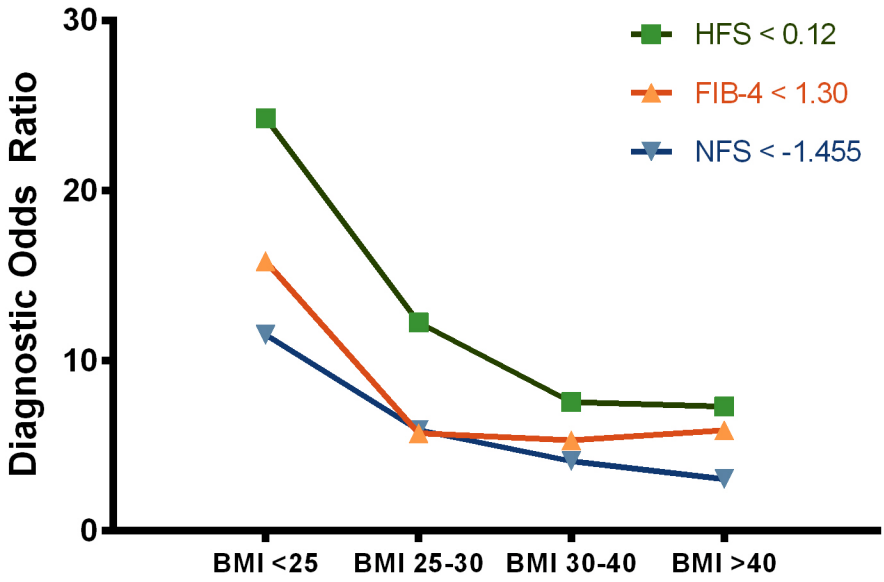


Figure 1b

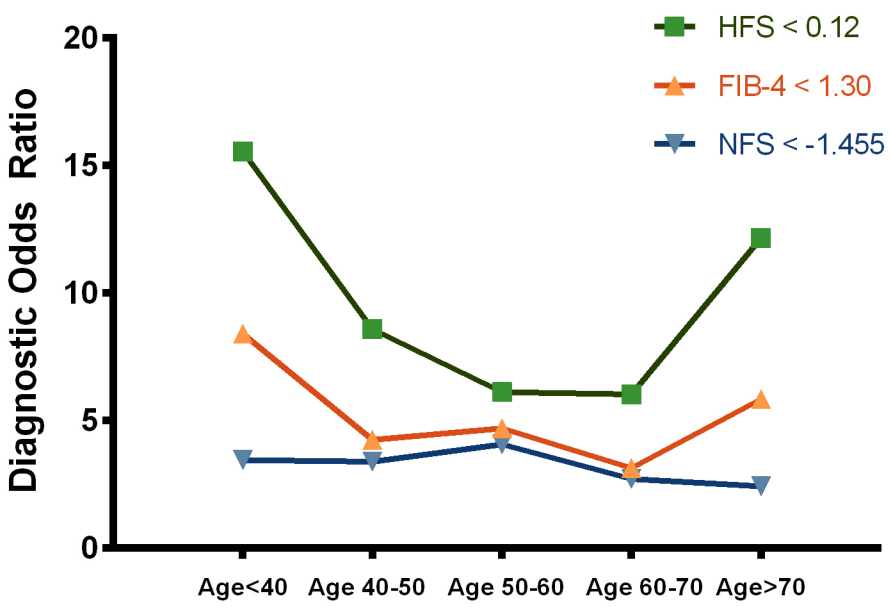


Figure 1c

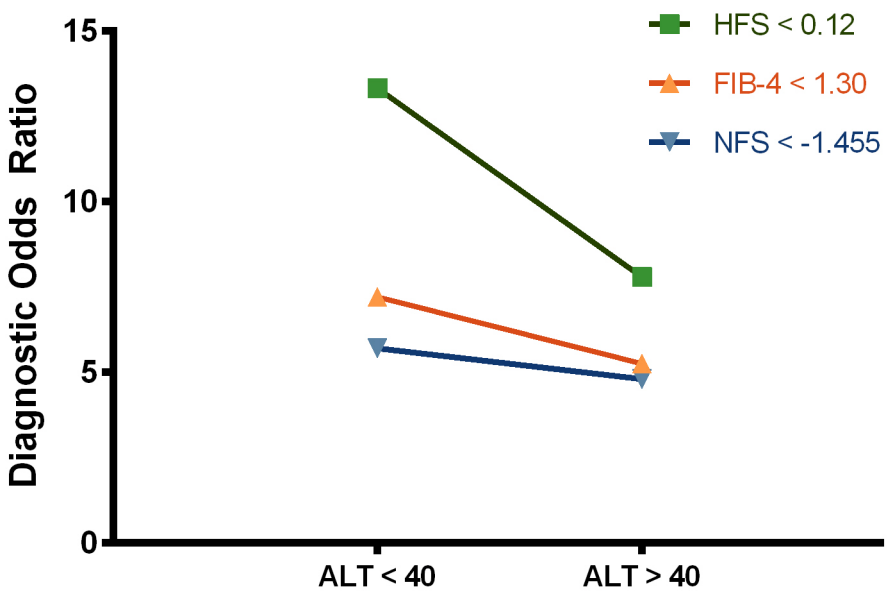


Figure 1d

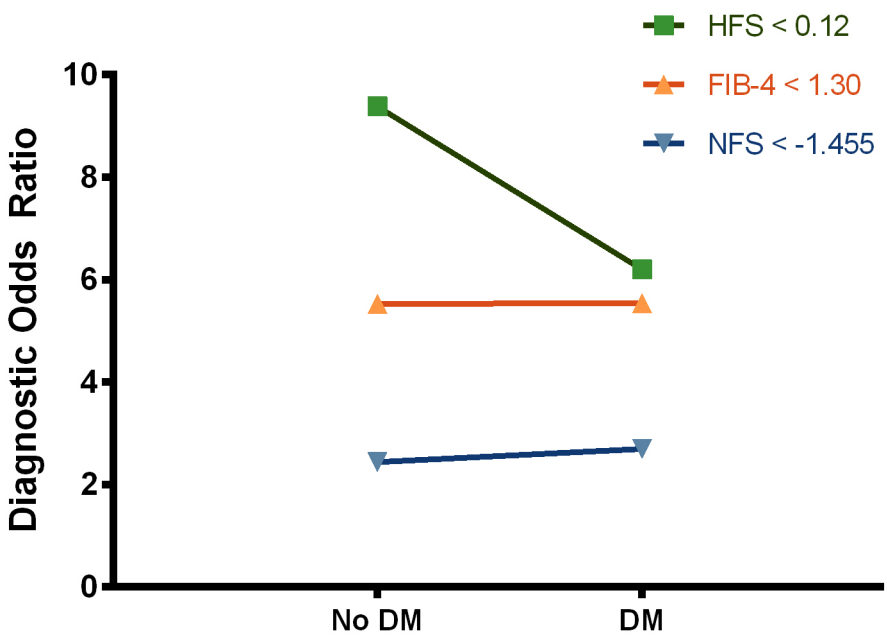


Figure 2a

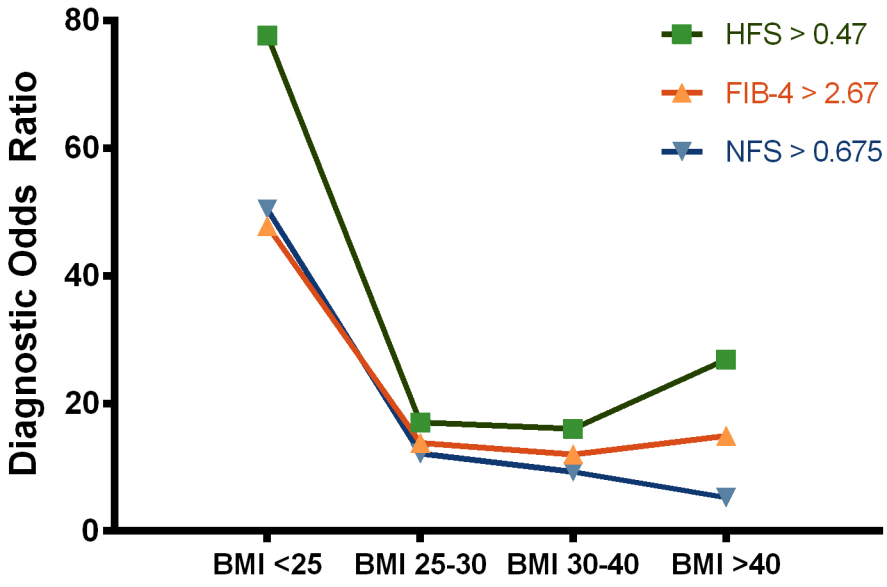


Figure 2b

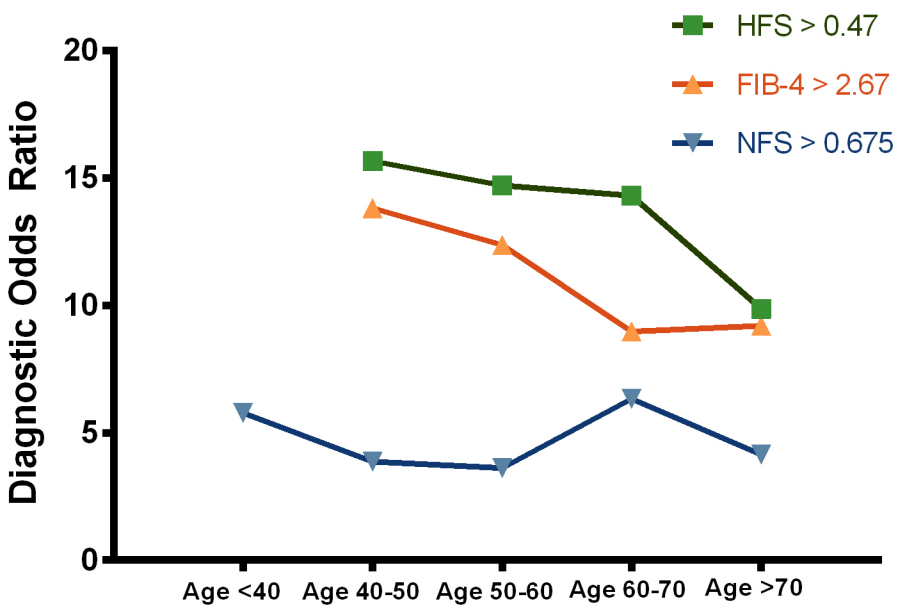


Figure 2c

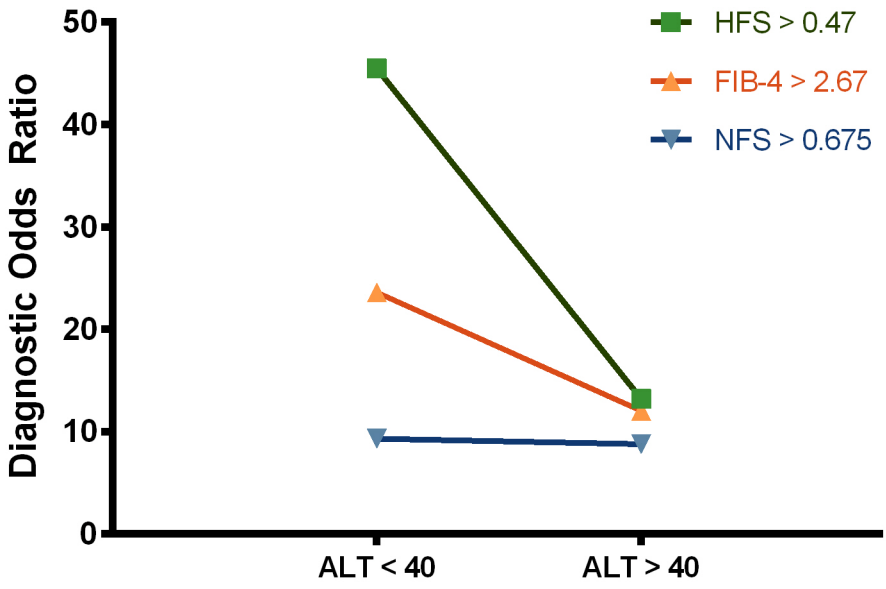


Figure 2d

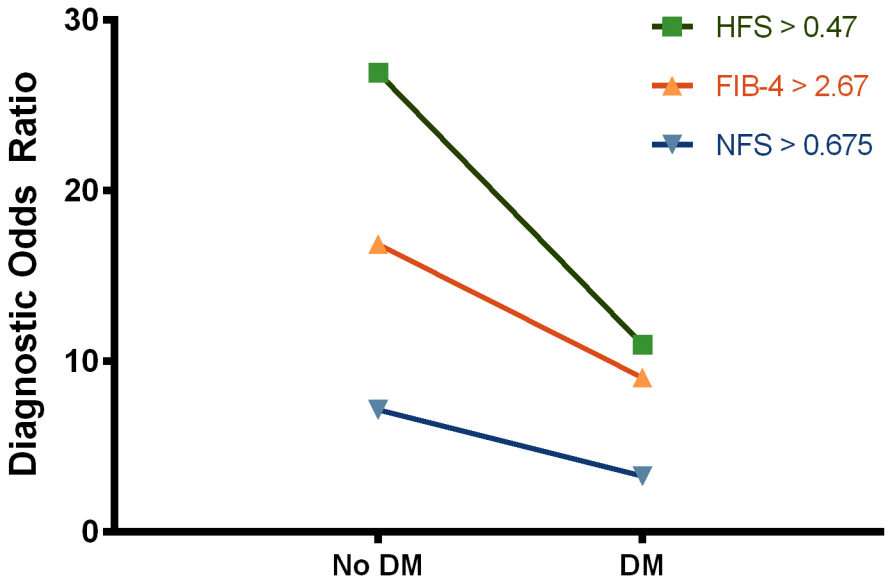
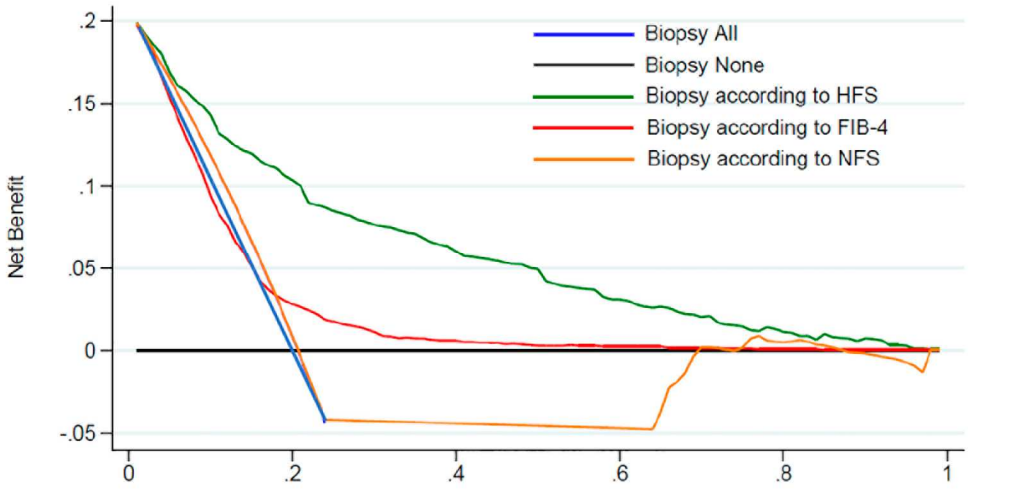


Figure 3



Biopsy All	+0.008	-0.322	-0.983	-2.966	-78.33
HFS	+0.104	+0.060	+0.031	+0.011	+0.001
FIB-4	+0.028	+0.006	+0.002	+0.000	+0.000
NFS	+0.008	-0.203	-0.068	+0.005	+0.000

Characteristic	Estimation Cohort (n=758)	Validation cohort (N=1694)	P value
Male sex	44.9% (340/758)	58.9% (997/1694)	0.0001
Age; years ± SD	53.9 ± 12.4	51 ± 13.3	0.0001
BMI ± SD	36.4 ± 10.1	31.7 ± 6.9	0.0001
Obesity (BMI≥30)	64.9% (491/757)	52.3% (882/1688)	0.0001
Arterial Hypertension	43.4% (326/752)	47.3% (679/1436)	0.080
Type 2 Diabetes Mellitus	27.6% (209/758)	37.8% (634/1679)	0.0001
Glucose ± SD (mg/dL)	110 ± 36	113 ± 43	0.047
HOMA-IR ± SD	4.7 ± 4.3	6.3 ± 10	0.0001
Total cholesterol ± SD (mg/dL)	195 ± 44	194 ± 48	0.731
HDL-c ± SD (mg/dL)	53 ± 22	45 ± 19	0.0001
Triglycerides ± SD (mg/dL)	155 ± 81	166 ± 104	0.004
Albumin ± SD (g/dL)	4.38 ± 0.4	4.40 ± 0.4	0.292
Bilirubin ± SD (mg/dL)	0.75 ± 1.01	0.69 ± 0.42	0.033
Creatinine ± SD (mg/dL)	0.83 ± 0.3	0.85 ± 0.3	0.126
Platelet count ± SD (x 10^9/L)	251 ± 73	230 ± 66	0.0001
AST ± SD (IU/mL)	35 ± 26	46 ± 32	0.0001
ALT ± SD (IU/mL)	50 ± 40	66 ± 52	0.0001
NASH	47.2% (358/758)	43% (726/1688)	0.052
Significant fibrosis (F2-F4)	22% (167/758)	44.7% (758/1694)	0.0001
Advanced fibrosis (F3-F4)	12.1% (92/758)	24.4% (414/1694)	0.0001
Cirrhosis	2.9% (22/758)	7% (118/1694)	0.0001

Characteristic	Unadjusted (Univariable Analysis)	Adjusted (Multivariable Analysis)
Female sex	OR 2.14 (95%CI 1.33-3.42); p=0.002	OR 2.40 (95%CI 1.33-4.33); p=0.004
Age		
< 45 years old	Reference	Reference
45-64 years old	OR 3.80 (95%CI 1.60-9.05); p=0.003	OR 2.68 (95%CI 1.06-6.77); p=0.037
≥ 65 years old	OR 10.01 (95%CI 4.09-24.51); p=0.0001	OR 5.58 (95%CI 2.09-14.92); p=0.001
HOMA – DM		
HOMA < 2	Reference	Reference
HOMA 2 – 3.99	OR 1.69 (95%CI 0.58-4.91); p=0.333	OR 2.46 (CI95% 0.76-7.92); p=0.132
HOMA ≥ 4	OR 4.74 (95%CI 1.77-12.71); p=0.002	OR 4.47 (95%CI 1.49-13.42); p=0.008
Diabetes mellitus	OR 9.18 (95%CI 3.56-23.66); p=0.0001	OR 8.88 (95%CI 3.10-25.44); p=0.0001
Albumin		
≥ 4.5 g/dL	Reference	Reference
4 – 4.49 g/dL	OR 1.86 (95%CI 1.11-3.12); p=0.018	OR 1.03 (95%CI 0.56-1.88); p=0.929
< 4 g/dL	OR 3.81 (95%CI 2.01-7.25); p=0.0001	OR 2.45 (95%CI 1.14-5.29); p=0.022
Platelet count		
≥ 220 x 10^9/L	Reference	Reference
155 – 219 x 10^9/L	OR 2.25 (95%CI 1.35-3.74); p=0.002	OR 2.42 (95%CI 1.35-4.34); p=0.003
< 155 x 10^9/L	OR 12.50 (95%CI 6.54-23.89); p=0.0001	OR 9.33 (95%CI 4.01-21.67); p=0.0001
AST		
< 35 IU/mL	Reference	Reference
35 – 69 IU/mL	OR 2.94 (95%CI 1.79-4.83); p=0.0001	OR 2.45 (95%CI 1.37-4.38); p=0.002
≥ 70 IU/mL	OR 9.42 (95%CI 4.89-18.13); p=0.0001	OR 8.38 (95%CI 3.72-18.91); p=0.0001

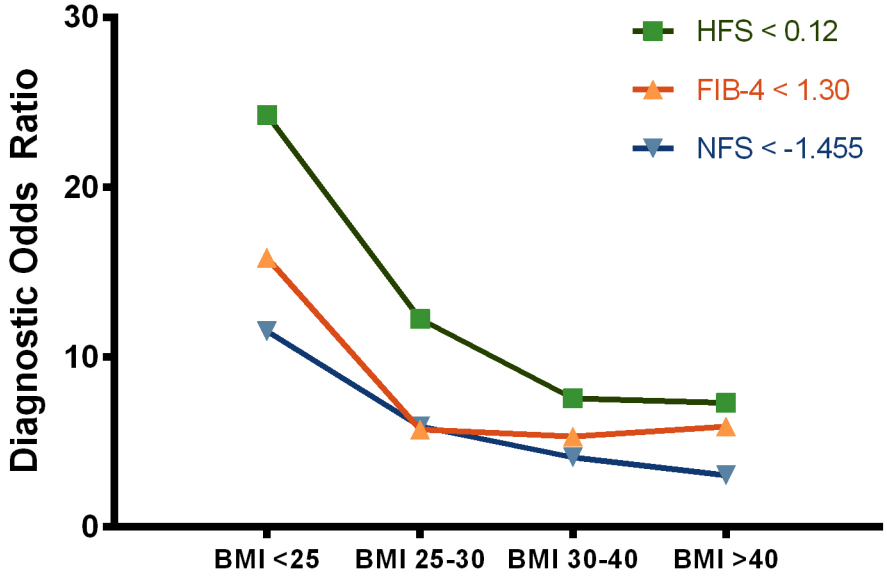
Estimation Cohort (n=758)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.850 (95%CI 0.807-0.893)	0.775 (95%CI 0.723-0.828) p=0.0025	0.772 (95%CI 0.713-0.832) p=0.0002

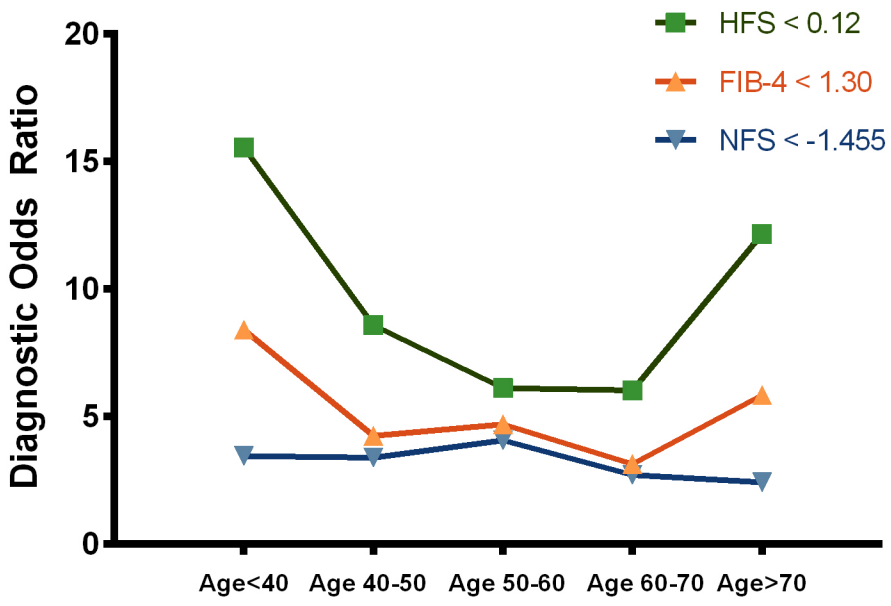
Validation Cohort (n=1694)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.844 (95%CI 0.819-0.869)	0.789 (95%CI 0.764-0.814) p<0.0001	0.801 (95%CI 0.776-0.826) p<0.0001

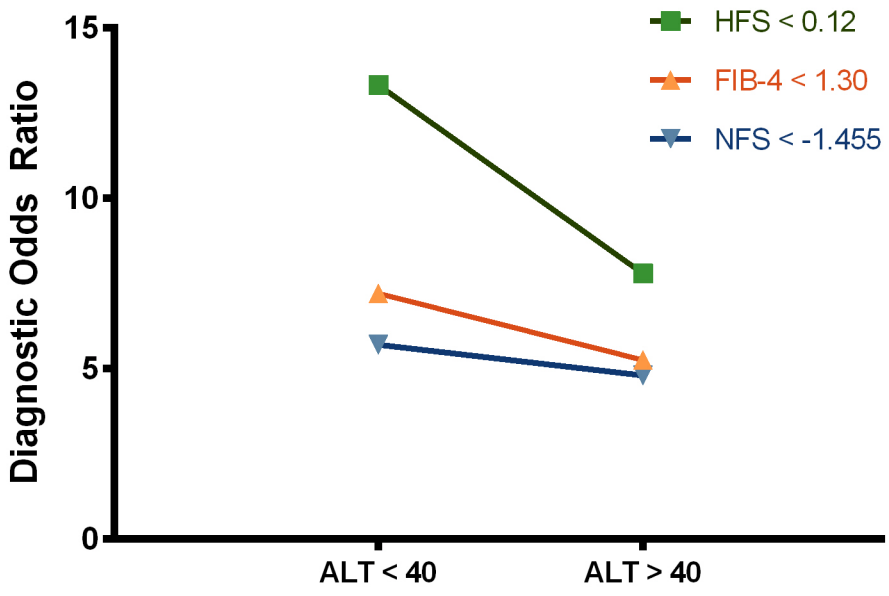
Overall Cohort (N=2452)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.848 (95%CI 0.826-0.869)	0.778 (95%CI 0.756-0.801) p<0.0001	0.802 (95%CI 0.780-0.825) p<0.0001

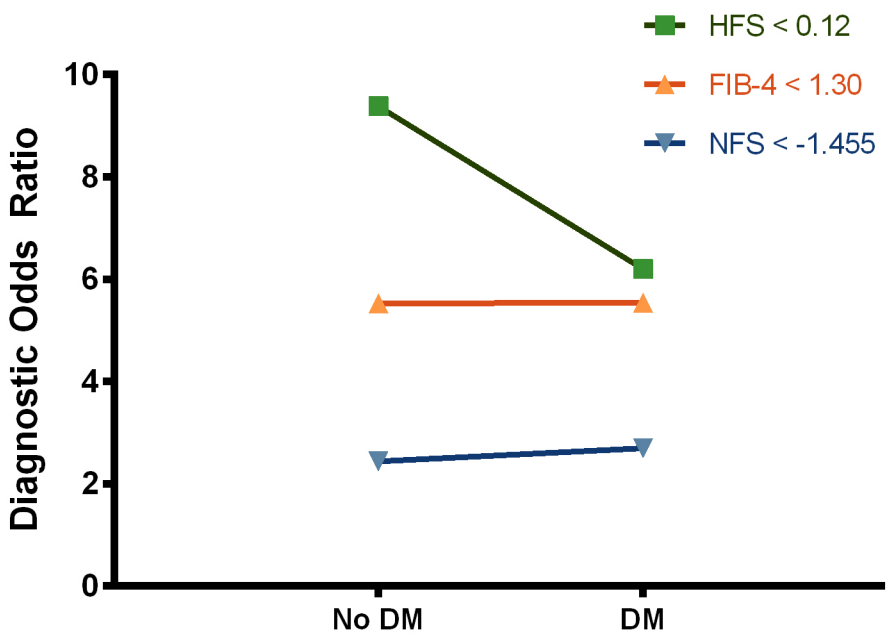
p value *vs.* Hepamet Fibrosis Score

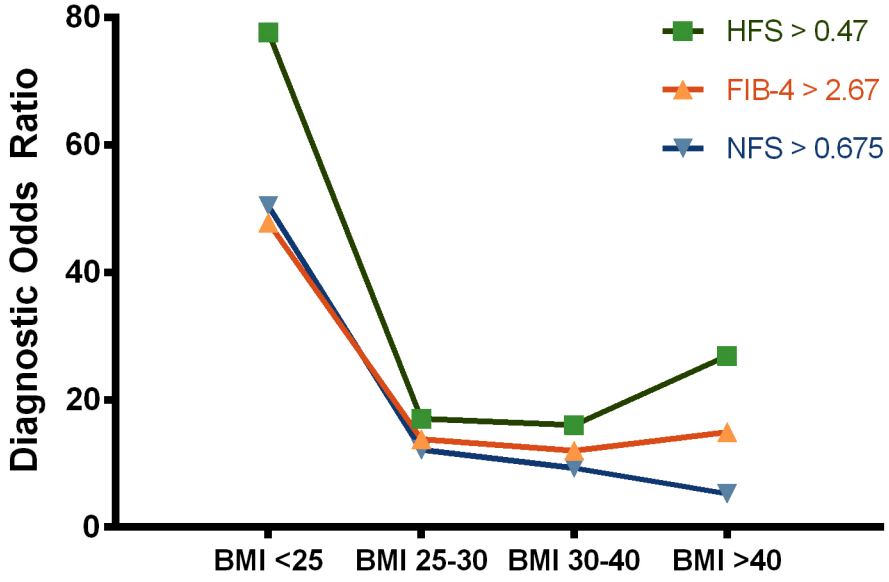
	Estimation Cohort		**Validation Cohort**	
Advanced Fibrosis (%)	12.1%		24.6%	
Cut-off	< 0.12	≥ 0.47	< 0.12	≥ 0.47
Sensitivity (%)	70.7	38	74.6	34.6
Specificity (%)	80.9	98	75.5	96.7
PPV (%)	33.9	72.9	49.8	77.2
NPV (%)	95.2	92	90.1	81.9
LR+	3.71	15.24	3.05	10.40
LR-	0.36	0.63	0.34	0.68

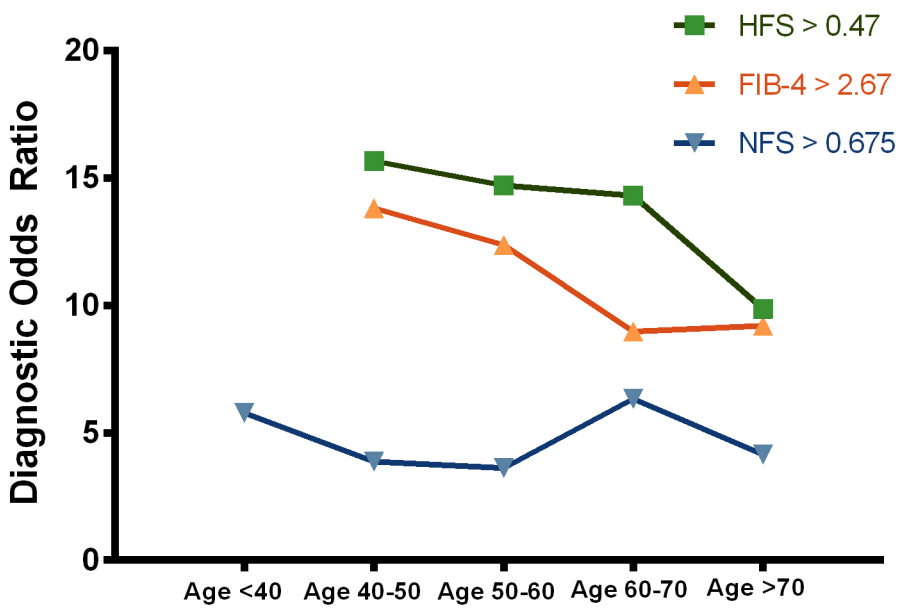




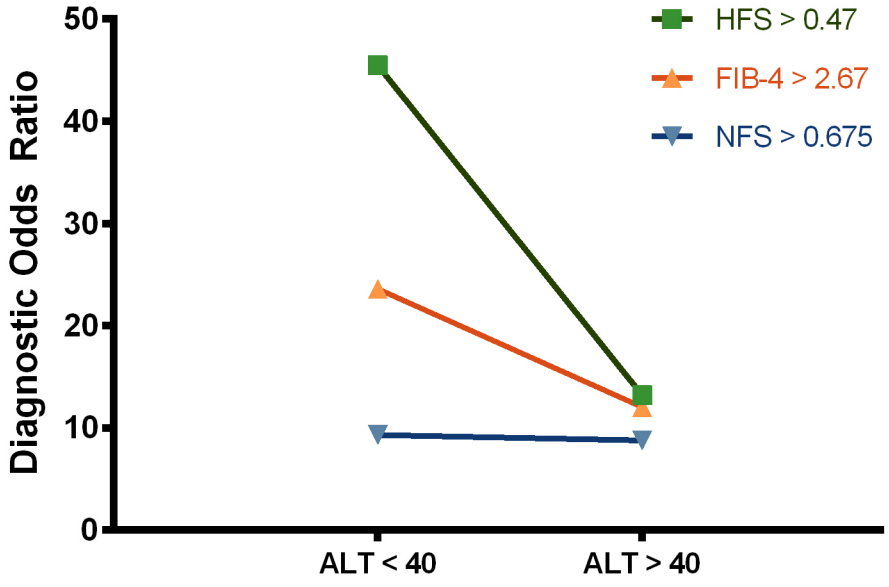


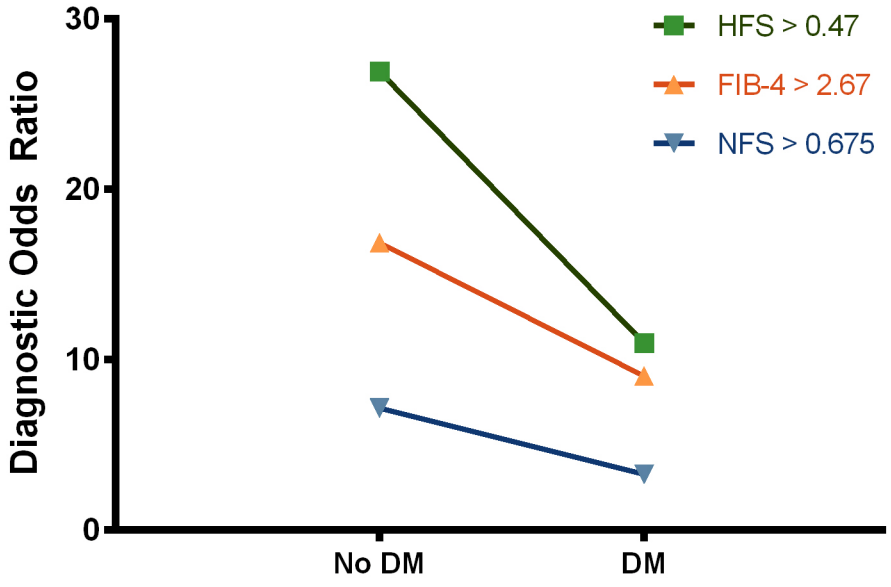


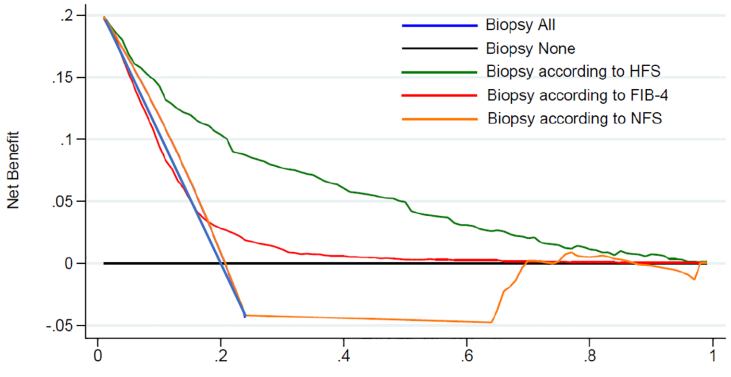




Age < 40 Age 40-50 Age 50-60 Age 60-70







Biopsy All	+0.008	-0.322	-0.983	-2.966	-78.33
HFS	+0.104	+0.060	+0.031	+0.011	+0.001
FIB-4	+0.028	+0.006	+0.002	+0.000	+0.000
NFS	+0.008	-0.203	-0.068	+0.005	+0.000

Characteristic	Spanish Cohort (n=758)	French cohort No. 1 (N=444)	French cohort No. 2 (N=386)	Cuban Cohort (n=344)	Italian Cohort (n=288)	Chinese Cohort (n=232)
Male sex	44.9%	60.4%	61.1%	42.2%	62.5%	72.4%
Age; years ± SD	53.9 ± 12.4	54.2 ± 12.3	56.1 ± 12.2	51.1 ± 12.8	46.2 ± 13.3	42.5 ± 12.4
BMI ± SD	36.4 ± 10.1	31.4 ± 6.5	32.5 ± 6	36 ± 8.3	29.9 ± 5.1	26.7 ± 4.3
Obesity (BMI≥30)	64.9%	50.7%	63.5%	74.7%	44%	13.4%
Arterial Hypertension	43.4%	48.1%	57.5%	50.9%	28.1%	27%
Type 2 Diabetes Mellitus	27.6%	45.9%	43.8%	43.9%	21.5%	24.1%
Glucose ± SD (mg/dL)	110 ± 36	116 ± 43	122 ± 47	118 ± 48	99 ± 31	103 ± 30
HOMA-IR ± SD	4.7 ± 4.3	4.8 ± 5	8.5 ± 14	7.9 ± 12.9	4.1 ± 3	5.9 ± 8
Total cholesterol ± SD (mg/dL)	195 ± 44	190 ± 46	197 ± 47	189 ± 52	206 ± 46	194 ± 46
HDL-c ± SD (mg/dL)	53 ± 22	45 ± 17	45 ± 14	44 ± 32	51 ± 17	40 ± 9
Triglycerides ± SD (mg/dL)	155 ± 81	150 ± 93	167 ± 113	174 ± 97	146 ± 78	210 ± 131
Albumin ± SD (g/dL)	4.38 ± 0.4	4.38 ± 0.4	4.25 ± 0.4	4.26 ± 0.5	4.60 ± 0.4	4.64 ± 0.3
Bilirubin ± SD (mg/dL)	0.75 ± 1.01	0.63 ± 0.47	0.68 ± 0.42	0.69 ± 0.40	0.67 ± 0.35	0.82 ± 0.38
Creatinine ± SD (mg/dL)	0.83 ± 0.3	0.90 ± 0.25	0.83 ± 0.18	0.90 ± 0.35	0.88 ± 0.34	0.76 ± 0.17
Platelet count ± SD (x 10^9/L)	251 ± 73	229 ± 63	223 ± 67	223 ± 69	232 ± 69	250 ± 58
AST ± SD (IU/mL)	35 ± 26	46 ± 30	46 ± 34	44 ± 21	46 ± 21	46 ± 32
ALT ± SD (IU/mL)	50 ± 40	60 ± 42	63 ± 38	61 ± 53	81 ± 51	73 ± 74
NASH	47.2%	46.5%	29.9%	31.7%	80.9%	28%
Significant fibrosis (F2-F4)	22%	52.3%	61.9%	35.8%	46.9%	12.5%
Advanced fibrosis (F3-F4)	12.1%	27.3%	35.8%	25.3%	20.8%	3.4%
Cirrhosis	2.9%	6.8%	7.3%	11.3%	7.3%	0%

Characteristic	Fibrosis F3-4 (n=92)	Fibrosis F0-2 (N=666)	Univariable Analysis	Multivariable Analysis
Female sex	70.7% (65/92)	53% (353/666)	0.001	OR 2.08 (95%CI 1.18-3.66); p=0.011
Age; years ± SD	61.1 ± 10.1	52.9 ± 12.3	0.0001	OR 1.05 (95%CI 1.03-1.08); p=0.0001
BMI ± SD	37.5 ± 10.2	36.2 ± 10.1	0.247	
Obesity (BMI≥30)	70.7% (65/92)	64.1% (426/665)	0.214	
Arterial Hypertension	64.4% (58/90)	40.5% (268/662)	0.0001	
Type 2 Diabetes Mellitus	54.3% (50/92)	23.9% (159/666)	0.0001	OR 1.66 (95%CI 0.92-3.00); p=0.093
Glucose ± SD (mg/dL)	129 ± 50	107 ± 33	0.0001	
HOMA-IR ± SD	8.6 ± 7	4.2 ± 3.4	0.0001	OR 1.16 (95%CI 1.10-1.23); p=0.0001
Total cholesterol ± SD (mg/dL)	185 ± 43	197 ± 44	0.017	
HDL-c ± SD (mg/dL)	50 ± 23	53 ± 22	0.244	
Triglycerides ± SD (mg/dL)	161 ± 69	154 ± 83	0.480	
Albumin ± SD (g/dL)	4.20 ± 0.45	4.40 ± 0.4	0.0001	OR 2.54 (95%CI 1.30-4.98); p=0.006
Bilirubin ± SD (mg/dL)	1.05 ± 2.55	0.71 ± 0.52	0.216	
Creatinine ± SD (mg/dL)	0.85 ± 0.4	0.83 ± 0.3	0.571	
Platelet count ± SD (x 10^9/L)	209 ± 85	257 ± 70	0.0001	OR 0.99 (95%CI 0.987-0.995); p=0.0001
AST ± SD (IU/mL)	50 ± 31	32 ± 25	0.0001	OR 1.02 (95%CI 1.01-1.03); p=0.0001
ALT ± SD (IU/mL)	62 ± 41	48 ± 40	0.002	

Spanish Cohort (n=758)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.850 (95%CI 0.807-0.893)	0.775 (95%CI 0.723-0.828)	0.772 (95%CI 0.713-0.832)

French Cohort No. 1 (n=444)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.800 (95%CI 0.751-0.849)	0.768 (95%CI 0.717-0.820)	0.764 (95%CI 0.710-0.817)

French Cohort No. 2 (n=386)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.810 (95%CI 0.766-0.853)	0.749 (95%CI 0.700-0.799)	0.765 (95%CI 0.716-0.815)

Italian Cohort (n=288)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.843 (95%CI 0.790-0.895)	0.785 (95%CI 0.711-0.858)	0.773 (95%CI 0.706-0.840)

Cuban Cohort (n=344)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.854 (95%CI 0.810-0.899)	0.768 (95%CI 0.709-0.828)	0.830 (95%CI 0.781-0.880)

Chinese Cohort (n=232)			
	Hepamet Fibrosis Score	NAFLD Fibrosis Score	FIB-4
Advanced Fibrosis (F0-2 vs. F3-4)	0.904 (95%CI 0.829-0.979)	0.812 (95%CI 0.709-0.915)	0.787 (95%CI 0.644-0.930)

ADVANCED FIBROSIS (prevalence 20.6%)						
	Hepamet Fibrosis Score		**NAFLD Fibrosis Score**		**FIB-4**	
Cut-off	< 0.12	≥ 0.47	< -1.455	> 0.675	< 1.30	≥ 2.67
Sensitivity (%)	73.9	35.2	70.5	32.9	66.9	29.6
Specificity (%)	77.4	97.2	63.6	93.2	74.8	97.3
PPV (%)	46	76.3	33.5	55.6	40.8	74.1
NPV (%)	91.9	85.2	89.3	84.2	89.7	84.2
LR+	3.27	13.22	1.94	4.81	2.66	10.03
LR-	0.31	0.67	0.46	0.72	0.44	0.72
Post-Test Probability (+) (%)	46	79.7	33.5	55.5	40.8	74.1
Post-Test Probability (-) (%)	6.4	13.5	10.7	15.7	10.3	15.8

	HFS *vs.* FIB-4		**HFS *vs.* NFS**	
	Values	P-Value	Values	P-Value
NRI (95% CI)	25.3% (16-33.7)	<0.0001	31.7% (15.1-48.2)	<0.0001
% of events correctly reclassified	2.2%	<0.0001	4.4%	<0.0001
% of non-events correctly reclassified	23.1%	<0.0001	27.3%	<0.0001
IDI (95% CI)	0.0700 (0.0624-0.0776)	<0.0001	0.1170 (0.1077-0.1263)	<0.0001