IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Impact of Obesity and Alanine Aminotransferase Levels on the Diagnostic Accuracy for Advanced Liver Fibrosis of Noninvasive Tools in Patients with Nonalcoholic Fatty Liver Disease**

**This is a pre print version of the following article:**

*Original Citation:*

*Availability:*

This version is available http://hdl.handle.net/2318/1725580 since 2020-01-28T12:32:02Z

*Published version:*

DOI:10.14309/ajg.0000000000000153

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

23 December 2024

**TITLE: Impact of Obesity and ALT Levels on the Diagnostic Accuracy for Advanced Liver Fibrosis of Noninvasive Tools in NAFLD Patients**

**AUTHORS**: Salvatore Petta[1], Vincent Wai-Sun Wong[2,3], Elisabetta Bugianesi[4], Anna Ludovica Fracanzani[5], Calogero Cammà[1], Jean-Baptiste Hiriart[6], Grace Lai-Hung Wong[2,3], Julien Vergniol[6], Anthony Wing-Hung Chan[7], Aurora Giannetti[1], Wassil Merrouche[6], Henry Lik-Yuen Chan[2,3], Brigitte Le-Bail[7,8], Rosa Lombardi[5], Salvatore Guastella[1], Antonio Craxì[1], Victor de Ledinghen[6,8].

**INSTITUTIONS**:

[1]Sezione di Gastroenterologia, Di.Bi.M.I.S., University of Palermo, Italy; [2]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong; [3]State Key Laboratory of Digestive Disease, The Chinese University of Hong Kong, Hong Kong; [4]Division of Gastroenterology, Department of Medical Sciences, University of Torino, Torino,Italy; [5]Department of Pathophysiology and Transplantation, Ca' Granda IRCCS Foundation, Policlinico Hospital, University of Milan, Italy; [6]Centre d'Investigation de la Fibrose hépatique, Hôpital Haut-Lévêque, Bordeaux University Hospital, Pessac, France; [7]Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Hong Kong; [8][7]INSERM U1053, Bordeaux University, Bordeaux, France; [8]Service de Pathologie, Hôpital Pellegrin, Bordeaux University Hospital, Bordeaux, France;

**CORRESPONDING AUTHOR**: Dr. Salvatore Petta, Section of Gastroenterology and Hepatology, Di.Bi.M.I.S., Policlinico Universitario Paolo Giaccone, Piazza delle Cliniche, 2, 90127 Palermo, Italy. Phone: +39 091 6552274. Fax +39 091 655 2156. E-mail: salvatore.petta@unipa.it.

**WORD COUNT:** 3970

**NUMBER OF FIGURES: 4**

**NUMBER OF TABLES/SUPPLEMENTAL TABLES: 4**/4

**LIST OF ABBREVIATIONS:** AUROC: Area Under ROC Curve; LSM: Liver Stiffness Measurement; NFS: NAFLD Fibrosis Score; NAFLD: Nonalcoholic fatty liver disease.

**CONFLICT OF INTEREST:** Vincent Wong, Grace Wong and Henry Chan have received lecture fees from Echosens.

**Author contributions:** S. Petta, V.W. Wong, E. Bugianesi, A.L. Fracanzani, C. Cammà, J-B Hiriart, G. Wong, J. Vergniol, A. Chan, A. Giannetti, W. Merrouche, H. Chan, B. Le-Bail, R. Lombardi, S. Guastella, A. Craxì, V. de Ledinghen had full control of the study design, data analysis and interpretation, and preparation of article. All authors were involved in planning the analysis and drafting the article. The final draft article was approved by all the authors.

**Abstract**

**Background:**Some evidence suggests an interference of obesity and alanine aminotransferase(ALT) levels on the diagnostic accuracy for advanced fibrosis of noninvasive tools like liver stiffness measurement(LSM) by FibroScan, FIB-4 and nonalcoholic fatty liver disease(NAFLD) fibrosis score(NFS).

**Aims:**We assessed whether the diagnostic accuracy of LSM, FIB-4 and NFS and of strategies based on these tools is affected by obesity and/or ALT levels.

**Methods:**We analyzed data from 968 patients with a histological diagnosis of NAFLD. The FIB-4, NFS and LSM by FibroScan were measured.

**Results:**LSM was better than both FIB-4 and NFS for staging F3-F4 fibrosis(AUC 0.863,0.777, and 0.765, respectively; p<0.001 for both), showing higher accuracy, higher negative predictive value (NPV) but lower positive predictive value (PPV). LSM worked less well in high ALT(>100 IU)(AUC 0.811 vs 0.877,p=0.04; PPV 57.5% vs 62.4%; NPV 90.7% vs 94%) or obese(AUC 0.786 vs 0.902,p<0.001; PPV 58.7% vs 64.8%; NPV 88.3% vs 95.2%) patients, this last picture being not affected by M or XL probe. Consistently, LSM worked better in terms of AUC and accuracy compared with both FIB-4 and NFS only in nonobese or high ALT patients, even if always keeping a slightly lower PPV value. The serial combination of FIB-4 or NFS with LSM as second test used in patients in the grey area of the first test kept -in most scenarios- similar PPV and NPV compared with LSM alone, while increasing the diagnostic accuracy of about 20% in all groups of patients, even if with a lower overall accuracy in obese(71.3% and 67.1% for FIB-4 and NFS as first test, respectively) compared to nonobese(81.9% and 82.4%for FIB-4 and NFS as first test, respectively) patients.

**Conclusion:**LSM has a better diagnostic accuracy for advanced fibrosis than both FIB-4 and NFS only in nonobese and/or low ALT patients. Serial combination strategies are better

than a single tool strategy regardless of obesity and ALT levels, though the accuracy is lower in obese patients.

**What is current knowledge**

- NAFLD Fibrosis Score (NFS), FIB-4, and liver stiffness measurement (LSM) by FibroScan are the most validated and best performing tools for the noninvasive diagnosis of advanced fibrosis in NAFLD.
- NFS and FIB-4 are not diagnostic in a relevant proportion of patients, while LSM can lead to false positive results, so their sequential combination is recommended.
- No data exist about the impact of obesity and ALT levels on the diagnostic accuracy for advanced fibrosis of NFS, FIB-4 and LSM in NAFLD.

**What is new here**

- LSM, and at a less extent NFS and FIB-4, have a higher diagnostic accuracy for advanced fibrosis in nonobese compared to obese NAFLD patients, LSM also better performing in patients with ALT≤100 IU compared to those with ALT >100.
- Overall LSM works better for diagnosis advanced fibrosis than both FIB-4 and NFS, even if this better performance was evident only in nonobese NAFLD patients.
- A serial combination of NFS or FIB-4 with LSM increases, regardless obesity and ALT levels, the diagnostic accuracy of about 20% respect to NFS or FIB-4 alone, even if with lower accuracy observed in obese NAFLD patients.

## Introduction

Traditionally considered the hepatic manifestation of the metabolic syndrome, non-alcoholic fatty liver disease (NAFLD) has dramatically increased in concert with the epidemics of both obesity and type 2 diabetes. It currently represents the most common liver disease in Western countries [1], with a burden of NAFLD-related cirrhosis about twice as high as that caused by chronic hepatitis C, and is projected to be the principal reason for hepatocellular carcinoma [2] and liver transplantation [3] within the decade. Furthermore, it is associated with an increased risk of extra-hepatic -mostly cardiovascular and cancer- morbidity and mortality [4].

Two retrospective cohort studies and a meta-analysis of the natural history of NAFLD patients have clearly shown that the severity of liver fibrosis estimated by liver histology is the strongest predictor not only of liver-related complications, but also of important extra-hepatic diseases, including cardiovascular disease and extra-hepatic malignancy [5-7]. Presently, liver biopsy is the gold standard for the assessment of liver fibrosis, however this procedure, because invasive, painful, and with potentially life-threatening complications cannot be implemented in all NAFLD individuals [8]. Consistently, different noninvasive scores/tools have been proposed to identify NAFLD patients with advanced fibrosis.

In this heterogeneous landscape, NAFLD fibrosis score (NFS) and FIB-4 -because based on easy-to-obtain clinico-metabolic variables, and liver stiffness measurement (LSM) by FibroScan -because not dangerous, easy-to-perform and widely diffused-, are the most available and validated noninvasive tools used in clinical practice to assess fibrosis severity in patients with NAFLD [9]. A recent meta-analysis reported that the diagnostic accuracy of LSM for advanced liver fibrosis is higher than that of NFS and FIB-4 [10]. However, while

the main limitation of both NFS and FIB-4 lies in the high proportion of patients falling in the grey area of the test, the most relevant clinical concern of LSM is represented by the risk of wrong classification mostly due to false positive results [11]. Consistently, EASL guidelines suggests to use LSM in patients with non-diagnostic noninvasive scores to improve the diagnostic accuracy [12].

Some studies suggested that obesity and severity of steatosis can affect the diagnostic accuracy of NFS and FIB-4 for advanced fibrosis in NAFLD [13]. Similarly, there are evidences about the interference of obesity [14,15], severity of steatosis [16] and skin-to-capsule distance [17] on the diagnostic accuracy of LSM for advanced fibrosis in NAFLD, as well as of increased transaminases levels in patients with liver diseases due to other etiologies [18,19].

Consistent to all the above, it is plausible that the diagnostic performance of clinical tests/strategies could differ according to the patient profile. For this purpose, we aimed to assess whether the diagnostic accuracy of LSM, FIB-4 and NFS and of strategies based on these tools is affected by obesity and/or ALT levels.

## MATERIALS & METHODS

### Patients

Data from 968 patients full filling the above reported inclusion criteria and prospectively recruited at the first diagnosis of biopsy-proven NAFLD were retrospectively reviewed and analyzed. Patients were recruited at the GI & Liver Unit of the University Hospital in Palermo (287 patients), at the University Hospital of Pessac in France (294 patients), at the Prince of Wales Hospital in Hong Kong (180 patients), at the Division of Gastroenterology Department of Medical Sciences University of Torino (142 patients), and

at the Department of Pathophysiology and Transplantation Ca' Granda IRCCS Foundation Policlinico Hospital University of Milan (65 patients), and with complete biochemical data and reliable LSM values.

Patients underwent liver biopsy for assessment of liver damage after ultrasonographic evidence of fatty liver. The diagnosis of NAFLD was based on alcohol consumption in the last year <20 g/day in females and <30 g/day in males, steatosis (≥5% of hepatocytes) at histology with/without necroinflammation and/or fibrosis. Exclusion criteria were as follows: (i) advanced cirrhosis (Child-Turcotte-Pugh B and C); (ii) hepatocellular carcinoma; (iii) other causes of liver disease or mixed etiologies (alcohol abuse, hepatitis C, hepatitis B, autoimmune liver disease, Wilson's disease, haemochromatosis or a1-antitrypsin deficiency); (iv) human immunodeficiency virus infection; (v) previous treatment with immunosuppressive drugs, and/or regular use of steatosis-inducing drugs, evaluated by a questionnaire (for example, corticosteroid, valproic acid, tamoxifen, amiodarone); or (vi) active intravenous drug addiction or use of cannabis.

The study was performed in accordance with the principles of the Declaration of Helsinki and its appendices, and with local and national laws. Approval was obtained from the hospital's Internal Review Boards and Ethics Committees, and written informed consent was obtained from all patients.

**Clinical, laboratory assessment, and histology**

Clinical and anthropometric data, including BMI, the presence of arterial hypertension and type 2 diabetes, were collected at the time of enrollment. The same day of liver biopsy, a 12-hour overnight fasting blood sample was drawn to determine serum levels of AST, ALT, GGT, PLT, albumin, total and HDL cholesterol, triglycerides, and plasma glucose concentration.

In each center, one liver-dedicated expert pathologist, who was unaware of patients' identity and history, coded and read histological slides. A minimum 15mm-length of the biopsy specimen or the presence of at least 10 complete portal tracts was required [20]. Steatosis was assessed as the percentage of hepatocytes containing fat droplets (minimum 5%), and as a categorical variable. Kleiner classification [21] was used to stage fibrosis from 0 to 4.

**Non-invasive fibrosis algorithms/tools**

The FIB-4 (age, AST, ALT, PLT), and NFS (age, IFG/Diabetes, BMI, PLT, albumin, AST/ALT) were calculated using the original reported formulas [22,23].

Transient elastography was performed with the FibroScan (Echosens, Paris, France) medical device, using the M probe (also named as standard probe). In a subgroup of patients both M and XL FibroScan probes were used because available. In each center, LSM was assessed on the same day of liver biopsy, before the procedure and after an overnight fast, by a trained operator who had previously performed at least 300 determinations in patients with chronic liver disease. As recently reported in the literature [24], we classified all LSM examinations into three reliability categories: "very reliable" (IQR/M ≤0.10), "reliable" (0.10< IQR/M ≤0.30, or IQR/M >0.30 with LSM median <7.1 kPa), and "poorly reliable" (IQR/M >0.30 with LSM median ≥7.1 kPa). Only patients with 10 valid measurements were included, and "poorly reliable" results were excluded from the analysis.

**Statistics**

Continuous variables were summarized as mean ± SD, and categorical variables as frequency and percentage.

The accuracy of each score for detection of advanced fibrosis (F3-F4) was assessed using receiver operating characteristic curves describedas AUC with 95% confidence intervals (95% CI). A patient was assessed as positive or negative according to whether the noninvasive marker value was greater than, less than, or equal to a given cut-off value. Connected with any cut-off value is the probability of a true positive (sensitivity) and the probability of a true negative (specificity). AUCs for both paired and unpaired curves are compared using the bootstrap method, with non-parametric resampling and with the percentile method, as described by Carpenter and Bithell (2000) [25], with 2000 replicates as recommended by the same authors. Cut-off points of LSM, NFS and FIB-4 for the F3-F4 model were derived from literature. Specifically, for LSM, cut-offs of <7.9 KPa and of ≥9.6 KPa for M probe, and of <5.7 KPa and of ≥9.3 KPa were used to rule-out and rule-in, respectively, severe fibrosis [23,26]; for NFS, cut-offs of <-1.455 and of >0.676 were used to rule-out and rule-in, respectively, severe fibrosis [20]; and for FIB-4, cut-offs of <1.30 KPa and of >2.67 were used to rule-out and rule-in, respectively, severe fibrosis [21]. Accordingly, false negative and false positive rates of the single test, and of their combination, as well as sensitivity, specificity, positive likelihood ratio (LHR), negative LHR, positive predictive value (PPV), and negative predictive value (NPV) are calculated.

Analysis was performed by R 3.5.1 [27].

## RESULTS

### Patient characteristics and histology

The baseline characteristics of the 968 NAFLD patients are shown in Table 1. Mean age was 50 years, with male preponderance (62.9%). Thirty-nine percent of patients were obese, diabetes was present in 37% of cases, and 20.7% of patients had ALT >100 IU.

At liver biopsy, 28.5% of patient had fibrosis ≥ 3 by Kleiner score.

**Is liver stiffness measurement better than NFS or FIB-4 for the diagnosis of severe liver fibrosis in the entire cohort?**

Figure 1 shows the accuracy, in terms of AUC, of the different noninvasive tools to detect fibrosis ≥ F3 in the entire cohort of 968 patients after excluding 99 subjects with invalid LSM (age 54.7±11.9 years, males 57.5%, BMI 36.1±6.5 Kg/m$^2$, obese 84.8%, ALT 68±50.3 IU, ALT>100 IU 15.1%, advanced fibrosis 42.4%).

In the entire cohort, LSM worked better than both FIB-4 and NFS ($p<0.001$ for both) with AUC values of 0.863, 0.777, and 0.765, respectively, while no differences were observed between NFS and FIB-4 ($p=0.32$) (Table 2). LSM had the highest accuracy and the highest negative predictive value (93.5%; 95% C.I. 91.4%-95.7%), while FIB-4 and NFS the higher positive predictive value (72.1%, 95% C.I. 62.9%-81.8% for FIB-4; 67.2%, 95% C.I. 56.1%-79.2% for NFS) (Table 3).

**Do Obesity and ALT levels affect the diagnostic performance of liver stiffness measurement, FIB-4 and NFS?**

Figure 2 and table 2 show the accuracy, in terms of AUC, of the different noninvasive tools to detect fibrosis ≥ F3 in sub-groups of patients according to obesity and/or ALT (>100 IU, third quartile) levels. The prevalence of advanced fibrosis was 38.2% (144/377) in obese, 22.3% (132/591) in nonobese, 31.8% (64/201) in high ALT, and 27.6% (212/767) in low ALT patients.

The diagnostic performance of LSM for advanced fibrosis, in terms of AUC, was better in nonobese compared to obese (0.902 vs 0.786, $p<0.001$) (Figure 2A and Table 2) patients, and in subjects with ALT levels ≤100 IU compared to their counterpart (0.877 vs 0.811, $p=0.04$) (Figure 2B and Table 2). Consistent with these data, PPV and NPV were

higher (64.8% and 95.2%, respectively) in nonobese compared with obese (58.7% and 88.3%, respectively) subjects, and in patients with ALT levels ≤100 IU (62.4% and 94%, respectively) compared to those with ALT>100 IU (57.7% and 90.7%, respectively) (Table 3). A similar trend was observed for accuracy (Table 3). When splitting patients according to both BMI and ALT levels, the AUC for advanced fibrosis progressively increased from the worst performance (patients with BMI>30 and ALT >100 IU (AUC 0.759) to the best one for BMI≤30 and ALT≤100 IU (AUC 0.916) (Figure 2C and Supplemental Table 1). As a consequence, PPV and NPV were better in the best (66.7% and 96.4%, respectively) respect to the worst (55% and 86.4%, respectively) class (Supplemental Table 2).

The lower accuracy of LSM in obese patients could raise the issue that this picture is related to the use of M probe instead of XL probe. To solve this concern, in a subgroup of 244 patients (55.7% males, mean age 54.3±12.6 years, 50.8% obese, 40.1% with advanced fibrosis) with available and reliable LSM by both M (mean LSM 12±8.9 KPa) and XL (mean LSM 9.2±6.8 KPa) probes, we compared the diagnostic performance of LSM according to probe and obesity. The AUC of LSM by M and XL probe was similar in the entire cohort of 244 patients (0.815 vs 0.812, p=0.86) (Supplemental Figure 1A) as well as in obese (0.745 vs 0.760, p=0.64) (Supplemental Figure 1B) and nonobese (0.892 vs 0.865, p=0.19) (Supplemental Figure 1C) patients, similar results being observed for accuracy (Supplemental Table 3). Consistently, when splitting the population according to BMI, the AUC was lower in obese compared to nonobese patients by using both M (0.745 vs 0.892, p=0.007) (Supplemental Figure 2A) and XL (0.760 vs 0.865, p=0.06) (Supplemental Figure 2B) probes, similar results were observed for accuracy (Supplemental Table 3).

The AUC of FIB-4 for advanced fibrosis was higher but not statistically significant in nonobese when compared to obese patients (0.800 vs 0.742, p=0.10) (Figure 2D and Table 2). When considering ALT levels the AUC were similar in patients with ALT ≤100 IU and in

those with ALT>100 IU (0.789 vs 0.737, p=0.228) (Figure 2E and Table 2). Along this line, PPV and NPV were better in nonobese (74.5% and 90%, respectively) respect to obese (69.2% and 77.9%, respectively) patients, while being similar in subjects with ALT ≤100 IU (73.1% and 86.4%, respectively) or >100 IU (68.4% and 81.8%, respectively) (Table 3). A similar trend was reported for accuracy (Table 3).

Similar to FIB-4, the AUC of NFS for advanced fibrosis was higher but not statistically significant in nonobese when compared to obese patients (0.767 vs 0.718, p=0.18) (Figure 2F and Table 2). When considering ALT levels the AUC were similar in patients with ALT ≤100 IU and in those with ALT>100 IU (0.774 vs 0.753, p=0.85) (Figure 2G and Table 2). Along this line, PPV and NPV were better in nonobese (76.5% and 89.7%, respectively) respect to obese (64% and 80.5%, respectively) patients, while being similar in subjects with ALT ≤100 IU (67.2% and 88.5%, respectively) or >100 IU (66.7% and 82.4%, respectively) (Table 3). A similar trend was reported for accuracy (Table 3).

When looking at the 99 patients (84.8% obese) failure to LSM the observed diagnostic performance of noninvasive scores for advanced fibrosis [NFS: AUC 0.718 (0.616-0.820), PPV 68%, NPV 85%, accuracy 34%, uncertainty area 54.5%, wrong classification rate 11.15; FIB-4: AUC 0.722 (0.623-0.820), PPV 100%, NPV 63.5%, accuracy 58%, uncertainty area 11.1%, wrong classification rate 31.3%] was similar to that observed in the before reported obese group of patients with reliable liver stiffness.

**Is the diagnostic performance of liver stiffness measurement better than that of FIB-4 and NFS according to BMI and ALT levels?**

LSM was confirmed better in terms of AUC when compared with both FIB-4 and NFS in patients without obesity (AUC 0.902, 0.800, 0.767, respectively; p<0.001 for both), and in those with ALT ≤100 IU (AUC 0.877, 0.789, 0.774, respectively; p<0.001 for both), while not

in their counterparts except that compared with NFS also in obese (Figure 3, Table 2). Consistent with these data, while keeping lower PPV and higher NPV respect to both FIB-4 and NFS independently of obesity, accuracy, uncertainty area and wrong classification rate of LSM were better than those of FIB-4 and NFS in nonobese patients, LSM being superior of NFS also in obese (Table 3). Similarly, while keeping lower PPV and higher NPV respect to both FIB-4 and NFS independently of ALT levels, LSM had better accuracy, uncertainty area and wrong classification rate in subjects with ALT≤100 IU, while being similar in the other groups (Table 3).

Consistent with all the above, when splitting patients according to both BMI and ALT levels, the AUC of LSM for advanced fibrosis was significantly higher than that of both FIB-4 and NFS in patients in the best class (BMI≤30 and ALT≤100 IU; AUC 0.916 for LSM, 0.816 for FIB-4, 0.783 for NFS; p<0.001 for both), but not in those in the worst class (BMI>30 and ALT >100 IU, AUC 0.759 for LSM, 0.715 for FIB-4, 0.741 for NFS; p n.s. for both) (Figures 3E and 3F, Supplemental Table 1). As a consequence, while keeping lower PPV and higher NPV respect to both FIB-4 and NFS in all classes, in the best class LSM had accuracy, uncertainty area and wrong classification rate better than both FIB-4 and NFS, while no differences were observed in the worst class (Supplemental Table 2).

**Is serial combination strategy of NFS or FIB-4 with LSM better than one test strategy according to BMI and ALT levels?**

In the entire cohort of NAFLD patients, the AUC of the logistic model including LSM and FIB-4 was better than that of LSM and FIB-4 alone (0.877 vs 0.863 and vs 0.777, p=0.01 and p<0.001 respectively), while the AUC of the logistic model including LSM and NFS was better than NFS alone but similar than that of LSM (0.867 vs 0.863 and vs 0.765, p<0.001 and p=0.60 respectively). Consistently, the use of LSM in patients who fell in the uncertainty

area of FIB-4 generated higher accuracy and lower grey area than LSM or FIB-4 alone (Table 4). The algorithm also generated PPV and a NPV of 60.7% and 91.0%, respectively (Table 4). They were similar to those obtained by LSM alone, while NPV was higher and PPV lower than FIB-4 alone (Table 4). Similarly, the use of LSM in patients who fell in the uncertainty area of NFS generated higher accuracy and lower grey area than LSM or NFS alone (Table 4). The algorithm also generated a PPV and a NPV of 65.7% and 88.7%, respectively (Table 4). They were similar to those obtained by NFS alone, while NPV was lower and PPV higher than LSM alone (Table 4).

Due to the evidence that LSM works better than NFS and FIB-4 for the diagnosis of advanced fibrosis only in patients without obesity and/or in those with ALT ≤100 IU, we tested whether the serial combination strategy of FIB-4 or NFS with LSM is always better than FIB-4 or LSM alone.

When splitting the cohort according to obesity or ALT levels we confirmed that in all subgroups the serial combination strategy increased the diagnostic performance for advanced fibrosis at a similar extent than in the entire population, even if the strategy worked overall better in nonobese (PPV 61.6% and 59.3%, NPV 93.1% and 94.7% for FIB-4 and NFS, respectively) compared to obese (PPV 59.8% and 71.5%, NPV 86.7% and 76.2% for FIB-4 and NFS, respectively) patients. In these subgroups of patients, combination strategies generated PPV and NPV similar to those obtained by LSM alone, but higher NPV and lower PPV compared with FIB-4 or NFS alone, except that for NFS in obese patients where the combination strategy had lower both PPV and NPV than NFS alone (Table 4). Along this line, the diagnostic performance of serial combination strategies was better in nonobese patients with ALT ≤100 IU (PPV 67.0% and 65.5%, NPV 94.4% and 94.1% for FIB-4 and NFS, respectively) respect to obese patients with ALT>100 IU (PPV 57.1% and 66.7%, NPV 80% and 82.6% for FIB-4 and NFS, respectively) (Supplemental Table 4).

## DISCUSSION

In this study on a large cohort of patients with histological diagnosis of NAFLD, we confirmed that LSM is better than FIB-4 and NFS for diagnosing advanced fibrosis, but this superiority is mostly observed in nonobese patients and/or in those with low ALT levels, while it is partially lost in subjects obese and/or with high ALT levels. We also demonstrated that a serial combination of FIB-4 or NFS with LSM improves overall and in all subgroups the diagnostic performance for advanced fibrosis, with the higher accuracy observed in nonobese and the lower in obese patients.

LSM, FIB-4 and NFS are the mostly available, used and validated nonivasive tools aimed to identify NAFLD patients with advanced fibrosis –i.e. at risk of hepatic and extra-hepatic complications-. In our cohort, we reported that LSM has a better diagnostic accuracy than both FIB-4 and NFS for diagnosis advanced fibrosis in NAFLD, this result largely confirming what reported in a recent meta-analysis [10].

In the present study we found that LSM is more accurate for diagnosing advanced fibrosis in nonobese patients and/or in those with lower ALT levels, compared to obese and/or high ALT subjects, the accuracy ranging from 50% to 76.9%. We previously reported that BMI could affect the performance of LSM for the diagnosis of advanced fibrosis by increasing the rate of false positive results [15, 28]; along this line Caussy and colleagues identified in BMI the main reason of disagreement between LSM and magnetic resonance elastography for staging fibrosis in NAFLD [16]. However, in 315 Asiatic NAFLD patients, LSM showed consistent diagnostic performance for advanced fibrosis, regardless of obesity [13]. Differences in baseline characteristics of populations and in prevalence of liver disease severity could explain the discrepancies among the observed results, even if the high

number of patients enrolled in our study could make a conclusive point about this topic. However, our results can raise the doubt that the lower accuracy of LSM in nonobese patients could be due to the use of M instead of XL probe. To solve this question we compared the accuracy of M and XL probes in a subgroup of patients, confirming that the two probes have similar accuracy, and that both had worse diagnostic performance in obese compared to nonobese patients. Further studies assessing skin-to-capsule distance, could add insights about this topic.

Evidence in chronic hepatitis B and C already demonstrated that high ALT levels affect the accuracy of LSM for fibrosis by overestimating liver damage [18,19]. However, we firstly confirmed this picture in a population of patients with chronic liver disease due to NAFLD. We also showed that, while ALT levels did not interfere with diagnostic ability of both FIB-4 and NFS –except that for a lower PPV by using FIB-4-, these last performed less well in obese patients with an overall accuracy ranging from 41.4% to 65.5%. These data agree with what suggested in a smaller study on an Asiatic population of NAFLD patients [13]. Our evidence about lower PPV and NPV of FIB-4 and NFS in obese patients, as well as the lower PPV of FIB-4 in high ALT patients, is not expression of a biological phenomenon. It can be explained by the fact that these variables –included in the scores- are associated between them and with advanced fibrosis, but are also present in absence of advanced fibrosis sometime lowering the accuracy of noninvasive scores.

A relevant feature from our study is that LSM works better than both FIB-4 and NFS in obese and/or low ALT patients, providing a gain in accuracy of 10%-15% mostly due to reduction in the proportion of patients falling in the uncertainty area. Conversely, the diagnostic performance of LSM was similar to that of both NFS and FIB-4 in high ALT patients, and similar to that of FIB-4 in obese patients, while remaining slightly superior then NFS in obese.

EASL guidelines on NAFLD [12] recommend to noninvasively assess liver fibrosis by using as first test a simple and available tool like FIB-4 or NFS, then applying LSM only in patients with noninvasive tests in the grey area. In the present study, we confirmed overall the superiority of a serial combination strategy respect to the use of a singular test, but, due the effect of obesity and/or ALT levels on the diagnostic performance of especially LSM, we investigated whether the combination strategy is to be preferred regardless obesity and ALT levels. Notably we found that a serial combination strategy is always better than the use of only one test by providing an increase in accuracy of about 20% mostly due to reduction in the uncertainty area, also showing a better perfomance in nonobese compared to obese patients with an accuracy ranging from 67% to 84%.

From a clinical point of view, our study shows that in NAFLD patients, when using only one test for diagnosing advanced fibrosis, LSM should be preferred to both FIB-4 and NFS in obese and/or low ALT patients, while it does not provide any advantage in high ALT patients, and provides a slight advantage in obese patients only when compared to NFS. In a setting where LSM is not available/reliable, clinicians should be aware that overall the proportion of patients in the uncertainty area will be higher than with LSM, and that the diagnostic accuracy in terms of PPV and NPV will be worse in obese patients. When applying a serial combination strategy as recommended by EASL guidelines, it should be the preferred strategy in all categories of patients, however taking into account a lower overall accuracy in obese patients (Figure 4). Notably, this strategy, in our cohort, leaded to avoid 72.2% of liver biopsy by using FIB-4 as first test in both obese and nonobese patients, and 73.3% of liver biopsy by using NFS as first score in nonobese and FIB-4 in obese patients.

This study has limitations. First is the potentially limited validity of the results in different populations and settings. It is plausible that the performance of the proposed

algorithms could change according to characteristics of patients, and to the prevalence of obesity, high ALT levels and advanced fibrosis. Criteria for biopsy selection and lack of data about patients who underwent liver biopsy but without confirmation of NAFLD could further affect the interpretation of our results and the validity of the proposed algorithm. Furtehrmore, our strategies should be well validated in general population where the prevalence of advanced fibrosis is low. Consistent with these criticisms, to strength our study, in our analysis we used published standardized cut-offs and not cut-offs calculated from data of our populations, and we used bootstrap method as internal validation. Other limitation of our study lie in the fact that interobserver concordance of LSM examination was not assessed, this issue potentially affecting the interpretation of our results. However, all tests were performed by expert operators following the same protocol and fulfilling validity criteria. Moreover, different relevant studies assessing FibroScan in NAFLD were based on multicenter cohorts and/or on multiple operators [29,30]. Finally, many different studies reported good interobserver concordance for LSM [31,32]. Another relevant limitation of our study is the lack of interobserver agreement assessment among pathologists for liver histology. This issue could strongly affect the clinical interpretation of our results because fibrosis by histology is the assumed gold standard for the outcome in our study. Consistently, even if data from the literature [21] and from our group [33] clearly reported that overall interobserver agreement for staging severe fibrosis in NAFLD is good, and again different relevant studies assessing FibroScan in histological-defined NAFLD were based on multicenter cohorts [11,29,30], we cannot exclude that lack of central reading does not affect our results. Finally, the lack of inclusion of patients with unreliable liver stiffness could overestimate the diagnostic accuracy of the proposed algorithms, this issue being a significant limitation of the proposed approach.

In conclusion, we demonstrated that obesity and/or ALT levels affect the diagnostic accuracy of noninvasive tools, especially LSM, this issue leading to a better accuracy of LSM respect to both FIB-4 and NFS only in nonobese and/or patients with low ALT levels. We also observed that serial combination strategies are better than a single tool strategy regardless of obesity and ALT levels, even though the accuracy was lower in obese patients.

**References**


1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global Epidemiology of Non-Alcoholic Fatty Liver Disease-Meta-Analytic Assessment of Prevalence, Incidence and Outcomes. Hepatology. 2015 Dec 28. doi: 10.1002/hep.28431.
2. Dyson J, Jaques B, Chattopadyhay D, et al. Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team. J Hepatol. 2014 Jan;60(1):110-7.
3. Wong RJ, Aguilar M, Cheung R, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. Gastroenterology. 2015 Mar;148(3):547-55.
4. Adams LA, Anstee QM, Tilg H, Targher G. Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. Gut. 2017 Jun;66(6):1138-1153.
5. Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology. 2015 May;61(5):1547-54.
6. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver Fibrosis, but no Other Histologic Features,

Associates with Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2015 Apr 29. pii: S0016-5085(15)00599-5.

7. Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. R. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology. 2017 May;65(5):1557-1565.
8. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al; LIDO Study Group. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005 Jun;128(7):1898-906.
9. Maida M, Macaluso FS, Salomone F, Petta S. Non-Invasive Assessment of Liver Injury in Non-Alcoholic Fatty Liver Disease: A Review of Literature. Curr Mol Med. 2016;16(8):721-737.
10. Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. Hepatology. 2017 Nov;66(5):1486-1501.
11. Petta S, Wong VW, Cammà C, Hiriart JB, Wong GL, Vergniol J, et al. Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver fibrosis in patients with NAFLD. Aliment Pharmacol Ther. 2017 Sep;46(6):617-627.
12. 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 Jun;64(6):1388-402.

13. Joo SK, Kim W, Kim D, Kim JH, Oh S, Lee KL, et al. Steatosis severity affects the diagnostic performances of noninvasive fibrosis tests in nonalcoholic fatty liver disease. Liver Int. 2018 Feb;38(2):331-341.
14. Petta S, Maida M, Macaluso FS, Di Marco V, Cammà C, Cabibi D, et al. The severity of steatosis influences liver stiffness measurement in patients with nonalcoholic fatty liver disease. Hepatology. 2015 Oct;62(4):1101-10.
15. Petta S, Di Marco V, Cammà C, Butera G, Cabibi D, Craxì A. Reliability of liver stiffness measurement in non-alcoholic fatty liver disease: the effects of body mass index. Aliment Pharmacol Ther. 2011 Jun;33(12):1350-60.
16. Caussy C, Chen J, Alquiraish MH, Cepin S, Nguyen P, Hernandez C, et al. Association Between Obesity and Discordance in Fibrosis Stage Determination by Magnetic Resonance vs Transient Elastography in Patients with Non-alcoholic Liver Disease. Clin Gastroenterol Hepatol. 2017 Nov 2. pii: S1542-3565(17)31304-6. doi:10.1016/j.cgh.2017.10.037.
17. Shen F, Zheng RD, Shi JP, Mi YQ, Chen GF, Hu X, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. Liver Int. 2015 Nov;35(11):2392-400.
18. Tapper EB, Cohen EB, Patel K, Bacon B, Gordon S, Lawitz E, et al. Levels of alanine aminotransferase confound use of transient elastography to diagnose

fibrosis in patients with chronic hepatitis C virus infection. Clin Gastroenterol Hepatol. 2012 Aug;10(8):932-937.e1.

19. Coco B, Oliveri F, Maina AM, Ciccorossi P, Sacco R, Colombatto P, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. J Viral Hepat. 2007 May;14(5):360-9.
20. Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. J Hepatol 2003;39:239–44.
21. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005; 411:313–21.
22. 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: 1265-9.
23. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology. 2007 Apr;45(4):846-54.
24. Boursier J, Zarski JP, de Ledinghen V, et al; Multicentric Group from ANRS/HC/EP23 FIBROSTAR Studies. Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology. 2013;57:1182-91.
25. James Carpenter and John Bithell (2000) “Bootstrap condence intervals: when, which, what? A practical guide for medical statisticians”. Statistics in Medicine 19, 1141–1164.DOI:10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F.

26. Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. Am J Gastroenterol. 2012 Dec;107(12):1862-71.
27. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
28. Wong GL, Chan HL, Choi PC, Chan AW, Lo AO, Chim AM, Wong VW. Clin Gastroenterol Hepatol. 2013 Mar;11(3):295-302.
29. Wong VW, Vergniol J, Wong GL, *et al*. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010; 51: 454-62.
30. Petta S, Wong VW, Cammà C, Hiriart JB, Wong GL, Marra F, et al. Improved Noninvasive prediction of Liver Fibrosis by Liver Stiffness Measurement in Patients with Nonalcoholic Fatty Liver Disease Accounting for Controlled Attenuation Parameter Values. Hepatology. 2017 Apr;65(4):1145-1155.
31. Afdhal NH, Bacon BR, Patel K, Lawitz EJ, Gordon SC, Nelson DR, et al. Accuracy of fibroscan, compared with histology, in analysis of liver fibrosis in patients with hepatitis B or C: a United States multicenter study. Clin Gastroenterol Hepatol. 2015 Apr;13(4):772-9.e1-3.
32. Boursier J, Konate A, Guilluy M, Gorea G, Sawadogo A, Quemener E, et al. Learning curve and interobserver reproducibility evaluation of liver stiffness

measurement by transient elastography. Eur J Gastroenterol Hepatol. 2008 Jul;20(7):693-701.

**33.** Petta S, Valenti L, Marra F, Grimaudo S, Tripodo C, Bugianesi E, et al. MERTK rs4374383 polymorphism affects the severity of fibrosis in non-alcoholic fatty liverdisease. J Hepatol. 2016 Mar;64(3):682-90.

**Table 1.** Characteristics of the population of 968 patients with histological diagnosis of nonalcoholic fatty liver disease.

| Variable | Entire Cohort n=968 | Fibrosis F0-F2 n=692 | Fibrosis F3-F4 n=276 | P value |
|---|---|---|---|---|
| **Male Gender (%)** | 62.9 | 66.3 | 54.3 | 0.001 |
| **Age - years** | 50.1 (13.0) | 47.5 (12.7) | 56.4 (11.6) | <0.001 |
| **BMI – Kg/m$^2$** | 29.3 (4.7) | 28.8 (4.8) | 30.5 (4.3) | <0.001 |
| **Fasting Glucose – mg/dL** | 109.6 (46.9) | 104.4 (47.0) | 122.8 (44.1) | <0.001 |
| **Impaired Fasting Glucose (%)** | 51.9 | 43.1 | 73.9 | <0.001 |
| **Diabetes (%)** | 37.0 | 27.7 | 60.1 | <0.001 |
| **Arterial Hypertension (%)** | 39.4 | 32.4 | 56.9 | <0.001 |
| **Total Cholesterol – mg/dL** | 201.2 (46.5) | 205.4 (45.3) | 190.8 (47.7) | <0.001 |
| **Triglycerides – mg/dL** | 159.1 (100.7) | 154.2 (96.9) | 171.6 (109.1) | 0.017 |
| **HDL Cholesterol – mg/dL** | 50.5 (20.2) | 51.6 (20.9) | 47.7 (18.1) | 0.008 |
| **AST – IU** | 46.1 (37.4) | 41.2 (31.6) | 58.3 (46.9) | <0.001 |
| **ALT – IU** | 76.1 (55.9) | 74.1 (48.3) | 81.1 (71.2) | 0.082 |
| **Platelets – X10$^3$/mmc** | 229.7 (69.7) | 238.29 (64.1) | 208.44 (78.3) | <0.001 |
| **Albumin – g/dL** | 4.5 (0.4) | 4.5 (0.4) | 4.4 (0.4) | 0.001 |
| **Total Bilirubin – mg/dL** | 0.7 (0.4) | 0.7 (0.4) | 0.7 (0.3) | 0.729 |
| **INR** | 1.01 (0.18) | 0.99 (0.12) | 1.07 (0.26) | <0.001 |
| **Liver Stiffness Measurement - KPa** | 10.0 (8.2) | 7.5 (4.3) | 16.3 (11.7) | <0.001 |

Data are presented as mean (standard deviation), or as percentage. Abbreviations: BMI, body mass index; HDL, high density lipoprotein; AST, aspartate aminotransferase; ALT, alanine aminotransferase.

**Table 2.** Area Under the Receiver Operating Characteristics Curves for Transient Elastography and Noninvasive Markers for the Diagnosis of Advanced Fibrosis in the Entire Cohort of Patients with Nonalcoholic Fatty Liver Disease, and in Sub-groups According to BMI or ALT levels.

| Tools | ROC AUC (C.I.) |
|---|---|
| | |
| **Entire Cohort (n=968; F3-F4 fibrosis 28.5%)** | |
| **Liver Stiffness Measurement** | 0.863 (0.84-0.886) |
| **FIB-4** | 0.777 (0.746-0.809) |
| **NAFLD Fibrosis Score** | 0.765 (0.733-0.799) |
| | |
| **BMI<30 (n=591; F3-F4 fibrosis 22.3%)** | |
| **Liver Stiffness Measurement** | 0.902 (0.876-0.928) |
| **FIB-4** | 0.8 (0.753-0.844) |
| **NAFLD Fibrosis Score** | 0.767 (0.717-0.813) |
| | |
| **BMI>=30 (n=377; F3-F4 fibrosis 38.2%)** | |
| **Liver Stiffness Measurement** | 0.786 (0.739-0.829) |
| **FIB-4** | 0.745 (0.697-0.794) |
| **NAFLD Fibrosis Score** | 0.719 (0.662-0.767) |
| | |
| **ALT<=100 (n=767; F3-F4 fibrosis 27.6%)** | |
| **Liver Stiffness Measurement** | 0.877 (0.852-0.904) |
| **FIB-4** | 0.789 (0.753-0.822) |
| **NAFLD Fibrosis Score** | 0.774 (0.737-0.81) |
| | |
| **ALT>100 (n=201; F3-F4 fibrosis 31.8%)** | |
| **Liver Stiffness Measurement** | 0.811 (0.748-0.868) |
| **FIB-4** | 0.737 (0.664-0.811) |
| **NAFLD Fibrosis Score** | 0.753 (0.673-0.83) |

**Table 3.** Accuracy of Fibroscan, FIB-4 and NAFLD Fibrosis Score for the Diagnosis of Severe Fibrosis in the Entire Cohort of Patients with Nonalcoholic Fatty Liver Disease and in Sub-groups According to BMI or ALT Levels.

| Tools | Cut-off | Sensitivity (%) | Specificity (%) | Positive predicted value (%) | Negative predicted value (%) | Positive likelihood ratio | Negative likelihood ratio | Uncertainty area (%) | Accuracy (%) | Wrong classification (%) | False positive (%) | False negative (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Entire cohort (n=968; F3-F4 fibrosis 28.5%)** | | | | | | | | | | | | |
| **LSM** | 9,6 | 72.5 (67.5-77.9) | 81.8 (79.2-84.7) | 61.3 (56.2-66.9) | 88.2 (85.8-90.7) | 4 (3.1-4.6) | 0.3 (0.2-0.4) | | | | 18.2 (15.3-20.8) | 27.5 (22.1-32.5) |
| | 7,9 | 88.4 (84.8-92.3) | 66.5 (62.9-70.1) | 51.3 (46.8-55.7) | 93.5 (91.4-95.7) | 2.6 (2.3-2.9) | 0.2 (0.1-0.2) | 13.9 (11.8-16.1) | 68.2 (65.4-71.2) | 16.3 (14-18.6) | 33.5 (29.9-37.1) | 11.6 (7.7-15.2) |
| **FIB-4** | 2,67 | 22.5 (17.4-27.2) | 96.5 (95.2-98) | 72.1 (62.9-81.8) | 75.7 (73-78.5) | 6.5 (2.2-8.7) | 0.8 (0.8-0.9) | | | | 3.5 (2-4.8) | 77.5 (72.8-82.6) |
| | 1,3 | 67.8 (62.3-73.6) | 76.3 (73-79.4) | 53.3 (47.9-58.8) | 85.6 (82.8-88.4) | 2.9 (2.3-3.3) | 0.4 (0.3-0.5) | 27.4 (24.6-30.1) | 61 (57.9-64) | 11.7 (9.6-13.6) | 23.7 (20.6-27) | 32.2 (26.4-37.7) |
| **NFS** | 0,676 | 16.3 (11.9-20.5) | 96.8 (95.6-98.2) | 67.2 (56.1-79.2) | 74.4 (71.5-77.2) | 5.1 (1.5-7.1) | 0.9 (0.8-0.9) | | | | 3.2 (1.8-4.4) | 83.7 (79.5-88.1) |
| | -1,455 | 74.3 (69.1-79.5) | 69.7 (66.1-73.2) | 49.4 (44.7-54.5) | 87.2 (84.4-89.9) | 2.5 (2.1-2.8) | 0.4 (0.3-0.4) | 36 (32.9-38.8) | 54.4 (51.1-57.5) | 9.6 (7.9-11.5) | 30.3 (26.8-33.9) | 25.7 (20.5-30.9) |
| **BMI<30 (n=591; F3-F4 fibrosis 22.3%)** | | | | | | | | | | | | |
| **LSM** | 9,6 | 69.7 (61.9-77.5) | 89.1 (86.4-92.1) | 64.8 (57.2-72.7) | 91.1 (88.4-93.7) | 6.4 (4.2-7.9) | 0.3 (0.3-0.4) | | | | 10.9 (7.9-13.6) | 30.3 (22.5-38.1) |
| | 7,9 | 86.4 (80.7-92.6) | 77.1 (73.4-81.2) | 52.1 (45.5-58.7) | 95.2 (93-97.5) | 3.8 (3-4.4) | 0.2 (0.1-0.3) | 10.8 (8.3-13.5) | 75.5 (72.1-79) | 11.5 (8.8-14) | 22.9 (18.8-26.6) | 13.6 (7.4-19.3) |
| **FIB-4** | 2,67 | 26.5 (18.7-33.8) | 97.4 (96.1-99) | 74.5 (63.2-88.4) | 82.2 (78.9-85.5) | 10.1 (0.3-17.9) | 0.8 (0.7-0.8) | | | | 2.6 (1-3.9) | 73.5 (66.2-81.3) |
| | 1,3 | 70.5 (63.1-78.6) | 76.7 (72.8-80.6) | 46.5 (39.2-53.3) | 90 (87.3-93.2) | 3 (2.3-3.6) | 0.4 (0.3-0.5) | 25.9 (21.4-36.9) | 65.5 (61.6-69.2) | 8.6 (6.4-10.8) | 23.3 (19.4-27.2) | 29.5 (21.4-36.9) |
| **NFS** | 0,676 | 9.8 (4.7-14.6) | 99.1 (98.5-100) | 76.5 (58.2-99.1) | 79.3 (75.8-82.7) | 11.3 (4.08-50.47) | 0.9 (0.9-1) | | | | 0.9 (0-1.5) | 90.2 (85.4-95.3) |
| | -1,455 | 68.9 (60.7-76.8) | 78 (74.3-81.8) | 47.4 (40.3-54.6) | 89.7 (86.7-92.8) | 3.1 (2.4-3.7) | 0.4 (0.3-0.5) | 29.6 (25.9-33.2) | 62.8 (58.9-66.8) | 7.6 (5.4-9.6) | 22 (18.2-25.7) | 31.1 (23.2-39.3) |
| **BMI>=30 (n=377; F3-F4 fibrosis 38.2%)** | | | | | | | | | | | | |
| **LSM** | 9,6 | 75 (68.1-82.4) | 67.4 (61.4-73.7) | 58.7 (51.7-65.7) | 81.3 (76.2-87) | 2.3 (1.7-2.7) | 0.4 (0.3-0.5) | | | | 32.6 (26.3-38.6) | 25 (17.6-31.9) |
| | 7,9 | 90.3 (85.9-95) | 45.5 (39.2-51.5) | 50.6 (44.5-56.9) | 88.3 (83.1-94) | 1.7 (1.4-1.8) | 0.2 (0.1-0.3) | 18.8 (14.9-22.8) | 56.8 (52-61.8) | 23.9 (19.6-28.1) | 54.5 (48.5-60.8) | 9.7 (5-14.1) |
| **FIB-4** | 2,67 | 18.8 (12.5-24.9) | 94.8 (92.2-97.7) | 69.2 (55.1-83.7) | 65.4 (60.5-70.7) | 3.6 (2.1-8) | 0.9 (0.8-0.9) | | | | 5.2 (2.3-7.8) | 81.2 (75.1-87.5) |
| | 1,3 | 65.3 (57.3-73.1) | 75.5 (70-81.3) | 62.3 (54.3-70) | 77.9 (72.3-83.4) | 2.7 (1.9-3.2) | 0.5 (0.4-0.6) | 29.7 (24.9-34.2) | 53.9 (49.1-58.6) | 16.5 (12.5-19.9) | 24.5 (18.7-30) | 34.7 (26.9-42.7) |
| **NFS** | 0,676 | 22.2 (15.1-28.6) | 92.3 (89.1-95.8) | 64 (50.4-77.2) | 65.7 (60.6-70.8) | 2.9 (0.3-4) | 0.8 (0.8-0.9) | | | | 7.7 (4.2-10.9) | 77.8 (71.4-84.9) |
| | -1,455 | 79.2 (72.6-85.8) | 53.2 (46.9-59.6) | 51.1 (44.3-57.7) | 80.5 (74.5-87) | 1.7 (1.4-1.9) | 0.4 (0.3-0.5) | 45.9 (40.8-50.7) | 41.4 (36.3-46.4) | 12.7 (9.3-15.9) | 46.8 (40.4-53.1) | 20.8 (14.2-27.4) |
| **ALT<=100 (n=767; F3-F4 fibrosis 27.6%)** | | | | | | | | | | | | |
| **LSM** | 9,6 | 72.6 (66.8-78.5) | 83.2 (80.3-86.3) | 62.4 (56.7-68.2) | 88.8 (86.1-91.6) | 4.3 (3.4- 5.1) | 0.3 (0.3-0.4) | | | | 16.8 (13.7-19.7) | 27.4 (21.5-33.2) |
| | 7,9 | 88.6 (84.5-92.9) | 68.7 (65-72.3) | 52 (46.9-57.2) | 94 (91.8-96.4) | 2.8 (2.4-3.2) | 0.2 (0.1- 0.2) | 13.3 (10.9-15.4) | 69.7 (66.6-72.9) | 15.3 (12.8-17.7) | 31.3 (27.7-35) | 11.4 (7.1-15.5) |
| **FIB-4** | 2,67 | 22.4 (16.7-27.8) | 96.8 (95.5-98.4) | 73.1 (62.7-84.2) | 76.5 (73.3-79.6) | 7.1 (1.6- 9.8) | 0.8 (0.7- 0.9) | | | | 3.2 (1.6-4.5) | 77.6 (72.2-83.3) |
| | 1,3 | 68.5 (62.5-74.7) | 76.7 (73.2-80.2) | 53 (47.2-59.2) | 86.4 (83.6-89.3) | 2.9 (2.4-3.4) | 0.4 (0.3-0.5) | 27.3 (24.2-30.4) | 61.6 (58.4-65.1) | 11 (8.7-13.2) | 23.3 (19.8-26.8) | 31.5 (25.3-37.5) |
| **NFS** | 0,676 | 19.6 (14.3-24.7) | 96.3 (94.9-98) | 67.2 (55.6-79) | 75.8 (72.7-79) | 5.3 (1.3-7.4) | 0.8 (0.8-0.9) | | | | 3.7 (2-5.1) | 80.4 (75.3-85.7) |
| | -1,455 | 77.2 (71.4-82.7) | 67.3 (63.4-71.3) | 47.5 (42.1-52.6) | 88.5 (85.5-91.5) | 2.4 (2-2.7) | 0.3 (0.3-0.4) | 37 (33.3-40.3) | 54.1 (50.8-57.6) | 9 (7-10.9) | 32.7 (28.7-36.6) | 22.8 (17.3-28.6) |
| **ALT>100 (n=201; F3-F4 fibrosis 31.8%)** | | | | | | | | | | | | |
| **LSM** | 9,6 | 71.9 (60-84.2) | 75.2 (67.5-83.2) | 57.7 (46.1-69.1) | 85 (78.4-92.1) | 2.9 (1.5-3.7) | 0.4 (0.2-0.5) | | | | 24.8 (16.8-32.5) | 28.1 (15.8-40) |
| | 7,9 | 87.7 (80-96.9) | 56.2 (47.6-65) | 48.5 (38.9-57.4) | 90.7 (84.8-98) | 2 (1.5-2.4) | 0.2 (0.04-0.4) | 16.9 (11.2-22.5) | 61.2 (53.9-68.5) | 20.8 (14.6-26.4) | 43.8 (35-52.4) | 12.3 (3.1-20) |
| **FIB-4** | 2,67 | 22.8 (11.5-33.1) | 95 (91.7-99.2) | 68.4 (48-89.8) | 72.3 (65-79.1) | 4.6 (1.9-15.3) | 0.8 (0.7-0.9) | | | | 5 (0.8-8.3) | 77.2 (66.9-88.5) |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1,3 | 64.9 (53-77.4) | 74.4 (66.7-82.7) | 54.4 (42.2-66.8) | 81.8 (75.2-89.5) | 2.5 (1.3-3.3) | 0.5 (0.3-0.6) | 27.5 (20.8-33.7) | 57.9 (50.6-64.6) | 14.6 (9.6-19.7) | 25.6 (17.3-33.3) | 35.1 (22.6-47) |
| **NFS** | 0,676 | 3.5 (-1.9-7) | 99.2 (97.4-100) | 66.7 (0-100) | 68.6 (62-75.4) | 4.2 (0-10) | 1 (0.9-1) | | | | 0.8 (0-2.7) | 96.5 (90.9-100) |
| | -1,455 | 63.2 (50.9-76.3) | 81 (74.3-88) | 61 (48.6-73.6) | 82.4 (75.9-89.5) | 3.3 (1.2-4.4) | 0.5 (0.3-0.6) | 31.5 (24.2-38.2) | 56.2 (48.9-63.5) | 12.4 (7.3-16.9) | 19 (12-25.7) | 36.8 (23.7-49.1) |

**Table 4.** Accuracy of Serial Combination of FIB-4 and NAFLD Fibrosis Score with Fibroscan, for the Diagnosis of Severe Fibrosis in the Entire Cohort of Patients with Nonalcoholic Fatty Liver Disease, and in Sub-groups According to BMI or ALT Values.

| Tools | Uncertainty area (%) | Accuracy (%) | Wrong classification (%) | False positive (%) | False negative (%) | Negative Predictive Value (%) | Positive Predicitve Value (%) |
|---|---|---|---|---|---|---|---|
| **Entire cohort (n=968; F3-F4 fibrosis 28.5%)** | | | | | | | |
| **NFS→LSM in NFS [-1.455; 0.676]** | 7.23 (5.58; 8.88) | 76.45 (73.86; 79.13) | 16.32 (13.95; 18.7) | 8.78 (7.02; 10.43) | 7.54 (5.79; 9.19) | 88.77 (86.39- 91.24) | 65.73 (60.08 – 71.62) |
| **FIB4→LSM in FIB4 [1.30;2.67]** | 5.89 (4.24; 7.33) | 77.79 (75.41; 80.48) | 16.32 (13.85; 18.49) | 10.23 (8.26; 12.09) | 6.1 (4.55; 7.64) | 91.05 (88.95 -93.27) | 60.71 (55.02 – 66.78) |
| | | | | | | | |
| **BMI<30 (n=591; F3-F4 fibrosis 22.3%)** | | | | | | | |
| **NFS→LSM in NFS [-1.455; 0.676]** | 5.58 (3.55; 7.45) | 82.4 (79.36; 85.79) | 12.01 (9.14; 14.55) | 8.12 (5.92; 10.32) | 3.89 (2.2; 5.41) | 94.77 (92.72 -96.93) | 59.32 (50.35 – 88.64) |
| **FIB4→LSM in FIB4 [1.30;2.67]** | 5.25 (3.38; 6.94) | 81.9 (78.85; 85.11) | 12.86 (10.15; 15.4) | 7.78 (5.58; 9.81) | 5.08 (3.38; 6.6) | 93.18 (90.93 -95.68) | 61.67 (52.70 – 70.45) |
| | | | | | | | |
| **BMI>=30 (n=377; F3-F4 fibrosis 38.2%)** | | | | | | | |
| **NFS→LSM in NFS [-1.455; 0.676]** | 9.81 (6.9; 12.73) | 67.11 (62.33; 71.62) | 23.08 (18.83; 27.32) | 9.81 (6.9; 12.73) | 13.26 (9.81; 16.71) | 76.19 (70.44 -82.19) | 71.54 (64.18 – 78.95) |
| **FIB4→LSM in FIB4 [1.30;2.67]** | 6.9 (4.24; 9.28) | 71.35 (66.84; 75.86) | 21.75 (17.51; 25.73) | 14.06 (10.34; 17.77) | 7.69 (5.04; 10.34) | 86.76 (82.44 -91.44) | 59.85 (51.35 – 68.51) |
| | | | | | | | |
| **ALT<=100 (n=767; F3-F4 fibrosis 27.6%)** | | | | | | | |
| **NFS→LSM in NFS [-1.455; 0.676]** | 7.3 (5.48; 9) | 76.92 (74.05; 79.92) | 15.78 (13.17; 18.25) | 7.56 (5.61; 9.39) | 8.21 (6.13; 10.04) | 87.91 (85.26 -90.77) | 69.47 (62.78 – 76.10) |
| **FIB4→LSM in FIB4 [1.30;2.67]** | 6.26 (4.44; 7.82) | 79.53 (76.53; 82.53) | 14.21 (11.61; 16.69) | 9.13 (7.04; 11.08) | 5.08 (3.52; 6.65) | 92.60 (90.44 -94.91) | 63.54 (56.75 – 70.34) |
| | | | | | | | |

| ALT>100 (n=201; F3-F4 fibrosis 31.8%) | | | | | | | |
|---|---|---|---|---|---|---|---|
| **NFS→LSM in NFS [-1.455; 0.676]** | 6.97 (3.48; 10.45) | 74.63 (68.66; 80.6) | 18.41 (12.44; 23.38) | 13.43 (8.46; 17.91) | 4.98 (1.99; 7.46) | 92.25 (88.07 -97.28) | 53.45 (40.79 – 66.47) |
| **FIB4→LSM in FIB4 [1.30;2.67]** | 4.48 (1.49; 6.97) | 71.14 (65.17; 77.61) | 24.38 (17.91; 29.85) | 14.43 (9.45; 18.91) | 9.95 (5.47; 13.93) | 92.60 (90.47 -94.93) | 63.54 (57.13 – 70.45) |

**Figure Legends**

**Figure 1. Receiver operating characteristic (ROC) curve of noninvasive tools for the diagnosis of advanced fibrosis.** ROC of FIB-4, NFS, and liver stiffness measurement (LSM) for advanced fibrosis in the entire cohort of patients with Nonalcoholic fatty liver disease.

**Figure 2. Receiver operating characteristic (ROC) curve of noninvasive tools for the diagnosis of advanced fibrosis in subgroups of patients.** ROC of LSM according to obesity (A), ALT levels (B), and both obesity and ALT levels (C). ROC of FIB-4 according to obesity (D) and ALT levels (E). ROC of NFS according to obesity (F) and ALT levels (G).

**Figure 3. Comparison of ROC curves of noninvasive tools for the diagnosis of advanced fibrosis.** Nonobese (A), obese (B), ALT ≤100 IU (C), ALT >100 IU (D), nonobese and ALT≤100 IU (E), obese and ALT>100 IU (F).

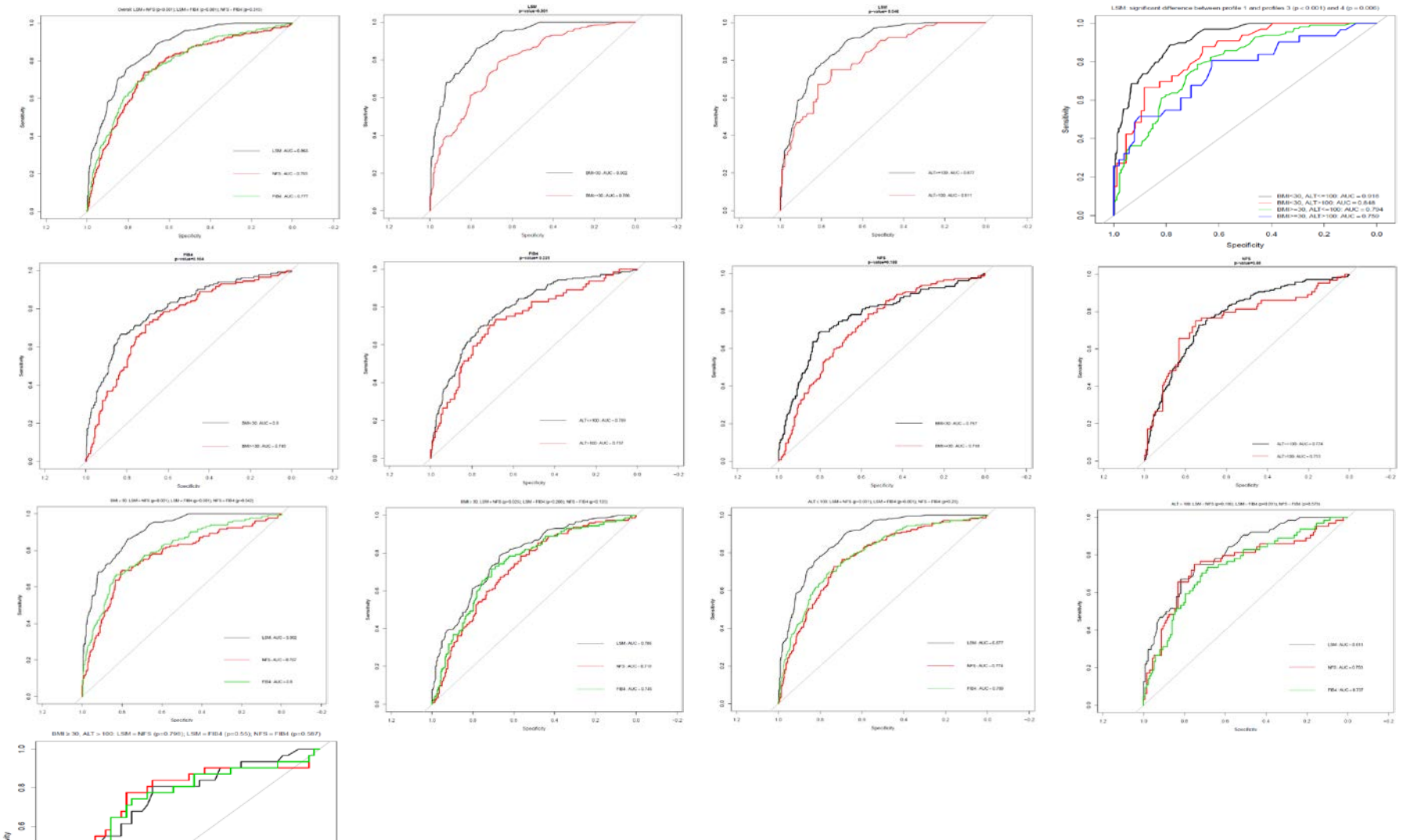

**Figure 4. Proposed Algorithm of Serial Combination Strategy of FIB-4 or NFS with LSM According to Obesity.**

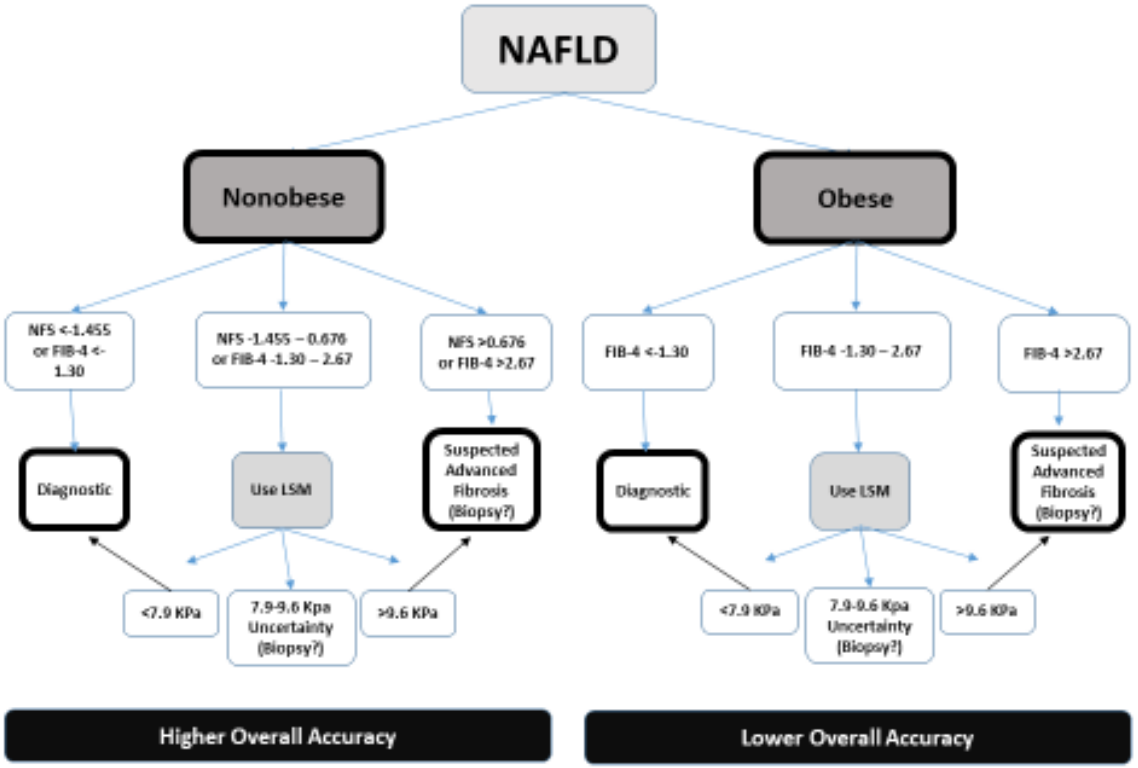


Figure 4