The Validation of Conventional Non-Invasive Fibrosis Scoring Systems in Patients with Metabolic Associated Fatty Liver Disease

Yinlian Wu
First Affiliated Hospital of Fujian Medical University

Rahul Kumar
Changi General Hospital

Mingfang Wang
First Affiliated Hospital of Fujian Medical University

Medha Singh
Tan Tock Seng Hospital

Jiaofeng Huang
First Affiliated Hospital of Fujian Medical University

Yueyong Zhu
First Affiliated Hospital of Fujian Medical University

Su Lin (sumer5129@fjmu.edu.cn)
Fujian Medical university  https://orcid.org/0000-0001-7517-9859

Research article

Keywords: metabolic associated fatty liver disease, FIB-4, NFS, APRI, BARD

DOI: https://doi.org/10.21203/rs.3.rs-53999/v1

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: Non-invasive fibrosis scores are not yet validated in the newly defined metabolic associated fatty liver disease (MAFLD). This study evaluated the diagnostic performance of four non-invasive scores including AST to platelet ratio index (APRI), fibrosis-4 index (FIB-4), BMI, AST/ALT ratio, and diabetes score (BARD), and NAFLD fibrosis score (NFS) in patients with MAFLD.

Methods: Consecutive patients with histologically-confirmed MAFLD were included. The discrimination ability of different non-invasive scores was compared.

Results: A total of 417 patients were included, 156 (37.4%) of them had advanced fibrosis (METAVIR ≥F3). The area under receiver operating characteristic curve (AUROC) of FIB-4, NFS, APRI and BARD for predicting advanced fibrosis were 0.736, 0.724, 0.671 and 0.609 respectively. The AUROC between FIB-4 and NFS were similar (P=0.523), while the difference between FIB-4 and APRI (P=0.001) and FIB-4 and BARD (P<0.001) was statistically significant. The best thresholds of FIB-4, NFS, APRI and BARD for diagnosis of advanced fibrosis in MAFLD were 1.05, -2.1, 0.42 and 2. A subgroup analysis showed that FIB-4, APRI and NFS performed worse in pure MAFLD than HBV-MAFLD group.

Conclusions: APRI and BARD score do not perform well in MAFLD. The FIB-4 and NFS could be more useful but new threshold is needed. Novel non-invasive scoring system for fibrosis is required for MAFLD.

Background

Non-alcoholic fatty liver disease (NAFLD), defined as excessive fat accumulation in liver cells in the absence of other liver diseases, has become a new epidemic due to its growing prevalence. As of to date, NAFLD is believed to affect more than a quarter of the global population. The natural history of NAFLD is highly variable, however it is believed to progresses through various fibrosis stages to end up in liver cirrhosis in a significant number of patients. The development and grade of liver fibrosis are strongly related with the adverse outcomes of NAFLD. Thus, it is critical to identify patients with advanced fibrosis to optimize the management of NAFLD.

Liver biopsy is currently regarded as the "gold standard" for the diagnosis of liver fibrosis. However, due to the high prevalence of NAFLD, it is nearly impossible to perform biopsy for each patient. Moreover, the inherent issue including safety, sampling errors, and the inter- and intra-observer variation in reporting limits its application. These limitations warrant the need for non-invasive scores for assessing liver fibrosis.

Numerous non-invasive assessment tools have been developed for diagnosis of advanced fibrosis. The most widely used non-invasive scores include aspartate aminotransferase (AST) to platelet ratio index (APRI), fibrosis-4 index (FIB-4), body mass index (BMI), AST/ALT ratio, and diabetes score (BARD), and NAFLD fibrosis score (NFS). Most of them have been tested in subjects with NAFLD, showing great diagnostic accuracy in predicting fibrosis.

Metabolic associated fatty liver disease (MAFLD) is a recently proposed concept to replace NAFLD. Significantly different from NAFLD, the MAFLD criteria does not require the exclusion of any chronic liver diseases, however the presence of metabolic associated disease or dysfunction is required. It is known that metabolic profiles are associated with risk for severe fibrosis in patient with NAFLD and MAFLD population had higher non-fibrosis scores than NAFLD. Thus, in the light of new concept of MAFLD which incorporates metabolic disorder, the performance of those non-invasive models requires re-evaluation and further validation. This study aimed to evaluate the utility of conventional non-invasive scoring systems in MAFLD.
Methods

Ethics

The study protocol was approved by the Institutional Review Board of the First Affiliated Hospital of Fujian Medical University and was in accordance with the Declaration of Helsinki. Written consent form was obtained from all patients. The data was anonymized prior to the analysis.

Patients

Consecutive patients with histologically-confirmed MAFLD admitted to the First Affiliated Hospital of Fujian Medical University from 2005 to 2015 were retrospectively reviewed and included in this study.

Histologic evaluation

All patients enrolled in this study underwent percutaneous liver biopsy under ultrasonic guidance. The liver specimens were fixed in formalin, embedded in paraffin and stained with hematoxylin and eosin, and Masson’s trichrome. For the liver biopsy to be considered adequate, the minimum biopsy length was 15 mm and at least 6 portal areas were required\textsuperscript{19}. Histopathological slides were reviewed independently by two pathologists experienced in reading liver histopathology slides and were blinded to the patient’s clinical data.

Fatty liver was defined as the presence of steatosis in at least 5% of hepatocytes. The liver fibrosis were graded as 0 to 4 points according to Metavir fibrosis stage\textsuperscript{20}, which is 0=absence of fibrosis; 1 = perisinusoidal or periportal; 2 = perisinusoidal and portal/periportal; 3 = bridging fibrosis; 4 = cirrhosis. Advanced fibrosis was defined as stage 3 or 4 fibrosis (bridging fibrosis or cirrhosis).

Diagnosis of MAFLD

The diagnosis of MAFLD was based on the following criteria\textsuperscript{15}: a histological evidence of hepatic steatosis and the presence of one of the following three conditions: BMI $\geq 23$ kg/m$^2$, presence of type 2 diabetes mellitus (T2DM), or evidence of metabolic dysregulation. The metabolic dysregulation was defined by the presence of at least two of the following conditions: 1) waist circumference $\geq$ 90cm in men and 80cm in women; 2) blood pressure $\geq$ 130/85mmHg or specific drug treatment; 3) plasma triglycerides $\geq$ 1.70 mmol/L or specific drug treatment; 4) plasma high-density lipoprotein cholesterol (HDL-C) $< 1.0$ mmol/L for men and $< 1.3$ mmol/L for women or specific drug treatment; 5) prediabetes (i.e., fasting glucose levels 5.6 to 6.9 mmol/L, or 2-hour post-load glucose levels 7.8 to 11.0 mmol/L or glycated hemoglobin (HbA1c) 5.7% to 6.4%); 6) homeostasis model assessment-insulin resistance (HOMA-IR) score $\geq 2.5$; 7) plasma high-sensitivity C-reactive protein (hs-CRP) level $> 2$ mg/L.

According to the result of hepatitis B surface antigen (HBsAg) seropositivity, patients were divided into “pure MAFLD” group (HBsAg negative) and “HBV-MAFLD” group (HBsAg positive for > 6 months) for the purpose of subgroup analysis.

Demographic and laboratory evaluation

The following characteristics were collected on the time of biopsy from all patients: age, gender, BMI, waist circumference, history of diabetes and hypertension. Results of the following laboratory parameters were collected: blood cell count, AST, alanine aminotransferase (ALT), $\gamma$-glutamyl transpeptidase (GGT), albumin, globulin, bilirubin,
fasting plasma glucose (FPG), total cholesterol, HDL-C, low-density lipoprotein cholesterol (LDL-C), triglyceride, urea, creatinine, uric acid, fasting insulin, HbA1c, hs-CRP and HBsAg.

All biochemical assessments were performed by standard laboratory methods. Non-invasive liver fibrosis assessment included APRI, FIB-4, BARD, and NFS, which were calculated based on previously published formulas which are listed in Table 1\(^\text{10,12,21,22}\) .

**Statistical analysis**

Continuous variables are expressed as mean ± standard deviation or median (interquartile range) and were compared using Student’s \(t\)-test in the case of normally distributed data or Mann-Whitney test in the remaining cases. Categorical variables are expressed as counts (percentages) and evaluated by Chi-squared test or the Fisher’s exact test. The diagnostic accuracy of conventional non-invasive scoring systems was evaluated by the receiver operating characteristic (ROC) curve. The best cut-off points to determine the presence of advanced fibrosis were chosen based on Youden’s index. The discrimination ability of different models were compared using area under ROC curve (AUROC), positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, and accuracy. Statistical analyses were performed using the SPSS software, version 18.0 (SPSS, Chicago, IL, USA) and MedCalc software version15.2.2 (MedCalc Software, Mariakerke, Belgium). A \(P\)-value<0.05 was considered statistically significant.

**Results**

**Baseline characteristics of patients**

A total of 417 patients with biopsy-proven MAFLD were included in this study (Consort diagram: Fig. 1). The mean age was 40.54 ± 10.95 years, and the BMI was 25.48 ± 2.66 kg/m\(^2\). Of them, 354 (84.9%) were male, 82 (19.7%) had T2DM, 42 (10.1%) had hypertension and 156 (37.4%) had advanced fibrosis (Table 2). Patients with advanced fibrosis were significantly older, had higher AST, GGT, and lower albumin, triglyceride and platelet count. In addition, the FIB-4 \((P<0.001)\), NFS \((P<0.001)\), APRI \((P=0.003)\) and BARD \((P<0.001)\) were all significantly higher in patients with advanced fibrosis, compared with patients with no/mild fibrosis.
Table 1
An overview of formulas and cutoffs for determining noninvasive marker panels for detection of liver fibrosis\textsuperscript{10–12, 21,22}.

| Formula | Equation |
|---------|----------|
| FIB-4   | \((\text{Age [years]} \times \text{AST [IU/L]})/(\text{platelet count [10}^9/\text{L}] \times \sqrt{\text{ALT [IU/L]}})\) |
|         | 1.3      |
|         | 2.67\textsuperscript{11} |
|         | 1.45     |
|         | 3.25\textsuperscript{22} |
| APRI    | \((\text{AST/ULN}/\text{platelet count [10}^9/\text{L}]) \times 100\) |
|         | 0.5      |
|         | 1.5\textsuperscript{21} |
| NFS     | -1.675 + 0.037 \times \text{age (years)} + 0.094 \times \text{BMI (kg/m2)} + 1.13 \times \text{IFG/diabetes (yes = 1, no = 0)} + 0.99 \times \text{AST/ALT} - 0.013 \times \text{platelet count (× 10}^9/\text{L}) - 0.66 \times \text{albumin (g/dL)}\) |
|         | -0.676   |
|         | 1.455\textsuperscript{10} |
| BARD    | Scale 0–4 |
|         | BMI \geq 28 kg/m^2 = 1 point |
|         | AST/ALT \geq 0.8 = 2 points |
|         | Diabetes = 1 point |
|         | 2\textsuperscript{12} |
Table 2
Baseline characteristics of patients.

|                          | Overall (n = 417) | No/mild fibrosis (n = 261) | Advanced fibrosis (n = 156) | p-value |
|--------------------------|-------------------|----------------------------|-----------------------------|---------|
| Age (years)              | 40.54 ± 10.95     | 39.42 ± 11.39              | 42.41 ± 9.94                | 0.007   |
| Male, n%                 | 354 (84.9%)       | 219 (83.9%)                | 135 (86.5%)                 | 0.468   |
| BMI (kg/m²)              | 25.48 ± 2.66      | 25.51 ± 2.56               | 25.42 ± 2.83                | 0.757   |
| Diabetes mellitus, n%    | 82 (19.7%)        | 46 (17.6%)                 | 36 (23.1%)                  | 0.175   |
| Hypertension, n%         | 42 (10.1%)        | 30 (11.5%)                 | 12 (7.7%)                   | 0.212   |
| ALB (g/L)                | 42.53 ± 4.96      | 43.65 ± 5.14               | 40.65 ± 3.98                | < 0.001 |
| GLO (g/L)                | 29.88 ± 14.11     | 29.11 ± 6.00               | 31.13 ± 21.53               | 0.160   |
| ALT (U/L)                | 65.00 (41.00-143.00) | 67.00 (42.00-144.50)       | 61.50 (40.25-141.00)        | 0.884   |
| AST (U/L)                | 42.00 (30.00-71.50) | 46.00 (34.00-91.00)       | 46.00 (34.00-91.00)         | 0.003   |
| GGT (U/L)                | 53.00 (31.00-98.50) | 51.00 (29.00-90.00)       | 57.00 (35.00-110.50)        | 0.033   |
| FPG (mmol/L)             | 5.34 ± 1.42       | 5.28 ± 1.20                | 5.45 ± 1.73                 | 0.251   |
| Cr (umol/L)              | 73.41 ± 13.71     | 74.00 ± 13.50              | 72.41 ± 14.07                | 0.273   |
| UA (umol/L)              | 368.49 ± 82.08    | 377.02 ± 83.87             | 353.77 ± 76.98              | 0.007   |
| TG (mmol/L)              | 1.67 ± 1.12       | 1.78 ± 1.45                | 1.48 ± 0.71                 | 0.016   |
| HDL-C (mmol/L)           | 1.12 ± 0.30       | 1.12 ± 0.29                | 1.12 ± 0.33                 | 0.959   |
| PLT (× 10⁹/L)            | 204.00 ± 64.22    | 221.00 ± 62.76             | 175.52 ± 56.22              | < 0.001 |
| FIB-4                    | 1.48 ± 1.64       | 1.19 ± 1.65                | 1.97 ± 1.49                 | < 0.001 |
| APRI                     | 1.03 ± 1.72       | 0.84 ± 1.33                | 1.34 ± 2.19                 | 0.003   |
| NFS                      | -2.16 ± 1.38      | -2.56 ± 1.24               | -1.50 ± 1.33                | < 0.001 |
| BARD                     | < 0.001           |                            |                             |         |
| 0                        | 209 (50.1%)       | 145 (55.6%)                | 64 (41.0%)                  |         |
| 1                        | 86 (20.6%)        | 59 (22.6%)                 | 27 (17.3%)                  |         |

BMI, body mass index; ALB, albumin; GLO, globulin; AST, aspartate aminotransferase; ALT, alanine aminotransferase; GGT, γ-glutamyl transpeptidase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; FPG, fasting plasma glucose; Cr, creatinine; UA, uric acid; TG, triglyceride; PLT, platelet.
Performance of FIB-4,NFS, APRI and BARD for advanced fibrosis in MAFLD

The ROC curves were used to evaluate the utility of noninvasive scoring systems for identification of advanced fibrosis (Fig. 2). The AUROC was greatest for FIB-4 (0.736; 95% CI, 0.691–0.778), followed by NFS (0.724; 95% CI, 0.679–0.767), APRI (0.671; 95% CI, 0.623–0.715) and BARD (0.609; 95% CI, 0.560–0.656). The comparison between AUROCs showed that the discrimination abilities of FIB-4 and NFS were similar ($P$ = 0.523). Similar result was noted with NFS and APRI ($P$ = 0.080). The differences between FIB-4 and APRI ($P$ = 0.001), FIB-4 and BARD ($P$ < 0.001), as well as NFS and BARD ($P$ < 0.001) were statistically significant.

Table 3 summarizes the best cut-off points developed for prediction of advanced fibrosis by these four non-invasive models and the validation of previously reported cutoffs (for NAFLD) in this MAFLD cohort$^{10–12,21,22}$. The best cutoff points of FIB-4,APRI,NFS and BARD for the diagnosis of advanced fibrosis in MAFLD were 1.05, 0.42, -2.1 and 2, respectively. Most cutoff points were lower than prior reported thresholds for each model. Using the newly developed thresholds, the sensitivity, specificity, PPV and NPV of the above four models ranged 41.7%-81.4%, 44.4–78.2%, 46.7–55.8%, and 69.2–80.0% respectively (Table 3).
Table 3
Comparison of the diagnostic value among FIB-4, NFS, APRI, and BARD in MAFLD.

| Cutoffs | AUROC | Accuracy (%) | False negative rate (%) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | Youden index |
|---------|-------|--------------|--------------------------|-----------------|-----------------|---------|---------|--------------|
| FIB-4   | 1.05  | 0.736        | 66.2                     | 20.3            | 73.7            | 61.7    | 53.5    | 79.7         | 0.354        |
|         | 1.30  | 68.1         | 25.4                     | 57.7            | 74.3            | 57.3    | 74.6    | 0.320        |
|         | 1.45  | 68.8         | 26.5                     | 52.6            | 78.5            | 59.4    | 73.5    | 0.311        |
|         | 2.67  | 66.2         | 34.1                     | 17.3            | 95.4            | 69.2    | 65.9    | 0.127        |
|         | 3.25  | 66.7         | 34.4                     | 14.1            | 98.1            | 81.5    | 65.6    | 0.122        |
| NFS     | -2.1  | 0.724        | 68.1                     | 20.9            | 70.5            | 66.7    | 55.8    | 79.1         | 0.372        |
|         | -1.455 | 66.2         | 29.5                     | 44.9            | 78.9            | 56.0    | 70.6    | 0.238        |
|         | 0.676 | 63.8         | 36.5                     | 4.5             | 99.2            | 77.8    | 63.5    | 0.037        |
| APRI    | 0.42  | 0.671        | 58.3                     | 20.0            | 81.4            | 44.4    | 46.7    | 80.0         | 0.258        |
|         | 0.5   | 59.2         | 24.7                     | 71.2            | 52.5            | 47.2    | 75.3    | 0.237        |
|         | 1.5   | 63.3         | 34.6                     | 22.4            | 87.7            | 52.2    | 65.4    | 0.101        |
| BARD    | 2     | 0.609        | 64.5                     | 30.8            | 41.7            | 78.2    | 53.3    | 69.2         | 0.199        |

Best cutoff value in bold.

AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value.

With the previously reported cutoff value of 1.30 (lower) and 2.67 (higher) of FIB-4 for advanced fibrosis, the PPV was only 57.3% and 69.2%, and the NPV was 74.6% and 65.9%, respectively. Similar results were found for the NFS score, the PPV and NPV of the well-accepted threshold (-1.455) were only 56.0% and 70.6% (Table 3).

With the FIB-4 and NFS scores, irrespective of the numerical value of cutoff used, the accuracy was only 63.8–68.8%. The accuracy was even lower for APRI and BARD (58.3–64.5%). On the other hand, 20–35% cases were misdiagnosed as “no advanced fibrosis” when the histology actually showed presence of “advanced fibrosis” by these four scores (Table 3).

The performance of the non-invasive scores in “HBV-MAFLD” and “pure MAFLD” subgroups

According to the result of HBsAg, patients were divided into HBV-MAFLD (359, 86.1%) and pure MAFLD (58, 13.9%) subgroups. The AUROC of FIB-4, NFS and APRI in HBV-MAFLD group were 0.738, 0.725 and 0.671, all higher than in pure MAFLD group (FIB-4 0.658, NFS 0.692 and APRI 0.633), while the AUROC of BARD was lower in HBV-MAFLD group than pure MAFLD group (0.609 vs. 0.644). Using different thresholds mentioned above, the overall performance of the FIB-4, APRI and NFS, including sensitivity, specificity, PPV and NPV, were all better in HBV-MAFLD group. The BARD score performed better in pure MAFLD population (Table 4).
Table 4
Comparison of the diagnostic value among FIB-4, NFS, APRI, and BARD in HBV-MAFLD (group A) and pure MAFLD (group B) subgroups.

| Cutoffs | AUROC | Accuracy (%) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|---------|-------|--------------|-----------------|-----------------|---------|---------|
|         | A     | B            | A               | B               | A       | B       | A       | B       |
| FIB-4   |       |              |                 |                 |         |         |         |         |
| 1.05    | 0.738 | 0.658        | 67.7            | 56.9            | 74.8    | 55.6    | 62.7    | 57.1    | 58.2    | 19.2    | 78.2    | 87.5    |
| 1.30    | 0.725 | 0.692        | 68.2            | 67.2            | 58.5    | 44.4    | 75.0    | 71.4    | 61.9    | 22.2    | 72.3    | 87.5    |
| 1.45    | 0.707 | 0.644        | 68.5            | 70.7            | 54.4    | 22.2    | 78.3    | 79.6    | 63.5    | 16.7    | 71.2    | 84.8    |
| 2.67    | 0.662 | 0.536        | 63.2            | 84.5            | 17.0    | 22.2    | 95.3    | 95.9    | 71.4    | 50.0    | 62.3    | 87.0    |
| 3.25    | 0.632 | 0.496        | 63.5            | 86.2            | 13.6    | 22.2    | 98.1    | 98.0    | 83.3    | 66.7    | 62.1    | 87.3    |
| NFS     |       |              |                 |                 |         |         |         |         |
| -2.1    | 0.725 | 0.692        | 68.2            | 67.2            | 70.1    | 77.8    | 67.0    | 65.3    | 59.5    | 29.2    | 76.3    | 94.1    |
| -1.455  | 0.789 | 0.661        | 64.9            | 74.1            | 44.2    | 55.6    | 79.2    | 77.6    | 59.6    | 31.3    | 67.2    | 90.5    |
| 0.676   | 0.756 | 0.641        | 60.4            | 84.5            | 4.1     | 11.1    | 99.5    | 98.0    | 85.7    | 50.0    | 59.9    | 85.7    |
| APRI    |       |              |                 |                 |         |         |         |         |
| 0.42    | 0.671 | 0.633        | 59.3            | 51.7            | 81.6    | 77.8    | 43.9    | 46.9    | 50.2    | 21.2    | 77.5    | 92.0    |
| 0.5     | 0.653 | 0.614        | 60.2            | 55.2            | 70.7    | 77.8    | 52.8    | 51.0    | 51.0    | 22.6    | 72.3    | 92.6    |
| 1.5     | 0.663 | 0.594        | 61.0            | 77.6            | 23.1    | 11.1    | 87.3    | 89.8    | 55.7    | 16.7    | 62.1    | 84.6    |
| BARD    | 2      |              | 0.609           | 0.644           | 63.0    | 74.1    | 42.2    | 33.3    | 77.4    | 81.6    | 56.4    | 25.0    | 65.9    | 87.0    |

AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value.

Discussion

Non-invasive scoring systems are widely used to identify or exclude advanced fibrosis in patients with chronic liver disease. The main finding of our study is that, FIB-4 and NFS performed better than APRI and BARD in MAFLD patients. These scores are more useful in HBV-MAFLD than pure MAFLD population. However, the performance of all above models were not as good as previous reported in NAFLD.

The non-invasive fibrosis scoring systems are derived from widely available clinical, laboratory and anthropometric parameters. Although due to ease of use the APRI and BARD scores are more user-friendly, the discrimination abilities of the APRI and BARD in prediction of advanced fibrosis is not satisfactory in this group of MAFLD patients. The AUROC of BARD was 0.609 and the accuracy was less than 65%. It is not a surprising result as the two variables for calculating BARD score (BMI and diabetes) were also variables for the diagnosis of MAFLD. This led to low sensitivity and higher false positive for detecting advanced fibrosis in MAFLD patients, thus BARD score should not be recommended for MAFLD patients in clinical practice. APRI score, although was easy to calculate, did not perform well in MAFLD as well. The NPV did not exceed 80% and the PPV were only around 50% at any cutoff values tested. Thus, APRI should not be used in for assessment of advanced fibrosis in MAFLD population either.

The calculation of FIB-4 and NFS requires the use of more complex formulas and may not be easy to use. Even though the results of our study indicate that both FIB-4 and NFS significantly outperforms APRI and BARD for prediction of advanced fibrosis in MAFLD, the NPV and PPV of these two models did not reach the similar numbers
reported by previous studies using cohort of NAFLD patients\textsuperscript{10,11,23}. As per the previously reported studies on patients with NAFLD\textsuperscript{10,11,23}, when the cutoffs of the FIB-4 and NFS was set at 1.3 and −1.455 respectively, the NPV increased from 90–93\%, respectively. And when FIB-4 was set at 2.67, the PPV could reach 80\%. However, in our cohort of MAFLD patients, by using the aforementioned cutoffs, or the new threshold being found in present study, the NPV did not ever exceed 75\% and PPV not exceed 70\%. This finding is of utmost clinical importance and tells us that in the light of new concept of MAFLD, newer non-invasive fibrosis scores will need to be developed and validated to assess the presence of advanced fibrosis in patients with MAFLD. Another worrying yet important finding of our study is that the use of these non-invasive fibrosis scores even with the new cutoffs missed out advanced fibrosis in 20–35\% patients, mis-diagnosing them as “no advanced fibrosis” (Table 3) in this cohort of MAFLD patient.

As our cohort consisted of 359 (86.1\%) cases with chronic HBV infection, which is frequently seen in Asian countries\textsuperscript{24}, we performed a subgroup analysis to test the performance of the non-invasive scores in a “pure-MAFLD” group which is very close to previous NAFLD and “HBV-MAFLD”. Three out of the four non-invasive models (FIB-4, NFS and APRI) performed even worse in “pure-MAFLD” group than “HBV-MAFLD”. As MAFLD is a new entity, this result further reinforces the need to develop and validate novel scoring systems for fibrosis in MAFLD population.

The strength of our study is that, to our knowledge this is the first validation of conventional non-invasive fibrosis scoring systems in a large sample of histology-proven MAFLD. However, the results of this study should be interpreted in light of some limitations. First of all, a large proportion of included patients (86\%) had concomitant chronic HBV infection, which could be a seen as a potential limitation as the western MAFLD population differs substantially from Asian's in the HBsAg seropositivity rates, the results may only be applicable to a subset of the entire MAFLD pool namely "HBV-MAFLD", which is the most important subtype of MAFLD in clinical practice in Asia. Second, as our study is a single center study with only Asian population, the findings will need further validation in other Asian and western cohorts.

In conclusion, APRI and BARD score do not perform well and is not suitable for the diagnosis of advanced fibrosis in MAFLD. The FIB-4 and NFS could be more useful and we propose a new threshold of 1.05 and −2.1, respectively, which had the best diagnostic performance for advanced fibrosis. There is an urgent need of MAFLD specific novel non-invasive scoring system for predicting advanced fibrosis in patients with MAFLD including its subtypes.

### Abbreviations

MAFLD, metabolic associated fatty liver disease; NAFLD, non-alcoholic fatty liver disease; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; GGT, \(\gamma\)-glutamyl transpeptidase; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TG, triglyceride; TC, total cholesterol; HbA1c, glycosylated hemoglobin; FPG, fasting plasma glucose; UA, uric acid; HBV, hepatitis B virus; HBsAg, hepatitis B surface antigen; BMI, body mass index; APRI, AST to platelet ratio index; FIB-4, fibrosis-4 index; BARD, BMI, AST/ALT ratio, and diabetes score; NFS, NAFLD fibrosis score; PPV, positive predictive value; NPV, negative predictive value; ROC, receiver operating characteristic; AUROC, area under the receiver operating characteristic curve.

### Declarations

**Ethics approval and consent to participate**

The study protocol was approved by the Institutional Review Board of the First Affiliated Hospital of Fujian Medical University. Written consent form was obtained from all patients.
Consent for publication
Not applicable.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors have declared no conflict of interest.

Funding
This research is supported by Chinese National 13th Five-Year Plan's Science and Technology Projects (2017ZX10202201) and Qingzhong Medical Science Research Fund (B17344).

Authors’ Contributions
Study concept and design: Yinlian Wu, Rahul Kumar and Su Lin
Acquisition, cleaning of data: Yinlian Wu and Mingfang Wang
Statistical analysis: Yinlian Wu and Jiaofeng Huang
Drafting of the manuscript: Yinlian Wu, Su Lin, Medha Singh and Rahul Kumar
Critical revision: Yueyong Zhu, Medha Singh and Rahul Kumar
Supervision: Su Lin
All authors contributed to the manuscript for important intellectual content and approved the submission.

Acknowledgments
Not applicable.

References
1. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology. 2018;67(1):328–57.
2. Younossi Z, Anstee QM, Marietti M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nature reviews Gastroenterology hepatology. 2018;15(1):11–20.
3. Sarin SK, Kumar M, Eslam M, et al. Liver diseases in the Asia-Pacific region: a Lancet Gastroenterology & Hepatology Commission. The lancet Gastroenterology hepatology. 2020;5(2):167–228.
4. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2015; 149(2): 389 – 97.e10.
5. 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. Journal of hepatology. 2017;67(6):1265–73.

6. Vilar-Gomez E, Calzadilla-Bertot L, Wai-Sun Wong V, et al. Fibrosis Severity as a Determinant of Cause-Specific Mortality in Patients With Advanced Nonalcoholic Fatty Liver Disease: A Multi-National Cohort Study. Gastroenterology 2018; 155(2): 443 – 57.e17.

7. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128(7):1898–906.

8. EASL-ALEH Clinical Practice Guidelines. Non-invasive tests for evaluation of liver disease severity and prognosis. Journal of hepatology. 2015;63(1):237–64.

9. Castera L, Friedrich-Rust M, Loomba R. Noninvasive Assessment of Liver Disease in Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2019;156(5):1264–81.e4.

10. 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(4):846–54.

11. AG S, K AL, BN M, MJ T. C, AJ S. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clinical gastroenterology hepatology: the official clinical practice journal of the American Gastroenterological Association. 2009;7(10):1104–12.

12. Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. Gut. 2008;57(10):1441–7.

13. Kolhe KM, Amarapurkar A, Parikh P, et al. Aspartate transaminase to platelet ratio index (APRI) but not FIB-5 or FIB-4 is accurate in ruling out significant fibrosis in patients with non-alcoholic fatty liver disease (NAFLD) in an urban slum-dwelling population. BMJ open gastroenterology. 2019;6(1):e000288.

14. Eslam M, Sanyal AJ, George J. MAFLD: A Consensus-Driven Proposed Nomenclature for Metabolic Associated Fatty Liver Disease. Gastroenterology. 2020;158(7):1999–2014.e1.

15. Eslam M, Newsome PN, Sarin SK, et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. Journal of hepatology 2020.

16. Fouad Y, Waked I, Bollipo S, Gomaa A, Ajlouni Y, Attia D. What's in a name? Renaming 'NAFLD' to 'MAFLD'. Liver international: official journal of the International Association for the Study of the Liver 2020.

17. Petta S, Eslam M, Valenti L, et al. Metabolic syndrome and severity of fibrosis in nonalcoholic fatty liver disease: An age-dependent risk profiling study. Liver international: official journal of the International Association for the Study of the Liver. 2017;37(9):1389–96.

18. Lin S, Huang J, Wang M, et al. Comparison of MAFLD and NAFLD diagnostic criteria in real world. Liver international: official journal of the International Association for the Study of the Liver 2020.

19. Bravo AA, Sheth SG, Chopra S. Liver biopsy. N Engl J Med. 2001;344(7):495–500.

20. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatology. 1996;24(2):289–93.

21. Wai CT, Greenson JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology. 2003;38(2):518–26.

22. 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(6):1317–25.

23. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP. Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut. 2010;59(9):1265–9.
Figure 1

Flow chart of case selection.
Figure 2

The ROC curves of different scores for advanced fibrosis.