Title

N-terminal Propeptide of Type 3 Collagen-Based Sequential Algorithm can identify high-risk steatohepatitis and fibrosis in MAFLD

Short Title: Sequential algorithm, steatohepatitis and fibrosis staging

Authors:

Liang-Jie Tang [1], Gang Li [1], Mohammed Eslam [2], Pei-Wu Zhu [3], Sui-Dan Chen [4], Howard Ho-Wai Leung [5], Ou-Yang Huang [1], Grace Lai-Hung Wong [6, 7], Yu-Jie Zhou [8], Morten Karsdal [9], Diana Julie Leeming [9], Pei Jiang [10], Cong Wang [10], Hai-Yang Yuan[1], Christopher D. Byrne [11], Giovanni Targher [12], Jacob George [2], Vincent Wai-Sun Wong [6, 7, *], Ming-Hua Zheng [1, 13, 14*]

Affiliations:

[1]MAFLD Research Center, Department of Hepatology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[2]Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital, Westmead, and University of Sydney, Sydney, Australia;

[3]Department of Laboratory Medicine, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[4]Department of Pathology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[5]Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Hong Kong, China;

[6]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China;

[7]State Key Laboratory of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China;

[8]Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Renji Hospital, School of Medicine, Shanghai Jiao Tong University; Shanghai Institute of Digestive Disease, Shanghai, China;

[9]Nordic Bioscience Biomarkers and Research A/S, Herlev, Denmark;

[10]Fosun Diagnostics (Shanghai) Co., Ltd, Shanghai, China;

[11]Southampton National Institute for Health Research Biomedical Research Centre, University Hospital Southampton, Southampton General Hospital, Southampton, UK;

[12]Section of Endocrinology, Diabetes and Metabolism, Department of Medicine, University and Azienda Ospedaliera Universitaria Integrata of Verona, Verona, Italy;

[13]Institute of Hepatology, Wenzhou Medical University, Wenzhou, China;

[14]Key Laboratory of Diagnosis and Treatment for the Development of Chronic Liver Disease in Zhejiang Province, Wenzhou, China

***Corresponding author:**

Ming-Hua Zheng, MD, PhD

MAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of

Wenzhou Medical University; No. 2 Fuxue Lane, Wenzhou 325000, China.

E-mail: zhengmh@wmu.edu.cn; fax: (86) 577-55578522; tel: (86) 577-55579622.

Vincent Wai-Sun Wong, MD

Department of Medicine and Therapeutics, 9/F Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong, China.

E-mail: wongv@cuhk.edu.hk; fax: (852) 2637 3852; tel: (852) 3505 4205.

Abbreviation list:

ALT, alanine aminotransferase; AST, aspartate aminotransferase; AAR, aspartate aminotransferase to alanine aminotransferase ratio; APRI, aspartate aminotransferase-platelet ratio index; AUROC, area under the receiver operating characteristic curve; BMI, body mass index; CI, 95% confidence interval; DCA, decision curve analysis; FIB-4, fibrosis-4; FLIP, fatty liver inhibition of progression; MAFLD, metabolic dysfunction-associated fatty liver disease; NFS, NAFLD fibrosis score; NPV, negative predictive value; PPV, positive predictive value; PRO-C3, N-terminal propeptide of type 3 collagen; VCTE, Vibration-Controlled Transient Elastography.

Total word count: 3937 words

Number of tables/supplementary tables: 2/5

Number of figures/supplementary figures: 4/2

Funding sources:

This paper was funded by grants from the National Natural Science Foundation of China (82070588), High Level Creative Talents from Department of Public Health in Zhejiang Province (S2032102600032) and Project of New Century 551 Talent Nurturing in Wenzhou. GT is supported in part by grants from the School of Medicine, University of Verona, Verona, Italy. CDB is supported in part by the Southampton NIHR Biomedical Research Centre (IS-BRC-20004), UK. Vincent Wong is supported in part by a Direct Grant from The Chinese University of Hong Kong (2020.045). ME and JG are supported by the Robert W. Storr Bequest to the Sydney Medical Foundation, University of Sydney; a National Health and Medical Research Council of Australia (NHMRC) Program Grant (APP1053206), an Investigator Grant (APP1196492) and Project and ideas grants (APP2001692, APP1107178 and APP1108422).

CRediT authorship contribution statement:

Liang-Jie Tang: Conceptualization, Investigation, Data curation, Writing - original draft. **Gang Li**: Conceptualization, Writing - original draft. **Mohammed Eslam**: Investigation, Writing - review & editing. **Pei-Wu Zhu**: Data curation, Writing - review & editing. **Sui-Dan Chen**: Data curation, Writing - review & editing. **Howard Ho-Wai Leung**: Data curation, Writing - review & editing. **Ou-Yang Huang**: Data curation, Writing - review & editing. **Grace Lai-Hung Wong**: Investigation, Writing - review & editing. **Yu-Jie Zhou**: Investigation, Writing - review & editing. **Morten Karsdal**: Investigation, Writing - review & editing. **Diana Julie Leeming**:

Investigation, Writing - review & editing. **Pei Jiang**: Investigation, Writing - review & editing. **Cong Wang**: Investigation, Writing - review & editing. **Hai-Yang Yuan**: Investigation, Writing - review & editing. **Christopher D. Byrne**: Investigation, Writing - review & editing. **Giovanni Targher**: Investigation, Writing - review & editing. **Jacob George**: Conceptualization, Methodology, Writing - review & editing. **Vincent Wai-Sun Wong**: Conceptualization, Methodology, Investigation, Writing - review & editing, Project administration. **Ming-Hua Zheng**: Conceptualization, Methodology, Investigation, Data curation, Writing - review & editing, Project administration.

All authors contributed to the manuscript for important intellectual content and approved the final submission of the manuscript.

Potential competing interests:

Although the PRO-C3 ELISA test was carried out at Nordic Bioscience under a research collaboration, we confirm that cohort generation, research conceptualization, analysis, and manuscript drafting were carried out independently of the Nordic Bioscience team. ADAPT has not been developed as a proprietary test. The PRO-C3 ELISA test is not currently commercially available, but can be obtained as a Nordic Bioscience research test for research use only. Diana Julie Leeming and Morten Karsdal are employed by, and own stock at Nordic Bioscience. Grace Lai-Hung Wong and Vincent Wai-Sun Wong have served as speakers and/or consultants for Echosens. Other authors have no conflicts of interest.

Abstract

Background & Aims: With metabolic dysfunction-associated fatty liver disease (MAFLD) incidence and prevalence sharply increasing globally, there is an urgent need for non-invasive diagnostic tests to accurately screen high-risk MAFLD patients for liver inflammation and fibrosis. We aimed to develop a novel sequential algorithm based on N-terminal propeptide of type 3 collagen (PRO-C3) for disease risk stratification in patients with MAFLD.

Methods: A derivation and independent validation cohort of 327 and 142 patients with biopsy-confirmed MAFLD were studied. We compared the diagnostic performances of various non-invasive scores in different disease states, and a novel sequential algorithm was constructed by combining the best performing non-invasive scores.

Results: For patients with high-risk progressive steatohepatitis (i.e., steatohepatitis + NAFLD activity score≥4 + F≥2), the AUROC of FAST score was 0.801 (95% confidence interval (CI): 0.739-0.863), and the negative predictive value (NPV) was 0.951. For advanced fibrosis (≥F3) and cirrhosis (F4), the AUROCs of ADAPT and Agile 4 were 0.879 (95%CI: 0.825-0.933) and 0.943 (95%CI: 0.892-0.994), and the NPV were 0.972 and 0.992. Sequential algorithm of ADAPT+Agile 4 combination was better than other combinations for risk stratification of patients with severe fibrosis (AUROC=0.88), with similar results in the validation cohort. Meanwhile, in all subgroup analyses (stratifying by sex, age, diabetes, NAS, BMI and ALT), ADAPT+Agile 4 had a good diagnostic performance.

Conclusions: The new sequential algorithm reliably identifies liver inflammation and fibrosis in MAFLD, making it easier to exclude low-risk patients and recommending high-risk MAFLD patients for clinical trials and emerging pharmacotherapies.

Keywords:

Metabolic dysfunction-associated fatty liver disease; Steatohepatitis; Sequential algorithm; Fibrosis staging

Introduction

Metabolic dysfunction-associated fatty liver disease (MAFLD) affects up to nearly 30% of the adult population worldwide and its prevalence has reached more than 70% in patients of severe obesity and type 2 diabetes [1-3]. The newly proposed definition of MAFLD has been endorsed by global multi-stakeholders [4]. MAFLD has created a huge clinical and economic burden and has become the most common global cause of chronic liver diseases [5-7]. Although only a small fraction of patients with MAFLD will progress to advanced fibrosis or cirrhosis [8, 9], the absolute number is huge given the large number of MAFLD patients worldwide [10, 11]. Liver fibrosis, in particular F3-4 fibrosis, has a strong association with adverse liver-related prognosis [12, 13]. Therefore, it is important to accurately identify high-risk patients with MAFLD to guide monitoring and treatment.

Presently, the diagnosis of active liver inflammation and progressive fibrosis mainly relies on serum biomarkers, imaging methods, non-invasive scores and/or liver biopsy [14]. The first three methodologies are the main screening tests that are widely used in clinical practice [15]. Although these tests have a high negative predictive value, they also have a high false positive rate, and they only target one disease state (such as the degree of steatosis, inflammation, or the stage of fibrosis) [16-18], so that they cannot assess the overall condition of the liver. Nevertheless, liver biopsy is the gold standard for pathological evaluation of MAFLD, but its limitations mean that it is being used less frequently in clinical practice [19, 20]. The current key challenge is how to

maximize the advantages of the first three aforementioned non-invasive diagnostic methods, in order to assess liver inflammation and fibrosis in patients with MAFLD.

Based on the N-terminal propeptide of type 3 collagen (PRO-C3) and ADAPT score that have good performance in fibrosis staging in Asian MAFLD patients [21, 22], we performed a comparative analysis of several up-to-date non-invasive scores (i.e. FAST, Agile 3+ and Agile 4) with the specific aim of developing a new combined diagnostic model to screen high-risk patients with MAFLD [23, 24], in order to more conveniently and accurately assess the overall liver status in MAFLD patients. Moreover, the development of new sequential algorithms may optimize the positive predictive value to increase the diagnosis rate, and reduce the number of patients with indeterminate results who may need liver biopsy.

Methods

The derivation and validation cohorts

The derivation liver biopsy cohort included two Asian cohorts with a total of 327 adult patients aged 18-70 years with biopsy-proven MAFLD between January 2017 and December 2020 (i.e., 262 patients from Wenzhou and 65 patients from Hong Kong). The Wenzhou and Hong Kong cohorts included patients from previously published studies but also included additional patients with MAFLD [21, 22]. The independent validation liver biopsy cohort included 142 MAFLD patients from Wenzhou between June 2021 and December 2021. MAFLD was diagnosed by the

presence of hepatic steatosis on liver histology (defined as presence of more than 5% of steatotic hepatocytes) with at least one of the following three coexisting metabolic conditions, i.e., overweight/obesity, type 2 diabetes mellitus, or metabolic dysregulation [25]. In both selected cohorts of the study, we included only patients with a single aetiologic MAFLD, as we excluded those with excessive alcohol consumption (> 10 g per day for women and > 20 g per day for men), autoimmune liver diseases, chronic viral hepatitis (hepatitis B or C), secondary fatty liver, or prior history of malignancy in the past two years [26]. Patients who have not undergone vibration-controlled transient elastography (VCTE) measurement were also excluded from the study.

We recorded clinical information within one day of liver biopsy examination. Fasting venous blood samples for measurement of serum liver enzymes, total bilirubin, albumin, lipids and other blood biochemical tests were obtained. Height and weight were measured for the calculation of body mass index (BMI) = weight [kg]/height^2 [m^2]. Diabetes was defined as fasting glucose levels ≥ 7.0 mmol/L (≥ 126 mg/dL), haemoglobin A1c ≥ 6.5% (≥ 48 mmol/mol), a previous diagnosis of diabetes, or the use of any anti-hyperglycaemic drugs [27]. The study was approved by the institutional review boards of both centres and all participants gave their written informed consent [21, 22].

Liver stiffness measurement

Measurement of liver stiffness (LSM) and steatosis using VCTE (FibroScan equipped with CAP software; Echosens, Paris, France) was undertaken within a week of the diagnostic liver biopsy. The FibroScan M probe was used in the Wenzhou cohort, and the M and XL probes were used in the Hong Kong cohort with BMI < 30 and ≥ 30 kg/m^2. Ten or more measurements and an interquartile range/median ratio of liver stiffness ≤ 0.3 were required for valid LSMs.

Liver histological assessment of MAFLD

Eligible patients with MAFLD for liver biopsy were aged 18 or older, able to give written informed consent, and scheduled (independently from this study) to have a liver biopsy within 2 weeks before or after VCTE measurements for investigation of suspected MAFLD (usually due to elevated serum liver enzyme levels and hepatic steatosis on ultrasound). Ultrasound-guided percutaneous liver biopsy was performed using a 16G Tru Cut needle and tissue was stained with haematoxylin-eosin and Masson trichrome. The reading and scoring of biopsy specimens were performed by an experienced liver histopathologist in each centre, who was blinded to the clinical and biochemical data of participants. The NAFLD activity score (NAS) was calculated as the sum of hepatic steatosis, lobular inflammation and hepatocyte ballooning [19, 28, 29]. Presence of steatohepatitis was diagnosed using the fatty liver inhibition of progression (FLIP) definition (at least grade one for steatosis, ballooning, and lobular inflammation) [30]. The histologic stages of liver fibrosis were defined as: stage 0, no fibrosis; stage 1, peri-sinusoidal or portal venular fibrosis;

stage 2, peri-sinusoidal and portal vein/periportal fibrosis; stage 3, bridging fibrosis; and stage 4, cirrhosis [27]. Significant fibrosis was defined as fibrosis stage ≥ F2, and advanced fibrosis was defined as fibrosis stage ≥ F3.

Non-invasive scoring tests

The formation of N-terminal propeptide of type 3 collagen (PRO-C3) was assessed in serum using the PRO-C3 competitive ELISA kit from Nordic Bioscience, Herlev, Denmark, as previously described [31]. It has been established that the parameters that have a significant contribution to predicting the severity of liver fibrosis were liver stiffness measurement (LSM), aspartate aminotransferase-to-alanine aminotransferase ratio (AAR), platelet count (PLT), sex and diabetes status. After including the above-mentioned parameters, the **Agile 3+** and **Agile 4** algorithms included (diabetes status: yes = 1, no = 0 and sex: male = 1, female = 0) follow the two subsequent equations:

$$Agile\ 3{+}{=}\frac{e^{logit(p_{F\geq 3})}}{1+e^{logit(p_{F\geq 3})}}$$

with $e^{logit(p_{F\geq 3})} = -3.92368 + 2.29714 \times \ln(LSM) - 0.00902 \times PLT - 0.98633 \times AAR^{-1} + 1.08636 \times Diabetes\ status - 0.38581 \times Sex + 0.03018 \times Age.$

$$Agile\ 4 = \frac{e^{logit(p_{F=4})}}{1+e^{logit(p_{F=4})}}$$

with $e^{logit(p_{F=4})} = 7.50139 - 15.42498 \times \frac{1}{\sqrt{LSM}} - 0.01378 \times PLT - 1.41149 \times AAR^{-1} - 0.53281 \times Sex + 0.41741 \times Diabetes\ status.$

Additionally, we used multivariate logistic regression to create a novel non-invasive

score to identify patients with advanced fibrosis in MAFLD, which including PRO-C3, LSM, platelet levels and diabetes status.

$$MLR\ Model = -3.89497 + 0.09313 \times PRO - C3 + 0.20373 \times LSM - 0.00603 \times Platelets + 1.29126 \times Diabetes\ status$$

Other widely used non-invasive scores of fibrosis, namely FibroScan-AST (FAST) score [24], ADAPT score [31], APRI [32], FIB-4 [33], NAFLD fibrosis score (NFS) [34], and BARD (BMI, AAR, diabetes) score [35] were calculated using available clinical and laboratory variables.

Statistical analysis

According to their normal or not normal distribution, continuous data were reported as means and standard deviation or medians (interquartile ranges), respectively. Categorical data were reported as a numbers (or percentages). The overall diagnostic accuracy of the various non-invasive scores was evaluated by the receiver operator characteristic (ROC) curve analysis and expressed as the area under the receiver operator characteristic curve (AUROC). The diagnostic accuracies of the FAST score, the ADAPT score, the Agile 3+ and Agile 4 algorithms, and other non-invasive fibrosis scores were determined by calculating the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), as well as positive and negative likelihood ratios (LRs).

The DeLong test was performed on the AUROC curves by a bootstrap re-sampling

method with 500 repetitions. The dual cut-off value improves the diagnostic certainty of screening high-risk groups of fibrosis. The curve with the greatest probability was the best decision strategy to maximize the new advantage. We conducted a decision curve analysis (DCA) to evaluate which combination of the two algorithms is more advantageous, that is, whether the new decision tree curve was better at identifying the severity of fibrosis rather than harmful (screening and deselecting the population that are in most need of a liver biopsy, reducing the risk of an unnecessary invasive inspection). All statistical analyses were performed by R version 3.6.1 (https://www.r-project.org/).

Results

Patient Characteristics

The clinical and biochemical characteristics of Asian patients with biopsy-proven MAFLD in the derivation (n=327) and validation (n=142) cohorts are shown in **Table 1**. Most clinical and laboratory parameters, as well as serum PRO-C3 levels, FAST score, ADAPT score, Agile 3+ score and Agile 4 score did not differ significantly between the two cohorts of patients. Compared to those in the validation cohort, patients in the derivation cohort had a lower prevalence of steatohepatitis (56% vs.74%, P< 0.001), but a comparable prevalence of advanced fibrosis (13% vs.8%, P= 0.125) and NAS ≥ 4 (51% vs.58%, P= 0.158) on liver histology.

Diagnostic accuracy of the FAST score and other widely used non-invasive fibrosis

scores for steatohepatitis + NAS ≥ 4 + F ≥ 2

In the derivation cohort, the FAST score was compared with the APRI, FIB-4, NFS and BARD scores for the identification of patients with high-risk progressive steatohepatitis, which was defined as presence of steatohepatitis + NAS ≥ 4 + F ≥ 2. **Supplementary Figure 1A** shows the diagnostic accuracy of the FAST and other non-invasive scores. The AUROC of the FAST score was 0.801 (95% CI: 0.739-0.863) that was higher than that of the other widely used non-invasive scores: 0.687 (95% CI: 0.604-0.770, P< 0.001) for APRI, 0.594 (95% CI: 0.506-0.683, P< 0.001) for FIB-4, 0.625 (95% CI: 0.543-0.707, P< 0.001) for NFS, and 0.640 (95% CI: 0.565-0.714, P= 0.029) for BARD, respectively. In our cohort the cut-off value for sensitivity (≥ 90%) was 0.285 and for specificity (≥ 90%) was 0.605, respectively. Using a dual cut-off approach (**Supplementary Figure 2A**), we found that patients with FAST score < 0.285 were low-risk groups (NPV= 0.951); patients with FAST score > 0.605 were likely to have progressive steatohepatitis, while 122 (37.3%) of 327 patients were in the 'gray zone' between the two FAST cut-off values. The FAST score showed a good diagnostic performance between the validation and derivation cohort (AUROC= 0.742, 95% CI: 0.617-0.868, NPV= 0.964), and 58 (40.8%) patients were in the 'gray zone' (**Supplementary Table 1**).

Diagnostic performances of novel scores and other widely used non-invasive fibrosis scores for advanced fibrosis and cirrhosis

Based on the previously observed strong association between serum PRO-C3 levels

and the histological severity of liver fibrosis [21], we tested the diagnostic performance of the ADAPT score for identifying advanced fibrosis in the derivation cohort and compared it with Agile 3+ score and other widely used serum-based non-invasive scores (i.e., APRI, FIB-4, BARD and NFS scores). As shown in the **Supplementary Table 2**, the AUROC of the ADAPT score was 0.879 (95% CI: 0.825-0.933, NPV= 0.972) for advanced fibrosis, better than Agile 3+ score (AUROC=0.805, 95% CI: 0.725-0.886, P= 0.015, NPV= 0.952) and Agile 4 score (AUROC=0.826, 95% CI: 0.747-0.904, P= 0.019, NPV= 0.97), also higher than that of the other widely used non-invasive scores: 0.642 (95% CI: 0.550-0.734, P< 0.001) for APRI, 0.732 (95% CI: 0.643-0.820, P< 0.001) for FIB-4, 0.800 (95% CI: 0.723-0.877, P< 0.001) for NFS, and 0.709 (95% CI: 0.632-0.786, P< 0.001) for BARD, respectively (**Supplementary Figure 1B**). The ADAPT score also performed best in the validation cohort (AUROC=0.885, 95% CI: 0.772-0.998, NPV= 0.988) compared to other non-invasive scores (see the **Supplementary Table 2**). We also used the dual cut-offs approach for ruling in and ruling out MAFLD patients with advanced fibrosis (**Supplementary Figure 2B**).

Furthermore, the AUROC for the Agile 4 score was 0.943 (95% CI: 0.892-0.994) for cirrhosis (n= 18), 0.909 (95% CI: 0.843-0.974, P= 0.169) for the Agile 3+ score, 0.582 (95% CI: 0.438-0.727, P< 0.001) for the APRI score, 0.794 (95% CI: 0.685-0.903, P= 0.006) for the FIB-4 score, 0.905 (95% CI: 0.849-0.961, P= 0.113) for NFS, and was 0.801 (95% CI: 0.720-0.882, P= 0.134) for BARD in the derivation cohort

(**Supplementary Figure 1C**). In view of the excellent diagnostic performance of the Agile 4 score in patients with cirrhosis, we used the dual cut-offs (**Supplementary Figure 2C**) to identify patients with advanced fibrosis. Patients with Agile 4 values < 0.318 (≥ 90% sensitivity) did not have cirrhosis, and patients with Agile 4 values > 0.586 (≥ 90% specificity) were likely to have cirrhosis. Due to the small number of patients with cirrhosis (F4) in the validation group (n= 4), the comparison of diagnostic performance lacks confidence, but on the whole cohort, Agile 4 score still performed well (**Supplementary Table 3**).

Combination of Agile 4 and other scores in a sequential algorithm improves the diagnostic accuracy of advanced fibrosis

We compared combinations of multiple sequential algorithms, including MLR Model, PRO-C3, VCTE, and other clinical and hematological parameters. The combination of ADAPT + Agile 4 showed the best diagnostic performance, with an AUROC of 0.880 (95% CI: 0.824-0.935) and accuracy of 88.1% (NPV= 0.955) that was better than MLR Model (AUROC=0.876, 95% CI: 0.817-0.935, accuracy of 85.3%) and ADAPT + LSM (AUROC=0.879, 95% CI: 0.821-0.936, accuracy of 84.4%), and also higher than that of the other widely used non-invasive scores. At the same time, a requirement for liver biopsy in only 31.4% of patients (gray zone) and the specificity was the highest (90.5%) (**Supplementary Table 4**). Taking advantages of these parameters (VCTE can be performed on an outpatient basis and hematological parameters can be obtained on the day of admission at no additional cost), sequential

combinations of non-invasive scores were used to assess the overall liver status of patients with MAFLD.

Based on the performance of the FIB-4 and NFS scores for diagnosing advanced fibrosis, we further evaluated the performance of the sequential combination algorithm of ADAPT, Agile 3+, FIB-4, NFS scores and Agile 4 score for the diagnosis of advanced fibrosis. The best performance was obtained by using ADAPT + Agile 4 (AUROC= 0.880, 95% CI: 0.824-0.935), which was better than combinations of the other sequential predictive algorithms: 0.823 (95% CI: 0.745-0.901, P= 0.009) for (Agile 3+) + Agile 4, 0.785 (95% CI: 0.703-0.866, P= 0.002) for FIB-4 + Agile 4, and 0.812 (95% CI: 0.736-0.889, P= 0.002) for NFS + Agile 4, respectively (**Figure 1A**). The diagnostic accuracy of this sequential algorithm that utilized the ADAPT + Agile 4 scores for the detection of advanced fibrosis was also superior to replacing ADAPT by any of Agile 3+, FIB-4, or NFS, with a diagnostic accuracy of 88.1%, compared with 86.9%, 85.9%, and 80.7%, respectively. The same results were also confirmed in the validation group (AUROC= 0.905, 95% CI: 0.788-1.000, NPV= 0.984), ADAPT + Agile 4 could best rule out advanced fibrosis in MAFLD (see the **Table 2**).

The decision curve analysis (DCA) used for assessing the sequential algorithms is presented in **Figure 1B**. This figure analyzes the clinical utility of the ADAPT + Agile 4 sequential algorithm compared with (Agile 3+) + Agile 4, FIB-4 + Agile 4, and NFS + Agile 4 in identifying patients with advanced fibrosis (F ≥ 3). DCA showed that

from a threshold probability > 10%, we could obtain more net benefit by using a variance graph. In particular, if a patient's threshold probability was > 10% and < 70%, the use of the ADAPT + Agile 4 score for predicting the risk of advanced fibrosis had more benefits than the reference strategy (using Agile 3+, FIB-4, or NFS scores combined with Agile 4 score).

We combined the algorithms of ADAPT (lower value of 3.705 and upper value of 4.93) and Agile 4 (lower value of 0.318 and upper value of 0.586), and then compared it with other combinations. This new combination algorithm approach could more accurately deselect patients with high-risk steatohepatitis (i.e., the number of true positive (TP) groups has increased from 18%, 21% and 31% to 32%), while reducing the number of patients requiring liver biopsy (including false positive (FP)), false negative (FN) and indeterminate groups from 34%, 32% and 26% to 20%) (**Figure 2**).

Diagnostic performance of ADAPT + Agile 4 in patient subgroups

Circulating levels of hyaluronic acid, type III procollagen, type IV collagen, and laminin were also used as routine liver fibrosis tests in the Wenzhou cohort. The best performance was still obtained by combining ADAPT + Agile 4 (AUROC= 0.761, 95% CI: 0.657-0.864), which was better than established fibrosis markers and combinations of other sequential algorithms: 0.508 (95% CI: 0.335-0.682) for hyaluronic acid, 0.661 (95% CI: 0.506-0.815) for type III procollagen, 0.635 (95% CI: 0.472-0.798) for type IV collagen, 0.570 (95% CI: 0.437-0.703) for laminin, 0.638

(95% CI: 0.468-0.808) for (Agile 3+) + Agile 4, 0.642 (95% CI: 0.480-0.804) for FIB-4 + Agile 4, and 0.663 (95% CI: 0.499-0.825) for NFS + Agile 4, respectively. In particularly, the ADAPT + Agile 4 combination algorithm had 100% sensitivity and 100% NPV for excluding advanced fibrosis in the Wenzhou cohort (see the **Supplementary Table 5**).

The ADAPT + Agile 4 combination algorithm had a good diagnostic performance for advanced fibrosis among different subgroups of MAFLD patients. In men, the AUROC of this sequential algorithm was 0.845 (95% CI: 0.771-0.919), and among women the AUROC was 0.953 (95% CI: 0.907-0.999). Similarly, in different patient subgroups, such as those with age < 44 years (AUROC=0.822, 95% CI: 0.726-0.917), age ≥ 44 years (AUROC=0.894, 95% CI: 0.829-0.958), NAS < 4 (AUROC=0.889, 95% CI: 0.794-0.984), NAS ≥ 4 (AUROC=0.880, 95% CI: 0.814-0.945), BMI < 25 kg/m^2 (AUROC=0.914, 95% CI: 0.804-1.000), BMI ≥ 25 kg/m^2 (AUROC=0.865, 95% CI: 0.798-0.931), ALT < 40 U/L (AUROC=0.821, 95% CI: 0.701-0.942), ALT ≥ 40 U/L (AUROC=0.915, 95% CI: 0.858-0.972), presence of diabetes (AUROC=0.871, 95% CI: 0.796-0.946) or absence of diabetes (AUROC=0.769, 95% CI: 0.660-0.878), the ADAPT + Agile 4 combination algorithm showed comparable good performances for predicting advanced fibrosis in patients with MAFLD (**Figure 3**).

The combined hierarchical screening using FAST, combined with ADAPT and

Agile 4 significantly improves the diagnostic accuracy of fibrosis staging

We utilized three non-invasive scores that had good performance in different stages of liver fibrosis, and developed a combined diagnostic method for staging the severity of fibrosis in MAFLD patients. Firstly, calculating the FAST score with FibroScan, we propose that patients with a FAST score < 0.605 could undergo routine follow-up (NPV= 0.9). Then, for patients with higher FAST scores who may have progressive steatohepatitis, those with an ADAPT value < 4.93 or in the gray area need to confirm by liver biopsy whether it is in the stage of severe inflammation and fibrosis. Finally, the Agile 4 score was used to assess the risk of cirrhosis in patients with advanced fibrosis with an ADAPT value > 4.93 (NPV= 0.955), and patients with an Agile 4 value < 0.586 were excluded from cirrhosis and were referred for tertiary care referral (NPV= 0.99). The performance of using this combined diagnostic method is better than using non-invasive algorithms or serum markers alone, and may reduce the number of patients requiring liver biopsy (see **Figure 4**).

Discussion

In this study we combined three non-invasive diagnostic scores (i.e. the FAST, ADAPT and Agile 4 scores) for identifying MAFLD progressive steatohepatitis and advanced fibrosis to classify severity of liver disease. The diagnostic performance and accuracy of this sequential prediction algorithm is better than other non-invasive scores or sequential algorithms (including the MLR Model, the combined ADAPT + LSM or the combined (Agile 3+) + Agile 4), and its superiority has been further

demonstrated in the independent validation cohort. This sequential prediction algorithm effectively reduces the need for liver biopsy, and provides a more comprehensive classification of MAFLD patients, so that patients with different risks could be better targeted for clinical management.

Although the degree of liver fibrosis determines the poor prognosis of MAFLD patients [36], increased necrotizing inflammatory activity causes progressive damage and may influence treatment response [37, 38]. The high-risk MAFLD patients with progressive steatohepatitis (defined as steatohepatitis + NAS $\geq$ 4 + F $\geq$ 2) could be selected and treated as a secondary or tertiary care referral. The diagnostic performance of the FAST score in the derivation cohort (AUROC= 0.801) could better exclude high-risk patients with MAFLD (NPV= 0.951), and better predict progressive steatohepatitis patients who may have a poor prognosis (NPV= 0.9). At the same time, two thresholds were used to reduce the number of patients in the gray or intermediate zones.

As the level of fibrosis is the most critical indicator that affects the prognosis of MAFLD patients [39], it is particularly important to assess the severity of the disease as accurately as possible. The ADAPT score based on the new serum marker PRO-C3 is good at diagnosing advanced fibrosis (AUROC = 0.879). We used the ADAPT score for further risk assessment of patients with progressive liver inflammation and significant fibrosis screened in the first step (i.e. those with

steatohepatitis + NAS ≥ 4 + F ≥ 2). Our results showed that an ADAPT score < 3.705 could basically eliminate the risk of advanced fibrosis (sensitivity = 89.7%, NPV = 0.972), but it is necessary to confirm by the liver biopsy with patients in the gray zone, through pathological examination to confirm the pathological conditions of the liver, because it is not ruled out that individual patients are still at risk of severe fibrosis. For patients with a high ADAPT value, we used the Agile 4 score for final screening (AUROC = 0.943). After excluding cirrhosis (sensitivity = 88.9%, NPV = 0.992), we will advise patients to make tertiary referral as soon as possible for better medical management.

Our newly proposed three-step sequential algorithm only requires patients to carry out a simple blood test for measurement of serum PRO-C3 level and undergo the FibroScan-measured CAP and LSM measurements. Combined with some basic patient data the algorithm allows evaluation of disease status of patients accurately and quickly. In the face of such a large MAFLD population, our proposed three-step sequential algorithm could effectively screen out low-risk patients and high-risk patients who need different management, so as to reduce the unnecessary social and economic burden of the liver disease.

The applicability of the ADAPT+ Agile 4 sequential algorithm was also verified in various subgroups of patients with MAFLD. In different patient subgroups, the diagnostic performance of the ADAPT+ Agile 4 sequential algorithm for identifying

advanced fibrosis was better than other non-invasive diagnostic score combinations (AUROC = 0.88, accuracy = 88.1%, NPV = 0.955). The combined sequential algorithm only utilized a few simple serological markers and the results of FibroScan to perform a hierarchical screening of MAFLD patients from mild to severe, so as to more accurately implement suitable treatment measures and lifestyle interventions.

Our study has some limitations that should be mentioned. Firstly, liver biopsy may be affected by sampling variability, intra-observer and inter-observer variability. Secondly, we conducted the study only in Asian patients. In addition, although we compared the diagnostic performance of several non-invasive fibrosis scores and included three models in developing the new combined sequential algorithm, we were unable to compare our newly proposed combined sequential algorithm with magnetic resonance elastography (MRE) [40-42] and other newer serological markers for fibrosis diagnosis (such as CK-18 M30, CK-18 M65, CHI3L1, and M2BPGi) [43]. Thirdly, analysing data from larger subsets in each of the MAFLD sub-groups to examine for the performance of x, y and z, would be an important area for future research. Finally, our patients were from tertiary hospital-based centres, so that the effectiveness of the test in primary care needs to be tested.

In conclusion, our study shows that the combined model of FAST+ ADAPT+ Agile 4 scores can be effectively used for risk stratification of liver disease in MAFLD and can reduce the need for unnecessary liver biopsy. Further studies are certainly needed

to corroborate these findings in other cohorts of patients with MAFLD of different countries.

Reference

1. Liu J, Ayada I, Zhang X, Wang L, Li Y, Wen T et al. Estimating Global Prevalence of Metabolic Dysfunction-Associated Fatty Liver Disease in Overweight or Obese Adults. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2021
2. Eslam M, Sarin SK, Wong VW, Fan JG, Kawaguchi T, Ahn SH et al. The Asian Pacific Association for the Study of the Liver clinical practice guidelines for the diagnosis and management of metabolic associated fatty liver disease. Hepatology international 2020;14:889-919
3. Zheng KI, Fan JG, Shi JP, Wong VW, Eslam M, George J et al. From NAFLD to MAFLD: a "redefining" moment for fatty liver disease. Chinese medical journal 2020;133:2271-2273
4. Méndez-Sánchez N, Bugianesi E, Gish RG, Lammert F, Tilg H, Nguyen MH et al. Global multi-stakeholder endorsement of the MAFLD definition. The lancet Gastroenterology & hepatology 2022
5. Méndez-Sánchez N, Díaz-Orozco L, Córdova-Gallardo J. Redefinition of fatty liver disease from NAFLD to MAFLD raised disease awareness: Mexican experience. Journal of hepatology 2021;75:221-222
6. Eslam M, Sanyal AJ, George J. MAFLD: A Consensus-Driven Proposed Nomenclature for Metabolic Associated Fatty Liver Disease. Gastroenterology 2020;158:1999-2014.e1991
7. Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030. Journal of hepatology 2018;69:896-904
8. Zheng KI, Eslam M, George J, Zheng MH. When a new definition overhauls perceptions of MAFLD related cirrhosis care. Hepatobiliary surgery and nutrition 2020;9:801-804
9. Rios RS, Zheng KI, Zheng MH. Non-alcoholic steatohepatitis and risk of hepatocellular carcinoma. Chinese medical journal 2021
10. Wang TY, Wang RF, Bu ZY, Targher G, Byrne CD, Sun DQ et al. Association of metabolic dysfunction-associated fatty liver disease with kidney disease. Nature reviews Nephrology 2022;18:259-268
11. Fouad Y, Gomaa A, Semida N, Ghany WA, Attia D. Change from NAFLD to MAFLD increases the awareness of fatty liver disease in primary care physicians and specialists. Journal of hepatology 2021;74:1254-1256
12. Zheng KI, Zheng MH. The uprising of metabolic dysfunction-associated fatty liver disease (MAFLD) in acute-on-chronic liver failure (ACLF). Hepatobiliary surgery and nutrition 2021;10:857-859
13. 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 (Baltimore, Md) 2015;61:1547-1554
14. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, Sanyal AJ. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2009;7:1104-1112
15. Zheng KI, Sun DQ, Jin Y, Zhu PW, Zheng MH. Clinical utility of the MAFLD definition. Journal

of hepatology 2021;74:989-991
16. Wu XX, Zheng KI, Boursier J, Chan WK, Yilmaz Y, Romero-Gómez M et al. acNASH index to diagnose nonalcoholic steatohepatitis: a prospective derivation and global validation study. EClinicalMedicine 2021;41:101145
17. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. Journal of hepatology 2018;68:305-315
18. Castera L, Friedrich-Rust M, Loomba R. Noninvasive Assessment of Liver Disease in Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2019;156:1264-1281.e1264
19. Younossi ZM, Gramlich T, Liu YC, Matteoni C, Petrelli M, Goldblum J et al. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 1998;11:560-565
20. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898-1906
21. Tang LJ, Ma HL, Eslam M, Wong GL, Zhu PW, Chen SD et al. Among simple non-invasive scores, Pro-C3 and ADAPT best exclude advanced fibrosis in Asian patients with MAFLD. Metabolism: clinical and experimental 2022;128:154958
22. Eslam M, Wong GL, Hashem AM, Chan HL, Nielsen MJ, Leeming DJ et al. A Sequential Algorithm Combining ADAPT and Liver Stiffness Can Stage Metabolic-Associated Fatty Liver Disease in Hospital-Based and Primary Care Patients. The American journal of gastroenterology 2021;116:984-993
23. Sanyal AJ, Foucquier J, Younossi ZM, Harrison SA, Newsome PN, Chan WK et al. Enhanced diagnosis of advanced fibrosis and cirrhosis in non-alcoholic fatty liver disease patients using FibroScan-based Agile scores. Journal of hepatology. 2022; Accepted.
24. Newsome PN, Sasso M, Deeks JJ, Paredes A, Boursier J, Chan WK et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. The lancet Gastroenterology & hepatology 2020;5:362-373
25. Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. Journal of hepatology 2020;73:202-209
26. Zhou YJ, Gao F, Liu WY, Wong GL, Mahadeva S, Raihan Nik Mustapha N et al. Screening for compensated advanced chronic liver disease using refined Baveno VI elastography cutoffs in Asian patients with nonalcoholic fatty liver disease. Alimentary pharmacology & therapeutics 2021;54:470-480
27. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. Journal of hepatology 2016;64:1388-1402
28. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology (Baltimore, Md) 2005;41:1313-1321
29. Fukusato T, Fukushima J, Shiga J, Takahashi Y, Nakano T, Maeyama S et al. Interobserver variation in the histopathological assessment of nonalcoholic steatohepatitis. Hepatology research : the official journal of the Japan Society of Hepatology 2005;33:122-127
30. Bedossa P. 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 (Baltimore, Md) 2014;60:565-575
31. Daniels SJ, Leeming DJ, Eslam M, Hashem AM, Nielsen MJ, Krag A et al. ADAPT: An Algorithm Incorporating PRO-C3 Accurately Identifies Patients With NAFLD and Advanced Fibrosis. Hepatology (Baltimore, Md) 2019;69:1075-1086
32. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, Conjeevaram HS et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology (Baltimore, Md) 2003;38:518-526
33. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology (Baltimore, Md) 2006;43:1317-1325
34. 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 (Baltimore, Md) 2007;45:846-854
35. 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;57:1441-1447
36. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 2015;149:389-397.e310
37. Friedman SL, Ratziu V, Harrison SA, Abdelmalek MF, Aithal GP, Caballeria J et al. A randomized, placebo-controlled trial of cenicriviroc for treatment of nonalcoholic steatohepatitis with fibrosis. Hepatology (Baltimore, Md) 2018;67:1754-1767
38. Ratziu V, Harrison SA, Francque S, Bedossa P, Lehert P, Serfaty L et al. Elafibranor, an Agonist of the Peroxisome Proliferator-Activated Receptor-α and -δ, Induces Resolution of Nonalcoholic Steatohepatitis Without Fibrosis Worsening. Gastroenterology 2016;150:1147-1159.e1145
39. Wang TY, George J, Zheng MH. Metabolic (dysfunction) associated fatty liver disease: more evidence and a bright future. Hepatobiliary surgery and nutrition 2021;10:849-852
40. Hsu C, Caussy C, Imajo K, Chen J, Singh S, Kaulback K et al. Magnetic Resonance vs Transient Elastography Analysis of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Pooled Analysis of Individual Participants. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2019;17:630-637.e638
41. Noureddin M, Truong E, Gornbein JA, Saouaf R, Guindi M, Todo T et al. MRI-based (MAST) score accurately identifies patients with NASH and significant fibrosis. Journal of hepatology 2022;76:781-787
42. Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H et al. Magnetic Resonance Imaging More Accurately Classifies Steatosis and Fibrosis in Patients With Nonalcoholic Fatty Liver Disease Than Transient Elastography. Gastroenterology 2016;150:626-637.e627
43. Shirabe K, Bekki Y, Gantumur D, Araki K, Ishii N, Kuno A et al. Mac-2 binding protein glycan isomer (M2BPGi) is a new serum biomarker for assessing liver fibrosis: more than a biomarker of liver fibrosis. Journal of gastroenterology 2018;53:819-826

Figure legends

Figure 1. (A) The area under the receiver operating characteristic curves (AUROC) for predicting F≥ 3 fibrosis. **(B)** Decision curve analysis (DCA) for the combined diagnosis of non-invasive fibrosis scores. The y-axis represents net benefits, calculated by subtracting the relative harms (true positives) from the benefits (false positives). The x-axis measures the threshold probability. A screening strategy is superior if it has the highest value compared with other models, including two simple strategies, such as all patients (sloping solid line) or no patients (horizontal solid line).

Figure 2. The Sankey diagrams showed the distribution of patients in the true positive (TP), true negative (TN), false positive (FP), false negative (FN), and indeterminate groups. Different thresholds are used for each test layer through the ADAPT score and the Agile 4 score. When applying the two tests, the lower threshold was used to rule out patients without advanced fibrosis and cirrhosis, and the upper threshold was used to rule in patients with advanced fibrosis and cirrhosis **(A)**. In the other three models, we compared the combination of Agile 3+, FIB-4 and NFS scores with Agile 4 scores respectively, and also used lower thresholds and upper thresholds to screen patients with advanced fibrosis and cirrhosis **(B, C, D)**. Two pairs of different thresholds were selected for this hybrid strategy: the lower threshold corresponds to 90% sensitivity; the upper threshold corresponds to 90% specificity. The table next to each panel contains the number and proportion of patients in each TP, TN, FP, and FN groups.

Figure 3. Forest plot for the AUROC of ADAPT + Agile 4 in different subgroups by

gender, age (years), diabetes status, ALT (U/L), BMI (kg/m^2), and NAS.

Figure 4. Flow chart for the 3-step serial combination of tests in the sequential algorithm. The FAST score is used as the first test, the ADAPT score is used as the second test, and the Agile 4 score is used as the third test to evaluate and judge the degree of liver inflammation and fibrosis in the biopsy cohort of MAFLD patients.

Supplementary Figure 1. The AUROC of FAST, ADAPT and Agile 4 scores in the derivation cohort.

Supplementary Figure 2. The dual cut-off approach of FAST, ADAPT and Agile 4 scores in the derivation cohort.