Research Article
MASLD and Alcohol-Related Liver Diseases

JOURNAL
OF HEPATOLOGY

Prognostic performance of the two-step clinical care pathway in metabolic dysfunction-associated steatotic liver disease

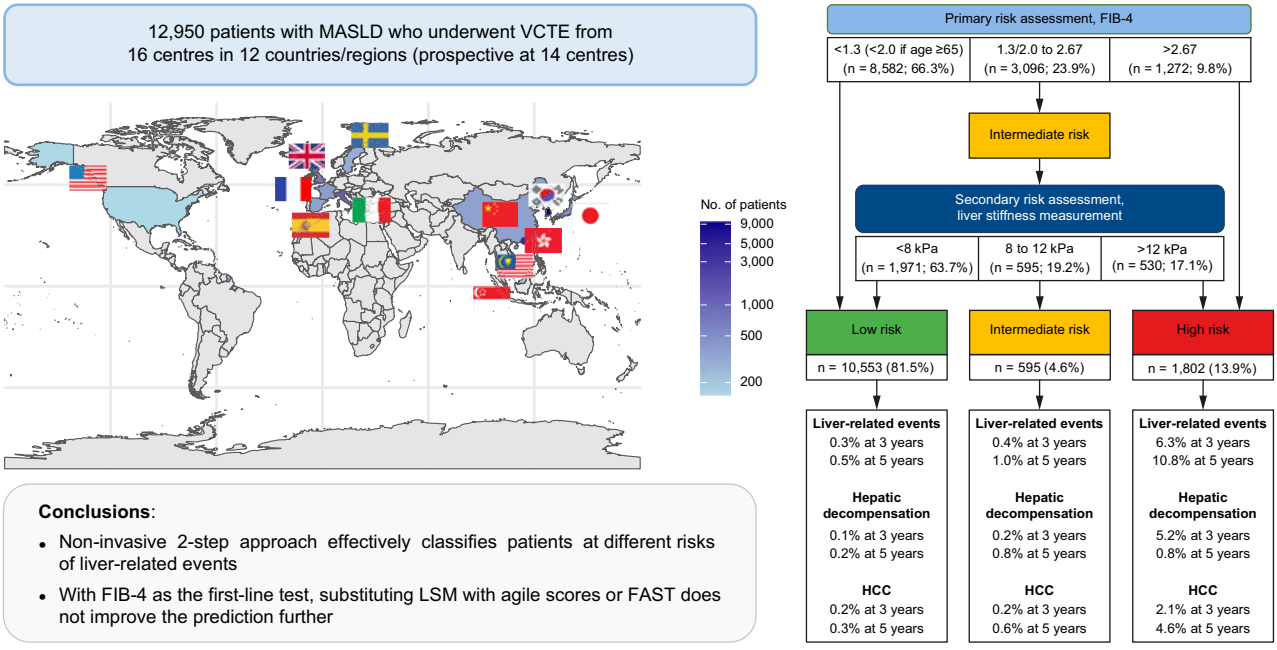
Authors

Terry Cheuk-Fung Yip, Hye Won Lee, Huapeng Lin, …, Manuel Romero-Gomez, Seung Up Kim, Vincent Wai-Sun Wong

Correspondence

KSUKOREA@yuhs.ac (S.U. Kim), wongv@cuhk.edu.hk (V.W.-S. Wong).

Graphical abstract



Highlights

- The FIB-4/LSM two-step approach can predict liver-related events.
- Reserving LSM for patients with intermediate FIB-4 is similarly prognostic as LSM for all patients.
- Substituting LSM by the Agile or FAST scores does not improve prognostication further.

Impact and implications

Metabolic dysfunction-associated steatotic liver disease (MASLD) is emerging as one of the leading causes of cirrhosis and hepatocellular carcinoma worldwide, but only a minority of patients will develop these complications. Therefore, it is necessary to use non-invasive tests instead of liver biopsy for risk stratification. Additionally, as most patients with MASLD are seen in primary care instead of specialist settings, cost and availability of the tests should be taken into consideration. In this multicentre study, the use of the Fibrosis-4 index followed by liver stiffness measurement by vibration-controlled transient elastography effectively identified patients who would later develop liver-related events. The results support current recommendations by various regional guidelines on a clinical care pathway based on non-invasive tests to diagnose advanced liver fibrosis.

https://doi.org/10.1016/j.jhep.2025.01.014
 J. Hepatol. 2025, 83, 304–314

Research Article
MASLD and Alcohol-Related Liver Diseases

JOURNAL
OF HEPATOLOGY

Prognostic performance of the two-step clinical care pathway in metabolic dysfunction-associated steatotic liver disease

Terry Cheuk-Fung Yip[1,2,†], **Hye Won Lee**[3,†], **Huapeng Lin**[4,†], Emmanuel Tsochatzis[5], Salvatore Petta[6], Elisabetta Bugianesi[7], Masato Yoneda[8], Ming-Hua Zheng[9], Hannes Hagström[10,11], Jérôme Boursier[12,13], José Luis Calleja[14], George Boon-Bee Goh[15], Wah-Kheong Chan[16], Rocio Gallego-Durán[17], Arun J. Sanyal[18], Victor de Lédinghen[19], Philip N. Newsome[20], Jian-Gao Fan[21], Laurent Castéra[22], Michelle Lai[23], Céline Fournier-Poizat[19], Grace Lai-Hung Wong[1,2], Grazia Pennisi[6], Angelo Armandi[7], Atsushi Nakajima[8], Wen-Yue Liu[24], Ying Shang[10], Marc de Saint-Loup[12], Elba Llop[14], Kevin Kim Jun Teh[15], Carmen Lara-Romero[17], Amon Asgharpour[18], Sara Mahgoub[25], Mandy Sau-Wai Chan[19], Clemence M. Canivet[12,13], Manuel Romero-Gomez[17], Seung Up Kim[3,*], Vincent Wai-Sun Wong[1,2,*]

Journal of Hepatology **2025**. vol. 83 | 304–314

Background & Aims: Current guidelines recommend a two-step approach for risk stratification in patients with metabolic dysfunction-associated steatotic liver disease (MASLD) involving Fibrosis-4 index (FIB-4) followed by liver stiffness measurement (LSM) by vibration-controlled transient elastography (VCTE) or similar second-line tests. This study aimed to examine the prognostic performance of this approach.

Methods: The VCTE-Prognosis study was a longitudinal study of patients with MASLD who had undergone VCTE examinations at 16 centres from the US, Europe and Asia with subsequent follow-up for clinical events. The primary endpoint was incident liver-related events (LREs), defined as hepatic decompensation and/or hepatocellular carcinoma.

Results: Of 12,950 patients (mean age 52 years, 41% female, 12.1% LSM >12 kPa), baseline FIB-4, at cut-offs of 1.3 (or 2.0 for age ≥65) and 2.67, classified 66.3% as low-risk and 9.8% as high-risk, leaving 23.9% in the intermediate-risk zone. After classifying intermediate FIB-4 patients as low-risk if LSM was <8.0 kPa and high-risk if LSM was >12.0 kPa, 81.5%, 4.6%, and 13.9% of the full cohort were classified as low-, intermediate-, and high-risk, respectively. At a median (IQR) follow-up of 47 (23-72) months, 248 (1.9%) patients developed LREs. The 5-year cumulative incidence of LREs was 0.5%, 1.0% and 10.8% in the low-, intermediate- and high-risk groups, respectively. Replacing LSM with Agile 3+, Agile 4, and FAST did not reduce the intermediate-risk zone or improve event prediction. Classifying intermediate FIB-4 patients by LSM <10 kPa (low-risk) and >15 kPa (high-risk) reduced the intermediate-risk zone while maintaining predictive performance.

Conclusions: The non-invasive two-step approach of FIB-4 followed by LSM is effective in classifying patients at different risks of LREs.

Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) affects over 30% of the general population and is one of the leading causes of cirrhosis and hepatocellular carcinoma (HCC).[1] Because of the large number of patients and the fact that only a small proportion of them will ultimately develop liver-related complications, it is important to use affordable and widely available non-invasive tests to assess the severity of MASLD,[2] especially as the majority of patients are seen in primary care instead of specialist settings.[3] Therefore, in 2021, the American Gastroenterological Association (AGA) recommended a two-step clinical care pathway for the assessment of MASLD.[4] The Fibrosis-4 index (FIB-4), based on age, aspartate aminotransferase (AST), alanine aminotransferase (ALT) and platelet count, serves as the inexpensive first-line test. Patients with low FIB-4 <1.3 have a low risk of advanced liver fibrosis and can be safely monitored by primary care. Patients with high FIB-4 >2.67 should be referred to hepatologists, whereas those with intermediate FIB-4 of 1.3-2.67 should undergo a specific second-line test such as vibration-controlled transient elastography (VCTE) or the enhanced liver fibrosis score. This approach has been largely adopted by other liver and endocrine societies.[5–8]

In the past few years, these non-invasive tests of liver fibrosis have been shown to be useful for not only diagnosing

* Corresponding authors. Addresses: Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea; Tel.: 82-2-2228-1944. (S.U. Kim), or Department of Medicine and Therapeutics, Prince of Wales Hospital, Shatin, Hong Kong, China; Tel.: 852-35054205.
E-mail addresses: KSUKOREA@yuhs.ac (S.U. Kim), wongv@cuhk.edu.hk (V.W.-S. Wong).
† These authors contributed equally to this work.
https://doi.org/10.1016/j.jhep.2025.01.014

advanced fibrosis but also predicting future liver-related events.[9–11] However, a number of practical questions remain. In particular, it is unclear if the two-step approach performs well in detecting patients at risk of liver-related events, and whether more widespread use of specific second-line tests can significantly improve prognostication. The use of dual FIB-4 cut-offs can also be confusing, and it is unclear if patients with intermediate and high FIB-4 should be handled differently. Furthermore, the Agile 3+ score, based on liver stiffness measurement (LSM) by VCTE, platelet count, AST, ALT, diabetes, sex and age, is more accurate than VCTE alone in diagnosing advanced fibrosis and classifies fewer patients in the intermediate zone.[12] Whether it is a better second-line test than VCTE alone deserves further evaluation.

For these reasons, we aimed to evaluate the prognostic performance of the two-step clinical care pathway and determine the best approach in the use of non-invasive tests.

Patients and methods

Study design and participants

This was a cohort study of patients with MASLD who had undergone VCTE examination at 16 tertiary centres from the US, Europe, and Asia, with data collected prospectively at 14 centres. Details of the study design have been reported previously.[10] In brief, we included adult patients aged 18 years or older with hepatic steatosis diagnosed by histologic methods or imaging studies and available FIB-4 and VCTE results. Most patients had hepatic steatosis diagnosed by abdominal ultrasonography. We excluded patients with other liver diseases such as chronic viral hepatitis, HIV infection, excessive alcohol consumption (>30 g/day in men and >20 g/day in women), secondary causes of hepatic steatosis, or a history of HCC, hepatic decompensation, liver resection, liver transplant, or other malignancies.

The study protocol was approved by the institutional review boards of the participating sites and conducted in accordance with the principles of the Declaration of Helsinki. The patients provided informed written consent for the prospective programs at the local sites, but consenting for the current secondary analysis was waived.

Assessments

At each clinic visit, the investigators recorded the medical history. BMI was calculated as weight (kg) divided by height (m) squared. A venous blood sample was taken after at least 8 h of fasting for kidney function, liver biochemistry, complete blood count, glucose and lipids. Liver stiffness was measured using the VCTE machine (FibroScan, Echosens, Paris, France) by trained operators, and the patients needed to have at least 10 valid acquisitions.[13] FIB-4, the Agile scores and the FibroScan-aspartate aminotransferase (FAST) score were calculated according to published formulas.[12,14,15] Among patients with liver biopsy, histological fibrosis stage was determined according to the Nonalcoholic Steatohepatitis Clinical Research Network scoring system.

Definition of non-invasive approaches

The AGA two-step approach was defined as performing FIB-4 in all patients, followed by LSM by VCTE using the cut-offs of 8 and 12 kPa among patients with FIB-4 1.3 (or 2.0 if age ≥65 years) to 2.67. Low-risk patients were patients with FIB-4 <1.3 (or 2.0 if age ≥65 years) or FIB-4 1.3/2.0-2.67 but LSM <8 kPa, while high-risk patients were those with FIB-4 >2.67 or FIB-4 1.3/2.0-2.67 and LSM >12 kPa. The remaining patients were classified as intermediate risk. In the AGA clinical care pathway, patients at intermediate risk were advised to be referred to hepatologists or receive a risk reassessment after 2-3 years.[4] Other two-step algorithms proposed by EASL and AASLD give slightly different recommendations for patients at intermediate risk in terms of the interval of the reassessment.[6,7] Modified approaches examined included i) All patients only received FIB-4; ii) Using cut-offs of 10 and 15 kPa instead of 8 and 12 kPa for LSM; iii) All patients received VCTE instead of FIB-4 in the first step; vi) Patients with FIB-4 ≥1.3 (or 2.0 if age ≥65 years) received VCTE in the second step; v) Replacing VCTE with Agile 3+ in the second step; vi) Replacing VCTE with Agile 4 in the second step; vii) Replacing VCTE with FAST in the second step; and viii) Replacing FIB-4 with the LiverRisk score in the first step.

Outcomes

The primary outcome was a composite endpoint of liver-related events including HCC, hepatic decompensation (ascites, variceal haemorrhage, hepatic encephalopathy or hepatorenal syndrome), liver transplant, and liver-related death. Secondary outcomes included HCC and hepatic decompensation, analysed separately. The diagnosis of the events was based on prospective follow-up, medical record review, or validated registries with positive predictive values (PPVs) of at least 90%.

Statistical analysis

The baseline date was defined as the latter of the date of first VCTE and the date of blood tests to avoid immortal time bias, as previously described.[10] Data were analysed using R (4.3.1, R Core Team 2023). Continuous variables were expressed in mean (SD) or median (25^{th} to 75^{th} percentile [P25-P75]), as appropriate, while categorical variables were presented as number (percentage). Cumulative incidence function of primary and secondary outcomes was estimated and compared by Gray's method and Gray's test, respectively; non-liver-related death was treated as a competing event for liver-related events and HCC, while non-liver-related death and HCC were treated as competing events for hepatic decompensation, as previously described.[10] The discriminatory performance of non-invasive approaches was assessed by time-dependent AUCs, accounting for competing events.[16] Integrated time-dependent AUC was used to summarise the time-dependent AUCs over 5 years of follow-up, calculated as an average of time-dependent AUCs weighted by the estimated probability density of the primary or secondary outcomes during follow-up; 95% CIs were estimated using nonparametric bootstrapping with 1,000 bootstrap samples. Integrated Brier score was used to assess the overall accuracy of the non-invasive approaches, respectively. The prognostic performance in terms of time-dependent sensitivity, specificity, PPV, and negative predictive value (NPV) at 3 and 5 years were evaluated, accounting for competing risks based on Fine and Gray's method. The modified approaches were evaluated for continuous net reclassification improvement (NRI) and integrated discrimination improvement index with reference to the original AGA two-step approach at 3 and 5 years using the inverse probability weighting estimator.[17,18] Decision curve analysis was used to assess the clinical benefit of using two-

step algorithms to inform the decision of referral to secondary care. It was performed based on the estimated incidence of liver-related events at 5 years by the two-step algorithms using Fine and Gray subdistribution hazard model. As the PPV and NPV of the two-step approach depend on the pre-test probability of the patient population, a simulation was performed as a sensitivity analysis to investigate the change in PPV and NPV under different risks of liver-related events in the patient population before performing the two-step approach to mimic different clinical settings. Subgroup analysis was performed among patients with liver biopsy. The AUC of detecting histological advanced fibrosis was estimated for each non-invasive approach with 95% CIs. All statistical tests were two-sided. Statistical significance was taken as p <0.05.

Results

Participants

From February 2004 to January 2023, we identified 17,949 patients with at least one VCTE examination. After excluding 4,999 patients according to the inclusion and exclusion criteria, 12,950 patients with both FIB-4 and VCTE results were included in the final analysis (Fig. 1). Their mean (SD) age was 51.7 (13.9) years; 5,316 (41.1%) were women (Table 1). A total of 4,429 patients (34.2%) had diabetes and 4,835 (37.3%) had hypertension. A total of 2,828 patients (21.8%) were from the US or Europe, while 10,122 (78.2%) were from Asia. Among 3,065 patients with liver biopsy, 1,023 (33.4%) had F3 or F4 fibrosis.

Events

At a median (P25-P75) follow-up of 47.4 (23.3 to 72.3) months, 248 (1.9%) patients developed liver-related events, including 174 (1.3%) patients developing hepatic decompensation and 109 (0.8%) developing HCC. In patients who developed liver-related events, 16.5%, 25.0%, and 58.5% had a controlled attenuation parameter <248 dB/m, 248-279 dB/m, and ≥280 dB/m, respectively. The 3- and 5-year cumulative incidences (95% CI) of liver-related events were 1.2% (1.0%-1.4%) and 2.0% (1.8%-2.4%), respectively. The 3- and 5-year cumulative incidences (95% CI) of hepatic decompensation were 0.9% (0.7%-1.1%) and 1.4% (1.1%-1.6%), respectively. The 3- and 5-year cumulative incidences (95% CI) of HCC were 0.5% (0.4%-0.6%) and 1.0% (0.8%-1.2%), respectively.

Performance of the two-step approach

Among 12,950 patients with MASLD, 8,582 (66.3%), 3,096 (23.9%), and 1,272 (9.8%) had FIB-4 <1.3 (or 2.5 if age ≥65

Included centres	Countries/ regions	No. of patients
Yonsei University School of Medicine	South Korea	9,556
The Chinese University of Hong Kong	Hong Kong	4,037
Royal Free Hospital	UK	709
University of Palermo	Italy	680
University of Torino	Italy	503
University of Yokohama	Japan	474
Wenzhou Medical University	China	366
Karolinska University Hospital	Sweden	302
Angers University Hospital	France	283
Singapore General Hospital	Singapore	211
University of Malaya	Malaysia	201
Hospital Puerta de Hierro	Spain	180
University of Seville	Spain	172
Virginia Commonwealth University	USA	161
University of Bordeaux	France	99
University of Birmingham	UK	15

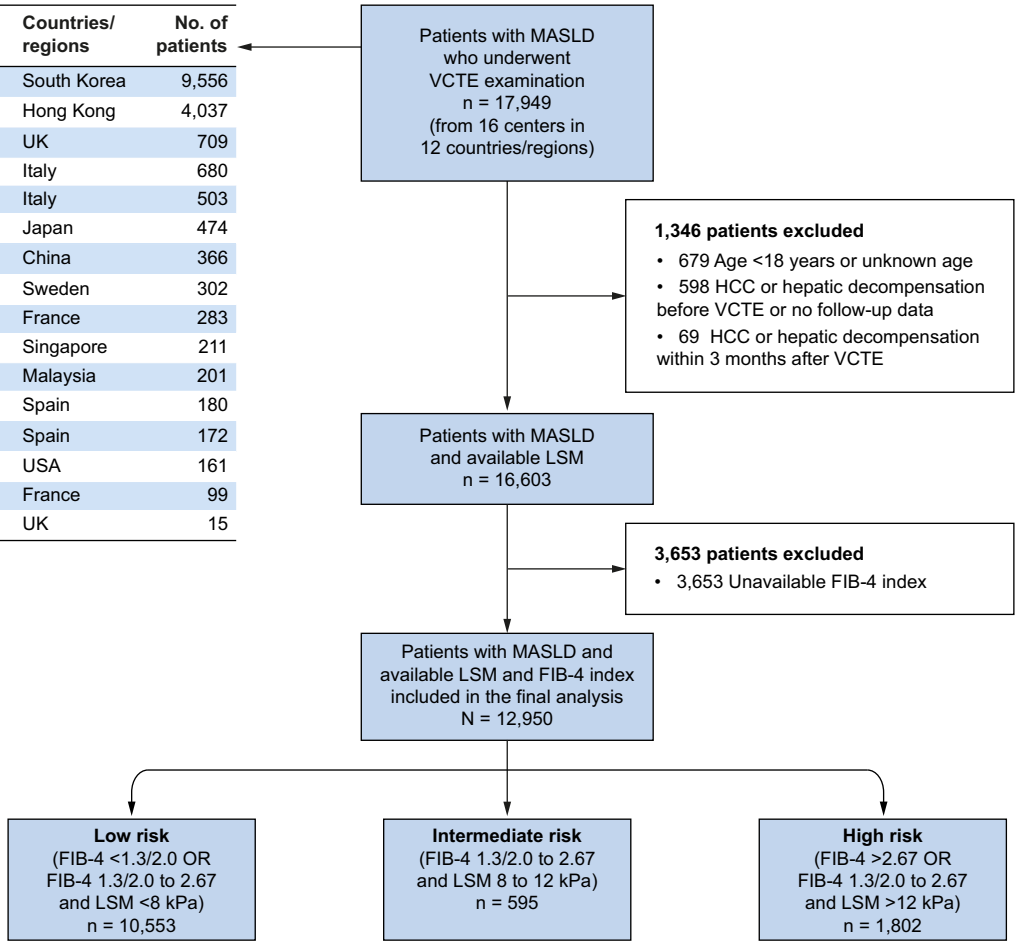


Fig. 1. Study participant flow. FIB-4, Fibrosis-4 index; HCC, hepatocellular carcinoma; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; VCTE, vibration-controlled transient elastography.

Table 1. Baseline characteristics of patients with metabolic dysfunction-associated steatotic liver disease.

Characteristics	All patients N = 12,950
Age (years)	51.7 (13.9)
Female sex, n (%)	5,316 (41.1)
Male sex, n (%)	7,634 (58.9)
BMI (kg/m^2)	27.2 (24.7 to 30.4)
Diabetes, n (%)	4,429 (34.2)
Hypertension, n (%)	4,835 (37.3)
ALT (IU/L)	38 (24 to 64)
AST (IU/L)	31 (23 to 47)
GGT (IU/L)	44 (27 to 78)
Albumin (g/L)	44.8 (3.8)
Total bilirubin (μmol/L)	12.0 (8.6 to 16.0)
Platelet (×10^9/L)	239 (200 to 282)
Creatinine (μmol/L)	72 (60 to 83)
FibroScan	
Liver stiffness measurement (kPa)	5.9 (4.6 to 8.3)
Controlled attenuation parameter (dB/m)	303 (274 to 335)
Non-invasive tests	
Fibrosis-4 index	1.11 (0.74 to 1.71)
LiverRisk score	6.29 (5.42 to 7.48)
Agile 3+	0.16 (0.06 to 0.44)
Agile 4	0.01 (0.00 to 0.06)
FibroScan-AST	0.28 (0.12 to 0.52)
Fibrosis stage[a] n = 3,065	
0	481 (15.7)
1	1,055 (34.4)
2	506 (16.5)
3	658 (21.5)
4	365 (11.9)
Follow-up duration (months)	47.4 (23.3 to 72.3)

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyltransferase.
Data were presented as n (%), mean (standard deviation), or median (25th to 75th percentile).
[a]Fibrosis stage was defined by histology using the Nonalcoholic Steatohepatitis Clinical Research Network scoring system.

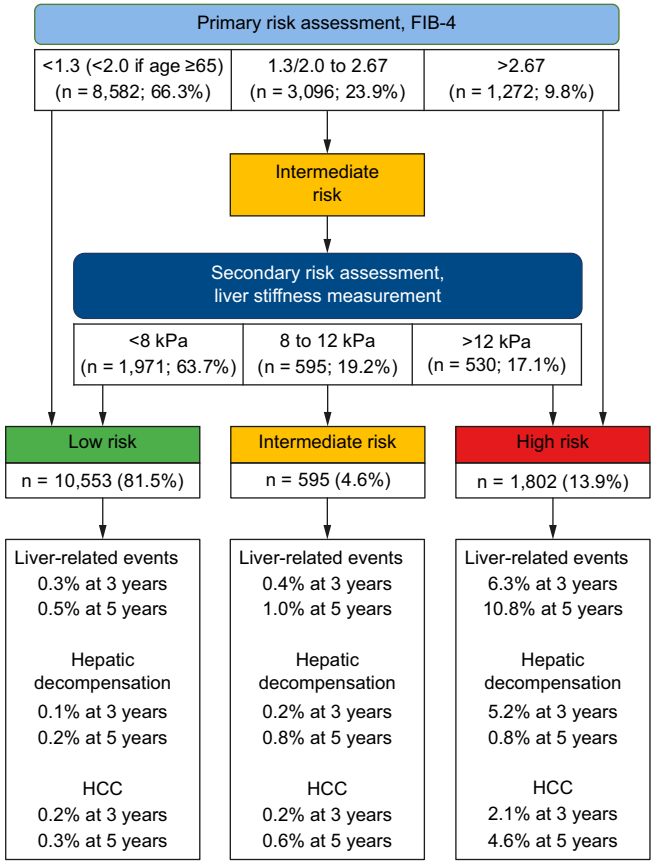


Fig. 2. Categorisation of patients with metabolic dysfunction-associated steatotic liver disease by the two-step approach. FIB-4, Fibrosis-4 index; HCC, hepatocellular carcinoma. (This figure appears in color on the web.)

years), 1.3/2.0 to 2.67, and >2.67, respectively (Fig. 2). If the 3,096 patients with intermediate FIB-4 of 1.3/2.0 to 2.67 were to undergo VCTE examination, 1,971 (63.7%) patients would be reclassified as low-risk (LSM <8 kPa), whereas 530 (17.1%) would be reclassified as high-risk (LSM >12 kPa).

Thus, the two-step approach classified 10,553 (81.5%) patients as low-risk, 595 (4.6%) as intermediate-risk, and 1,802 (13.9%) as high-risk. The two-step approach without the age-specific cut-off for FIB-4 classified 10,356 (80.0%) patients as low-risk, 720 (5.5%) as intermediate-risk, and 1,874 (14.5%) as high-risk (Fig. S1A).

According to the classification of the two-step approach, the 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.5% in the low-risk group, 0.4% and 1.0% in the intermediate-risk group, and 6.3% and 10.8% in the high-risk group (Fig. S2A). Corresponding results for hepatic decompensation and HCC are shown in Fig. S2B,C. Of the 25 patients who developed liver-related events in 3 years in the low-risk group, 17 patients had HCC and 10 patients had hepatic decompensation (two developed both HCC and hepatic decompensation). At 5 years, the sensitivity and NPV of the low-risk classification for excluding liver-related events were 79.0% and 99.5%, respectively (Table 2). The specificity and PPV of the high-risk classification in predicting liver-related events were 91.1% and 12.3%, respectively. By simulation of a different prevalence of advanced fibrosis and risk of liver-related events in the patient population before performing the two-step approach, *i.e.*, the pre-test probability, the PPV of the two-step approach increased while NPV decreased when the pre-test probability increased, and vice versa (Fig. S3A,B). When the underlying 5-year risk of liver-related events increased to 3%, the estimated PPV and NPV would be 21.9% and 98.8%, respectively. When the underlying 5-year risk of liver-related events dropped to 0.5%, the estimated PPV and NPV would be 4.3% and 99.8%, respectively. The use of LSM cut-offs of 10 and 15 kPa instead of 8 and 12 kPa in the second step yielded a slightly reduced intermediate-risk zone with a higher risk of liver-related events within that zone (Fig. S1B), and a comparable prognostic performance compared to the original AGA two-step approach (Table 2).

Prognostication by expansion of VCTE examination

The key reason for proposing the two-step approach was cost and availability considerations. As VCTE or other specific fibrosis tests might become more affordable and widely available in the future, it is important to study if expanding the use of specific fibrosis tests to more patients might improve prognostication further. If VCTE was performed on all 12,950 patients, 9,420 (72.7%), 1,966 (15.2%), and 1,564 (12.1%) would have LSM <8 kPa (low-risk), 8-12 kPa (intermediate-risk) and >12 kPa (high-risk), respectively (Fig. S1C). The 3- and 5-year

Table 2. Prognostic performance of the non-invasive two-step approach and alternative approaches in patients with metabolic dysfunction-associated steatotic liver disease.

Category	N (%)	Number of liver-related events (n, %)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Overall accuracy (%) (95% CI)
3 years							
FIB-4 for all patients	12,950	122					
Low-risk (FIB-4 <1.3/2.0 if age ≥65 years)	8,582 (66.3)	22 (0.3)	81.6 (74.3-88.6)	64.9 (63.9-65.8)	2.7 (2.2-3.2)	99.7 (99.5-99.8)	65.1 (64.2-66.1)
Intermediate-risk zone	3,096 (23.9)	29 (0.9)	—	—	—	—	—
High-risk (FIB-4 >2.67)	1,272 (9.8)	71 (5.6)	57.9 (49.0-67.0)	90.2 (89.6-90.8)	6.6 (5.2-8.1)	99.4 (99.3-99.6)	89.7 (89.1-90.3)
Non-invasive two-step approach	12,950	122					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and LSM <8 kPa)	10,553 (81.5)	25 (0.2)	79.0 (71.3-86.3)	81.0 (80.2-81.7)	4.7 (3.9-5.7)	99.7 (99.6-99.8)	80.9 (80.1-81.7)
Intermediate-risk zone	595 (4.6)	2 (0.3)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and LSM >12 kPa)	1,802 (13.9)	95 (5.3)	77.2 (69.3-85.0)	86.0 (85.3-86.7)	6.2 (5.0-7.5)	99.7 (99.6-99.8)	85.9 (85.2-86.6)
Non-invasive two-step approach with cut-offs of 10 and 15 kPa for LSM	12,950	122					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and LSM <10 kPa)	10,896 (84.1)	26 (0.2)	78.1 (70.1-85.5)	83.8 (83.1-84.6)	5.4 (4.4-6.6)	99.7 (99.6-99.8)	83.8 (83.0-84.5)
Intermediate-risk zone	445 (3.4)	6 (1.3)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and LSM >15 kPa)	1,609 (12.4)	90 (5.6)	73.0 (65.3-81.1)	87.6 (86.9-88.2)	6.5 (5.3-7.9)	99.6 (99.5-99.8)	87.4 (86.7-88.1)
VCTE for all patients	12,950	122					
Low-risk (LSM <8 kPa)	9,420 (72.7)	17 (0.2)	86.0 (79.6-91.8)	71.9 (70.9-72.8)	3.5 (2.9-4.2)	99.8 (99.7-99.9)	72.1 (71.1-73.0)
Intermediate-risk zone	1,966 (15.2)	13 (0.7)	—	—	—	—	—
High-risk (LSM >12 kPa)	1,564 (12.1)	92 (5.9)	74.3 (65.7-81.8)	87.8 (87.0-88.4)	6.8 (5.4-8.2)	99.7 (99.5-99.8)	87.6 (86.8-88.2)
VCTE for patients with FIB-4 ≥1.3/2.0	12,950	122					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 ≥1.3/2.0 and LSM <8 kPa)	10,980 (84.8)	26 (0.2)	78.2 (70.3-85.7)	84.3 (83.6-85.0)	5.6 (4.6-6.7)	99.7 (99.6-99.8)	84.2 (83.5-85.0)
Intermediate-risk zone	886 (6.8)	10 (1.1)	—	—	—	—	—
High-risk (FIB-4 ≥1.3/2.0 and LSM >12 kPa)	1,084 (8.4)	86 (7.9)	69.2 (60.2-77.4)	91.7 (91.1-92.2)	9.0 (7.3-10.9)	99.6 (99.5-99.7)	91.4 (90.8-92.0)
Non-invasive two-step approach with Agile 3+ as the second step	12,948	122					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and Agile 3+ <0.451)	10,374 (80.1)	25 (0.2)	78.9 (71.3-86.4)	79.2 (78.3-80.0)	4.3 (3.5-5.2)	99.7 (99.6-99.8)	79.2 (78.3-80.0)
Intermediate-risk zone	594 (4.6)	4 (0.7)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and Agile 3+ >0.678	1,980 (15.3)	93 (4.7)	75.5 (67.9-83.4)	84.2 (83.5-85.0)	5.4 (4.4-6.5)	99.7 (99.5-99.8)	84.1 (83.4-84.9)
Non-invasive two-step approach with Agile 4 as the second step	12,948	122					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and Agile 4 <0.251)	11,232 (86.7)	30 (0.3)	74.6 (67.0-82.4)	86.6 (85.9-87.3)	6.2 (5.1-7.5)	99.7 (99.5-99.8)	86.5 (85.8-87.2)
Intermediate-risk zone	324 (2.5)	10 (3.1)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and Agile 4 >0.564	1,392 (10.8)	82 (5.9)	66.5 (57.9-74.9)	89.3 (88.6-89.9)	6.9 (5.5-8.4)	99.6 (99.4-99.7)	89.0 (88.3-89.6)
Non-invasive two-step approach with FAST as the second step	11,345	102					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and FAST ≤0.35)	8,706 (76.7)	19 (0.2)	80.6 (72.9-88.1)	76.9 (76.0-77.8)	3.8 (3.1-4.7)	99.7 (99.6-99.8)	76.9 (76.0-77.9)
Intermediate-risk zone	992 (8.7)	12 (1.2)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and FAST ≥0.67	1,647 (14.5)	71 (4.3)	68.2 (58.4-77.4)	86.0 (85.2-86.7)	5.3 (4.2-6.6)	99.6 (99.4-99.7)	85.8 (85.0-86.5)
Non-invasive two-step approach with LiverRisk score as the first step	11,023	103					
Low-risk (LiverRisk score <6 OR LiverRisk score 6 to 10 and LSM <8 kPa)	8,396 (76.2)	15 (0.2)	85.3 (77.3-91.5)	75.9 (74.9-76.9)	4.0 (3.2-4.8)	99.8 (99.6-99.9)	76.1 (75.0-77.0)
Intermediate-risk zone	1,093 (9.9)	5 (0.5)	—	—	—	—	—
High-risk (LiverRisk score ≥10 OR LiverRisk score 6 to 10 and LSM >12 kPa)	1,534 (13.9)	83 (5.4)	79.7 (69.9-86.5)	86.4 (85.6-87.2)	6.4 (5.1-7.8)	99.7 (99.6-99.8)	86.3 (85.5-87.1)
5 years							
FIB-4 for all patients	12,950	248					
Low-risk (FIB-4 <1.3/2.0 if age ≥65 years)	8,582 (66.3)	46 (0.5)	82.8 (76.8-88.5)	65.4 (64.3-66.5)	4.7 (4.0-5.6)	99.5 (99.3-99.6)	65.9 (64.7-66.9)
Intermediate-risk zone	3,096 (23.9)	57 (1.8)	—	—	—	—	—
High-risk (FIB-4 >2.67)	1,272 (9.8)	145 (11.4)	60.3 (52.7-67.8)	91.1 (90.4-91.8)	12.3 (10.1-14.8)	99.1 (98.9-99.3)	90.3 (89.5-91.0)
Non-invasive two-step approach	12,950	248					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and LSM <8 kPa)	10,553 (81.5)	54 (0.5)	79.0 (72.8-85.3)	82.2 (81.3-83.1)	8.4 (7.1-9.9)	99.5 (99.3-99.6)	82.2 (81.3-83.0)
Intermediate-risk zone	595 (4.6)	8 (1.3)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and LSM >12 kPa)	1,802 (13.9)	186 (10.3)	76.5 (69.9-83.1)	87.2 (86.4-88.0)	11.0 (9.2-12.9)	99.4 (99.3-99.6)	86.9 (86.1-87.8)
Non-invasive two-step approach with cut-offs of 10 and 15 kPa for LSM	12,950	248					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and LSM <10 kPa)	10,896 (84.1)	58 (0.5)	77.6 (71.3-84.3)	85.0 (84.1-85.8)	9.7 (8.1-11.3)	99.5 (99.3-99.6)	84.8 (84.0-85.6)

(*continued on next page*)

Table 2. (*continued*)

Category	N (%)	Number of liver-related events (n, %)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Overall accuracy (%) (95% CI)
Intermediate-risk zone	445 (3.4)	11 (2.5)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and LSM >15 kPa)	1,609 (12.4)	179 (11.1)	72.6 (65.9-79.3)	88.6 (87.9-89.4)	11.7 (9.8-13.7)	99.4 (99.2-99.6)	88.3 (87.5-89.0)
VCTE for all patients	12,950	248					
Low-risk (LSM <8 kPa)	9,420 (72.7)	35 (0.4)	84.7 (79.0-89.7)	74.2 (73.2-75.1)	6.4 (5.4-7.5)	99.6 (99.4-99.7)	74.4 (73.4-75.4)
Intermediate-risk zone	1,966 (15.2)	23 (1.2)	—	—	—	—	—
High-risk (LSM >12 kPa)	1,564 (12.1)	190 (12.1)	75.1 (68.3-81.2)	89.1 (88.4-89.8)	12.5 (10.5-14.7)	99.4 (99.2-99.6)	88.8 (88.0-89.5)
VCTE for patients with FIB-4 ≥1.3/2.0	12,950	248					
Low-risk (FIB-4 <1.3/2.0 OR FIB-4 ≥1.3/2.0 and LSM <8 kPa)	10,980 (84.8)	58 (0.5)	77.1 (70.8-83.0)	85.8 (85.0-86.6)	10.1 (8.5-11.9)	99.5 (99.3-99.6)	85.6 (84.8-86.3)
Intermediate-risk zone	886 (6.8)	19 (2.1)	—	—	—	—	—
High-risk (FIB-4 ≥1.3/2.0 and LSM >12 kPa)	1,084 (8.4)	171 (15.8)	69.7 (62.3-76.6)	92.8 (92.2-93.4)	16.7 (14.0-19.8)	99.3 (99.1-99.5)	92.3 (91.6-92.9)
Non-invasive two-step approach with Agile 3+ as the second step	12,948	248					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and Agile 3+ <0.451)	10,374 (80.1)	53 (0.5)	79.7 (73.7-85.8)	80.6 (79.6-81.5)	7.9 (6.6-9.3)	99.5 (99.3-99.6)	80.6 (79.6-81.5)
Intermediate-risk zone	594 (4.6)	9 (1.5)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and Agile 3+ >0.678	1,980 (15.3)	186 (9.4)	75.6 (69.3-82.1)	85.6 (84.7-86.4)	9.8 (8.3-11.6)	99.4 (99.2-99.6)	85.3 (84.5-86.1)
Non-invasive two-step approach with Agile 4 as the second step	12,948	248					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and Agile 4 <0.251)	11,232 (86.7)	64 (0.6)	75.6 (69.0-82.2)	87.9 (87.1-88.7)	11.5 (9.6-13.5)	99.4 (99.2-99.6)	87.6 (86.9-88.4)
Intermediate-risk zone	324 (2.5)	21 (6.5)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and Agile 4 >0.564	1,392 (10.8)	163 (11.7)	65.9 (58.5-73.3)	90.3 (89.6-91.0)	12.4 (10.2-14.7)	99.2 (99.0-99.4)	89.8 (89.1-90.5)
Non-invasive two-step approach with FAST as the second step	11,345	188					
Low risk (FIB-4 <1.3/2.0 OR FIB-4 1.3/2.0 to 2.67 and FAST ≤0.35)	8,706 (76.7)	44 (0.5)	77.1 (69.9-84.7)	77.8 (76.6-78.7)	6.5 (5.3-7.8)	99.4 (99.2-99.6)	77.8 (76.7-78.8)
Intermediate-risk zone	992 (8.7)	22 (2.2)	—	—	—	—	—
High-risk (FIB-4 >2.67 OR FIB-4 1.3/2.0 to 2.67 and FAST ≥0.67	1,647 (14.5)	122 (7.4)	66.2 (58.2-74.5)	87.4 (86.5-88.2)	9.4 (7.6-11.3)	99.2 (99.0-99.5)	86.9 (86.0-87.7)
Non-invasive two-step approach with LiverRisk score as the first step	11,023	206					
Low-risk (LiverRisk score <6 OR LiverRisk score 6 to 10 and LSM <8 kPa)	8,396 (76.2)	31 (0.4)	83.0 (76.3-89.2)	78.4 (77.4-79.6)	7.1 (5.9-8.3)	99.6 (99.4-99.7)	78.5 (77.5-79.7)
Intermediate-risk zone	1,093 (9.9)	12 (1.1)	—	—	—	—	—
High-risk (LiverRisk score ≥10 OR LiverRisk score 6 to 10 and LSM >12 kPa)	1,534 (13.9)	163 (10.6)	78.7 (71.3-85.0)	88.2 (87.4-89.0)	11.6 (9.7-13.6)	99.5 (99.3-99.7)	88.0 (87.2-88.8)

FIB-4, Fibrosis-4 index; LSM, liver stiffness measurement; NPV, negative predictive value; PPV, positive predictive value; VCTE, vibration-controlled transient elastography.
The two sets of time-dependent sensitivity, specificity, PPV, and NPV referred to low risk *vs.* intermediate/high risk and low/intermediate *vs*. high risk in each of the algorithms.

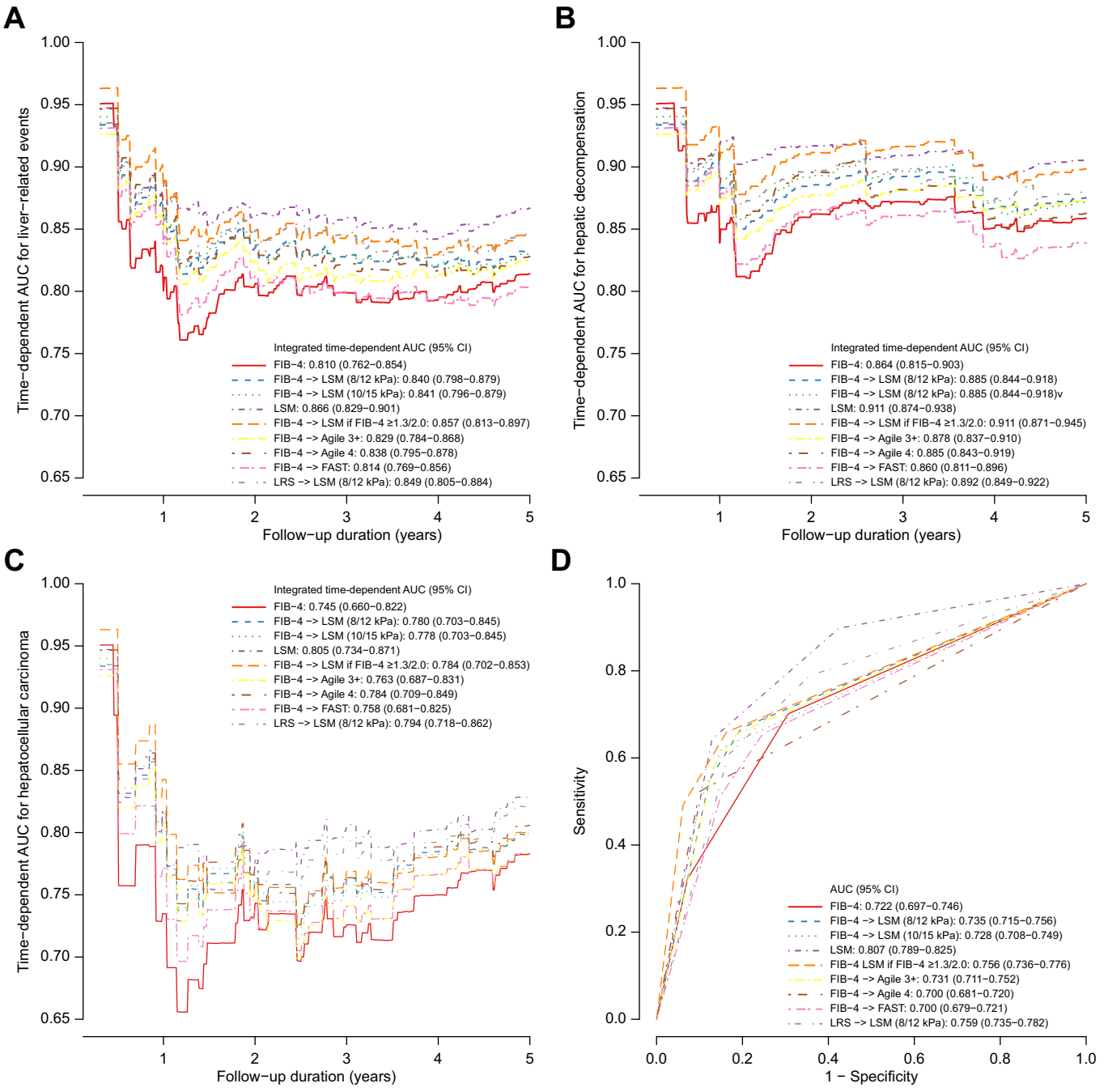


Fig. 3. Integrated time-dependent AUC of different approaches in 10,112 patients with all scores available. (A) Liver-related events, (B) hepatic decompensation, and (C) hepatocellular carcinoma. (D) AUC of different two-step approaches in 1,587 patients with all scores and liver histology available for diagnosing histological advanced fibrosis (F3-F4). FAST, FibroScan-aspartate aminotransferase; FIB-4, Fibrosis-4 index; HCC, hepatocellular carcinoma; LRS, LiverRisk score; LSM, liver stiffness measurement. (This figure appears in color on the web.)

cumulative incidences of liver-related events were 0.2% and 0.4% in the low-risk group, 0.9% and 1.3% in the intermediate-risk group, and 7.0% and 12.2% in the high-risk group. At 5 years, the sensitivity and NPV of LSM <8 kPa in excluding liver-related events were 84.7% and 99.6%, respectively (Table 2). The specificity and PPV of LSM >12 kPa in predicting liver-related events were 89.1% and 12.5%, respectively.

If VCTE was performed on all patients with FIB-4 ≥1.3 (or 2.0 if age ≥65 years) (n = 4,368) instead of those with intermediate FIB-4 of 1.3/2.0-2.67, 10,980 (84.8%), 886 (6.8%) and 1,084 (8.4%) would be classified as low, intermediate, and high risk (Fig. S1D). The 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.6% in the low-risk group, 1.5% and 2.1% in the intermediate-risk group, and 9.4% and

16.2% in the high-risk group, respectively. At 5 years, the sensitivity and NPV of FIB-4 <1.3/2.0 or LSM <8 kPa for excluding liver-related events were 77.1% and 99.5%, respectively (Table 2). The specificity and PPV of FIB-4 ≥1.3/2.0 and LSM >12 kPa for predicting liver-related events were 92.8% and 16.7%, respectively. On the other hand, among patients with FIB-4 <1.3/2.0 (n = 8,582), the 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.5% (Fig. S1E). Among the 8,582 patients, 7,022 (81.8%) and 1,560 (18.2%) had an LSM <8 kPa and ≥8 kPa, respectively; the 3- and 5-year cumulative incidence of liver-related events for those with LSM <8 kPa was 0.2% and 0.7%, respectively, while that for those with LSM ≥8 kPa was 0.4% and 1.3% respectively. The 3- and 5-year cumulative incidence of liver-related events of patients with FIB-4 <1.3 and LSM <8 kPa was 0.2% and 0.3%, respectively. Among 1,272 patients with FIB-4 >2.67, 427 (33.6%), 291 (22.9%), and 554 (43.6%) patients had LSM <8 kPa, 8-12 kPa, and >12 kPa, respectively. Their corresponding 5-year cumulative incidences of liver-related events were 1.1%, 4.2%, and 24.2%, respectively. Their corresponding 5-year cumulative incidences of hepatic decompensation were 0.4%, 2.9%, and 18.7%, respectively. Their corresponding 5-year cumulative incidences of HCC were 1.1%, 2.5%, and 9.6%, respectively (Fig. S4A-C).

Prognostication by the use of Agile or FAST scores as the second-line test

As a standalone test, both the Agile 3+ and Agile 4 scores outperformed LSM by VCTE alone in predicting liver-related events.[10] We therefore tested the prognostic performance of a two-step approach using the Agile scores instead of VCTE as the second-line test. In 12,948 patients with FIB-4, VCTE, and complete data for the calculation of the Agile scores, a FIB-4-Agile 3+ two-step algorithm classified 10,374 (80.1%) patients as low risk, 594 (4.6%) as intermediate risk, and 1,980 (15.3%) as high risk (Fig. S1F). The 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.5% in the low-risk group, 0.8% and 1.7% in the intermediate-risk group, and 5.6% and 9.6% in the high-risk group, respectively.

The Agile 4 score was designed to detect cirrhosis.[12] Not surprisingly, the FIB-4-Agile 4 two-step algorithm achieved a high specificity and classified the vast majority (11,232 patients, 86.7%) of patients as low risk, 324 (2.5%) as intermediate risk, and 1,392 (10.8%) as high risk (Table 2, Fig. S1G). The 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.6% in the low-risk group, 3.7% and 7.5% in the intermediate-risk group, and 7.0% and 12.0% in the high-risk group, respectively.

The FAST score was designed to detect at-risk MASH, *i.e.*, MASH with F2-F4 fibrosis.[15] The FIB-4-FAST two-step algorithm classified 8,706 (76.7%) patients as low risk, 992 (8.7%) as intermediate risk, and 1,647 (14.5%) as high risk (Fig. S1H). The 3- and 5-year cumulative incidences of liver-related events were 0.3% and 0.6% in the low-risk group, 1.6% and 2.3% in the intermediate-risk group, and 5.2% and 8.9% in the high-risk group, respectively. Table 2 shows the comparable overall prognostic performance of using Agile or FAST scores as the second-line test compared to the original two-step approach.

The LiverRisk score was designed to detect long-term liver-related outcomes in the general population.[19] The LiverRisk-LSM two-step algorithm classified 8,396 (76.2%) patients as low risk, 1,093 (9.9%) as intermediate risk, and 1,534 (13.9%) as high risk (Fig. S1I). The 3- and 5-year cumulative incidences of liver-related events were 0.2% and 0.4% in the low-risk group, 0.6% and 0.8% in the intermediate-risk group, and 6.5% and 11.0% in the high-risk group, respectively. Table 2 shows the overall prognostic performance comparable to the original AGA two-step approach.

Comparison of different two-step approaches

In 10,112 patients with all scores available, the original AGA two-step approach achieved an integrated time-dependent AUC of 0.840 (95% CI 0.798-0.879) for predicting liver-related events (Fig. 3A). Changing the LSM cut-offs to 10 and 15 kPa resulted in a similar integrated time-dependent AUC of 0.841 (95% CI 0.796-0.879). Expanding LSM to all patients or all patients with FIB-4 ≥1.3/2.0 increased the integrated time-dependent AUC marginally to 0.866 and 0.857, respectively, while decision curve analysis suggested that expanding LSM may improve clinical utility (Fig. S5A-C). Substituting LSM by the Agile 3+, Agile 4 and FAST scores as the second step did not increase the prediction either. The same trend was seen when the endpoints were analysed separately for hepatic decompensation and HCC (Fig. 3B,C). The integrated Brier scores of all six approaches for predicting liver-related events, hepatic decompensation and HCC were similar (Fig. S6A-C). Similarly, pairwise comparisons by net reclassification improvement and integrated discrimination improvement index showed marginal differences between the alternative approaches and the original AGA two-step approach (Table S1).

In the subgroup analyses of patients from Asia and Europe/US, the discriminative performance of the two-step algorithms was comparable to that in the overall analysis (Fig. S7A-F). In a subgroup of 1,587 patients with liver biopsy together with all scores available, the median (P25-P75) FIB-4 and LSM were 1.29 (0.79-2.10) and 8.9 (6.4-13.8), respectively. The original AGA two-step approach achieved an AUC of 0.735 (95% CI 0.715-0.756) in diagnosing F3-F4 fibrosis (Fig. 3D). The use of LSM cut-offs of 10 and 15 kPa instead of 8 and 12 kPa in the second step yielded a similar AUC of 0.728 (95% CI 0.708-0.749). Expanding LSM to all patients or all patients with FIB-4 ≥1.3/2.0 increased the AUC to 0.807 (95% CI 0.789-0.825) and 0.756 (95% CI 0.736-0.776), respectively. In contrast, substituting LSM by the Agile 3+, Agile 4 and FAST scores as the second step did not improve the diagnostic accuracy of F3-F4 fibrosis.

Discussion

In this large longitudinal multicentre study, we demonstrated that the two-step approach is effective in risk stratification and predicting future liver-related events in patients with MASLD. Using FIB-4 as the first step, only 24% of patients required a second-line test such as VCTE-LSM. Importantly, sparing patients from a second-line test had minimal impact on the accuracy of prognostication, suggesting that the two-step approach is equally clinically effective and probably more cost-effective than performing specific fibrosis tests on all patients.

The AGA two-step approach was designed with an emphasis on an effective clinical care pathway between hepatology practice and other disciplines such as primary care,

endocrinology and cardiology.[4] Currently, the American and European liver associations have adopted a similar approach with minor differences in the choices and settings of second-line tests.[6,7] Numerous studies have shown that FIB-4, the first step of the algorithm, has a high NPV in excluding advanced liver fibrosis and future liver-related events in patients with MASLD.[9,20] When used in a longitudinal manner, patients with persistently normal FIB-4 have very low risk of liver-related events.[21,22] Although FIB-4 appears to be confounded by age and diabetes and the optimal cut-off in special populations has been debated,[23–25] it remains a reasonable first-line test at the population level. Using LiverRisk score as the first step identified fewer low-risk individuals, failing to enhance the two-step algorithm's performance; a cut-off beyond 6 might be necessary to classify patients as low risk.

Ever since the publication of the two-step approach, there have been discussions on whether the algorithm should be simplified to improve its adoption. In particular, if the management of patients with intermediate (1.3-2.67) and high (>2.67) FIB-4 is the same (*i.e.*, referral to hepatology or arranging a second-line test), do we need two cut-offs? Along the same line, should a specific fibrosis test be performed in all patients? In our study, LSM for all patients, compared to LSM for patients with intermediate FIB-4 of 1.3-2.67 only, did increase the accuracy in identifying advanced fibrosis using liver histology as the reference standard, but the difference in the prediction of liver-related events was marginal. Likewise, LSM for all patients with FIB-4 ≥1.3 resulted in minimal improvement over LSM for patients with FIB-4 1.3-2.67 in terms of both detection of advanced fibrosis and prognostication. This is in line with the TARGET-NASH experience that a FIB-4 higher than 2.6 is sufficiently prognostic and can be used to guide management.[26] Furthermore, although specific fibrosis tests are routinely used in most tertiary hepatology practices, a concomitant abnormal FIB-4 >1.3 can increase the PPV from 60-80% to over 80-90% in diagnosing advanced fibrosis.[27] Therefore, even in specialist settings, taking FIB-4 into consideration can improve diagnostics and prognostication. In this study, expanding LSM for all patients with FIB-4 ≥1.3 yielded numerically the highest integrated time-dependent AUC and net benefit in the decision curve analysis.

The choice of 8 and 12 kPa as LSM thresholds for risk stratification in the second step of the AGA approach was based on a previous multicentre study and a meta-analysis,[20,28] aiming to rule out and rule in clinically significant fibrosis.[4] Another established pair of cut-offs of LSM to identify patients with advanced fibrosis would be 10 and 15 kPa as suggested by the Baveno VII consensus,[29] aiming to rule in and rule out compensated advanced chronic liver disease (cACLD). An ideal two-step algorithm would prioritise minimising the intermediate risk group while maintaining high NPV and PPV, ensuring a patient-friendly and cost-effective approach (Table S2). In this study, the use of LSM cut-offs of 10 and 15 kPa instead of 8 and 12 kPa in the second step yielded a comparable prognostic performance compared to the original AGA approach, and a slightly smaller intermediate-risk group with a higher risk of liver-related events. Using these higher cut-offs, about 3% more patients were classified as low risk, with a similarly low risk of liver-related events compared to the original AGA approach; only one more liver-related event occurred in the low-risk group (26 *vs*. 25). Regarding implementation, adopting 10 and 15 kPa could be favourable in terms of improved resource allocation and reduced patient anxiety. Aligning the concept of the AGA two-step approach with cACLD in Baveno VII consensus can also be beneficial for encouraging its future adoption. In fact, it may be reasonable to separately consider the threshold for initiating pharmacotherapy and for prognostication. While the cut-offs of 8 and 12 kPa are useful to identify patients with clinically significant fibrosis potentially suitable for pharmacotherapy, higher cut-offs of 10 and 15 kPa could be more suitable for prognostication of future liver-related events. In a previous analysis using the same dataset, our group showed that the VCTE-based Agile 3+ and Agile 4 scores outperformed LSM alone, histological fibrosis stage and simple fibrosis scores in predicting liver-related events.[10] This has been supported by similar studies from Europe and Japan.[30–32] However, when used as a second-line test after FIB-4, the Agile scores were not better than LSM alone in prognostication in the current study. One possible explanation is that the components of FIB-4 and the Agile scores overlap with each other. When the patient population has already been enriched by FIB-4, components of the Agile scores other than LSM are no longer as discriminating. A previous study showed that the specificity and PPV to identify high-risk patients can be further improved if the Agile score and LSM concordantly indicate high risk.[33] Future studies can investigate whether using both LSM and Agile scores as the second-line tests can improve prognostication among high-risk patients.

Our study has the strengths of a large sample size and representation from North America, Europe and Asia. The number of liver-related events allowed for granular analysis of predictors. On the other hand, it also has some limitations. First, all patients were from tertiary referral centres. Some patients developed liver-related events earlier during follow-up. We simulated how the prevalence of advanced fibrosis affects the prognostic performance of the two-step algorithm in a primary care setting with a lower prevalence of advanced fibrosis (Fig. S3A). When the two-step approach is applied in primary care settings, we anticipate that more patients will be classified as low risk, PPV and sensitivity will decrease, and the NPV and specificity will be even higher. Longitudinal data from primary care are important to investigate the two-step approach in the actual context of use. Second, VCTE is only one of the possible second-line tests for MASLD. In other settings, enhanced liver fibrosis test, magnetic resonance elastography, and other liver-specific biomarkers such as MACK-3, PRO-C3, ADAPT and ABC3D may be used to evaluate liver fibrosis as well.[34] Such diagnostic pathways should be evaluated. Third, the median follow-up of 47.4 months, though respectable, remains short compared with the natural history of MASLD. However, as current guidelines recommend interval examinations by non-invasive tests in patients at risk of disease progression,[5–7] the prediction of short-to medium-term prognosis is clinically relevant. Future studies that examine the evolution to intermediate and high risk among patients with low risk at baseline are warranted to inform the optimal reassessment interval.

In conclusion, the two-step approach is effective in predicting liver-related events. Restricting the second fibrosis test to intermediate FIB-4 patients is reasonable when resources are limited. With FIB-4 as the first-line test, substituting LSM by the Agile scores does not improve prediction further.

Affiliations

[1]*Medical Data Analytics Centre, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China;* [2]*State Key Laboratory of Digestive Disease, Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China;* [3]*Department of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine, Seoul, Republic of Korea;* [4]*Department of Gastroenterology and Hepatology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China;* [5]*University College London Institute for Liver and Digestive Health, Royal Free Hospital and UCL, London, United Kingdom;* [6]*Section of Gastroenterology, PROMISE, University of Palermo, Italy;* [7]*Department of Medical Sciences, Division of Gastroenterology, A.O.U. Città della Salute e della Scienza di Torino, University of Turin, Turin, Italy;* [8]*Department of Gastroenterology and Hepatology, Yokohama City University Graduate School of Medicine, Yokohama, Japan;* [9]*MAFLD Research Center, Department of Hepatology, First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;* [10]*Department of Medicine, Huddinge, Karolinska Institutet, Sweden;* [11]*Division of Hepatology, Department of Upper GI Diseases, Karolinska University Hospital, Huddinge, Stockholm, Sweden;* [12]*Hepato-Gastroenterology and Digestive Oncology Department, Angers University Hospital, Angers, France;* [13]*HIFIH Laboratory, SFR ICAT 4208, Angers University, Angers, France;* [14]*Department of Gastroenterology and Hepatology, Hospital Universitario Puerta de Hierro, Universidad Autonoma de Madrid, Spain;* [15]*Department of Gastroenterology and Hepatology, Singapore General Hospital, Singapore;* [16]*Gastroenterology and Hepatology Unit, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia;* [17]*UCM Digestive Diseases, Virgen del Rocio University Hospital, SeLiver Group, Institute of Biomedicine of Seville, Ciberehd, Department of Medicine, University of Seville, Seville, Spain;* [18]*Stravitz-Sanyal Institute for Liver Disease and Metabolic Health, Virginia Commonwealth University School of Medicine, Richmond, VA, USA;* [19]*Echosens, Paris, France;* [20]*Roger Williams Institute of Liver Studies, School of Immunology & Microbial Sciences, Faculty of Life Sciences and Medicine, King's College London, Foundation for Liver Research and King's College Hospital, London, United Kingdom;* [21]*Department of Gastroenterology, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai Key Lab of Pediatric Gastroenterology and Nutrition, Shanghai, China;* [22]*Université Paris Cité, UMR1149 (CRI), INSERM, Paris, France, Service d'Hépatologie, Hôpital Beaujon, Assistance Publique-Hôpitaux de Paris (AP-HP), Clichy, France;* [23]*Division of Gastroenterology & Hepatology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA;* [24]*Department of Endocrinology, First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;* [25]*National Institute for Health Research, Biomedical Research Centre at University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham, Birmingham, United Kingdom*

Abbreviations

AGA, American Gastroenterological Association; ALT, alanine aminotransferase; AST, aspartate aminotransferase; FIB-4, Fibrosis-4 index; HCC, hepatocellular carcinoma; LRE, liver-related events; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; NPV, negative predictive value; NRI, net reclassification improvement; PPV, positive predictive values; VCTE, vibration-controlled transient elastography.

Financial support

The study was supported in part by the Noncommunicable Chronic Diseases-National Science and Technology Major Project of China (2023ZD0508700), General Research Fund by the Hong Kong SAR Government (14106923), and Direct Grant of The Chinese University of Hong Kong (2024.065).

Conflict of interest

ET served as a consultant for Pfizer, NovoNordisk, Boehringer, and Siemens Healthineers; and a speaker for NovoNordisk, Echosens, and Dr Falk. SP served as a speaker or advisor for AbbVie, Echosens, MSD, Novo Nordisk, Pfizer, and Resalis. EB served as a consultant for Boehringer, MSD, Novo Nordisk, and Pfizer; and a speaker for MSD, Novo Nordisk, and Madrigal. She received research grants from Gilead Sciences. MY received research grant from Gilead Sciences and speaker for KOWA. AN received research grants from Mochida Pharmaceutical, Astellas Pharma, ASKA pharmaceutical, Biofermin Pharmaceutical and EA pharma; a speaker for Mochida Pharmaceutical, Kowa, Biofermin Pharmaceutical, MSD, Boehringer, Novo Nordisk, GlaxoSmithkline, EA pharma. HH served as a consultant for AstraZeneca; and a hepatic events adjudication committee member for KOWA and GW Pharama. His institution has received research funding from AstraZeneca, Echosens, Gilead Sciences, Intercept, MSD, and Pfizer. JB served as a consultant for AstraZeneca, Echosens, Intercept, and Siemens; a speaker for AbbVie, Gilead Sciences, Intercept, and Siemens; and an advisory board member for Bristol-Myers-Squibb, Intercept, Pfizer, MSD, and Novo Nordisk. His institution has received research funding from Diafir, Echosens, Intercept, Inventiva, and Siemens. JLC served as a consultant and speaker for Echosens, Gilead Sciences, and AbbVie. GB-BG served as a consultant for Roche and Ionis Pharmaceuticals; and a speaker for Echosens, Viatris, Abbott and Novo Nordisk. WKC served as a consultant for Abbott, Roche, AbbVie, Novo Nordisk, and Boehringer Ingelheim; and a speaker for Abbott, Novo Nordisk, Echosens, Hisky Medical, and Viatris. AJS served as a consultant for 89Bio, Akero, Allergan, Alnylam Pharmaceuticals, Amgen Inc, AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Genentech, Gilead Sciences, Histoindex, Intercept Pharmaceuticals, Inventiva, Madrigal, Merck, Novartis, Novo Nordisk, Pfizer, Poxel, Salix Pharmaceuticals, Siemens, Sun Pharmaceutical Industries Inc, Terns, and Valeant Pharmaceuticals; and a data safety monitoring board or advisory board member for Bard Peripheral Vascular Inc, NGM Biopharmaceuticals, and Sequana. He has received research funding from Albireo, Allergan, Echosens, Eli Lilly, Gilead Sciences, Intercept Pharmaceuticals, Mallinckrodt LLC, Merck, Novo Nordisk, Perspectum, Pfizer, Salix Pharmaceuticals, and Zydus; and holds the stocks of Durect, Exhalenz, Gen t, and Tiziana. MR-G served as a consultant for Siemens; and a speaker for Siemens and Echosens. He has received research funding from Siemens, Echosens, and Novo Nordisk. PN served as a consultant for Novo Nordisk, Boehringer Ingelheim, Gilead Sciences, Intercept, Poxel Pharmaceuticals, Pfizer, BMS, Eli Lilly, Madrigal, and GSK; and a speaker for Novo Nordisk and AiCME. He has received research funding from Novo Nordisk. LC served as a consultant for Boston pharmaceutical, Echosens, Gilead, GSK, Madrigal, MSD, Novo Nordisk, Pfizer, Sagimet and Siemens Healthineers and as speaker for Echosens, Gilead, Inventiva, Madrigal and Novo Nordisk. CF-P is an employee of Echosens. GL-HW served as a consultant for AstraZeneca, Gilead Sciences, GlaxoSmithKline and Janssen; and a speaker for Abbott, Abbvie, Bristol-Myers Squibb, Echosens, Furui, Gilead Sciences, GlaxoSmithKline and Roche. She has received research funding from Gilead Sciences. MS-WC is an employee of Echosens. VW-SW served as a consultant or advisory board member for AbbVie, Boehringer Ingelheim, Echosens, Gilead Sciences, Intercept, Inventiva, Novo Nordisk, Pfizer, Sagimet Biosciences, TARGET PharmaSolutions, and Visirna; and a speaker for Abbott, AbbVie, Gilead Sciences, Novo Nordisk, and Unilab. He has received a research grant from Gilead Sciences, and is a co-founder of Illuminatio Medical Technology.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

VW-SW designed the study. TC-FY, HL, HWL, ET, SP, EB, MY, M-HZ, HH, JB, JLC, GB-BG, W-KC, RG-D, AJS, VdL, PNN, J-GF, GL-HW, GP, AA, AN, W-YL, YS, MdS-L, EL, KKJT, CL-R, AA, SM, CMC, MR-G, SUK and VW-SW collected data in this study. ET, SP, EB, M-HZ, HH, JB, JLC, GB-BG, W-KC, AJS, VdL, PNN, MR-G, SUK and VW-SW supervised the project. TC-FY, and VW-SW were responsible for data analysis and data interpretation, and drafted the manuscript. TC-FY prepared the figures. All authors provided review and editing of the manuscript, and approved the final version of the manuscript.

Data availability statement

Data are available upon reasonable request to corresponding authors.

Role of the funding source

The funder of the study did not have a role in study design, data collection, data analysis, data interpretation, or manuscript preparation. Echosens provided logistic support in contacting investigators and organising investigator meetings but did not provide funding for this study.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2025.01.014.

References

Author names in bold designate shared co-first authorship

[1] **Wong VW**, **Ekstedt M**, Wong GL, et al. Changing epidemiology, global trends and implications for outcomes of NAFLD. J Hepatol 2023;79:842–852.
[2] Sanyal AJ, Castera L, Wong VW. Noninvasive assessment of liver fibrosis in NAFLD. Clin Gastroenterol Hepatol 2023;21:2026–2039.
[3] Wong VWS, Zelber-Sagi S, Cusi K, et al. Management of NAFLD in primary care settings. Liver Int 2022;42:2377–2389.
[4] Kanwal F, Shubrook JH, Adams LA, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. Gastroenterology 2021;161:1657–1669.
[5] Cusi K, Isaacs S, Barb D, et al. American association of clinical endocrinology clinical practice guideline for the diagnosis and management of nonalcoholic fatty liver disease in primary care and endocrinology clinical settings: Co-sponsored by the American association for the study of liver diseases (AASLD). Endocr Pract 2022;28:528–562.
[6] Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology 2023;77:1797–1835.
[7] European Association for the Study of the Liver, European Association for the Study of Diabetes. European Association for the Study of Obesity. EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). J Hepatol 2024;81:492–542.
[8] European Association for Study of the Liver. Asociacion Latinoamericana para el Estudio del Higado. EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. J Hepatol 2015;63:237–264.
[9] Mozes FE, Lee JA, Vali Y, et al. Performance of non-invasive tests and histology for the prediction of clinical outcomes in patients with non-alcoholic fatty liver disease: an individual participant data meta-analysis. Lancet Gastroenterol Hepatol 2023;8:704–713.
[10] **Lin H**, **Lee HW**, **Yip TC**, et al. Vibration-controlled transient elastography scores to predict liver-related events in steatotic liver disease. JAMA 2024;331:1287–1297.
[11] Kim BK, Bergstrom J, Loomba R, et al. Magnetic resonance elastography-based prediction model for hepatic decompensation in NAFLD: a multi-center cohort study. Hepatology 2023;78:1858–1866.
[12] Sanyal AJ, Foucquier J, Younossi ZM, et al. Enhanced diagnosis of advanced fibrosis and cirrhosis in individuals with NAFLD using FibroScan-based Agile scores. J Hepatol 2023;78:247–259.
[13] Wong VW, Irles M, Wong GL, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. Gut 2019;68:2057–2064.
[14] Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006;43:1317–1325.
[15] Newsome PN, Sasso M, Deeks JJ, 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. Lancet Gastroenterol Hepatol 2020;5:362–373.
[16] Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 2013;32:5381–5397.
[17] Pencina MJ, D'Agostino Sr RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011;30:11–21.
[18] Uno H, Tian L, Cai T, et al. A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. Stat Med 2013;32:2430–2442.
[19] Serra-Burriel M, Juanola A, Serra-Burriel F, et al. Development, validation, and prognostic evaluation of a risk score for long-term liver-related outcomes in the general population: a multicohort study. Lancet 2023;402:988–996.
[20] Mozes FE, Lee JA, Selvaraj EA, et al. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. Gut 2022;71:1006–1019.
[21] Hagstrom H, Talback M, Andreasson A, et al. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. J Hepatol 2020;73:1023–1029.
[22] Cholankeril G, Kramer JR, Chu J, et al. Longitudinal changes in fibrosis markers are associated with risk of cirrhosis and hepatocellular carcinoma in non-alcoholic fatty liver disease. J Hepatol 2023;78:493–500.
[23] McPherson S, Hardy T, Dufour JF, et al. Age as a confounding factor for the accurate non-invasive diagnosis of advanced NAFLD fibrosis. Am J Gastroenterol 2017;112:740–751.
[24] Wong VW, Tak WY, Goh GBB, et al. Performance of noninvasive tests of fibrosis among asians, hispanic, and non-hispanic whites in the STELLAR trials. Clin Gastroenterol Hepatol 2023;21:90–102 e106.
[25] Boursier J, Canivet CM, Costentin C, et al. Impact of type 2 diabetes on the accuracy of noninvasive tests of liver fibrosis with resulting clinical implications. Clin Gastroenterol Hepatol 2023;21:1243–1251 e1212.
[26] Sanyal AJ, Munoz B, Cusi K, et al. Validation of a clinical risk-based classification system in a large nonalcoholic fatty liver disease real-world cohort. Clin Gastroenterol Hepatol 2023;21:2889–2900 e2810.
[27] Chan WK, Treeprasertsuk S, Goh GB, et al. Optimizing use of nonalcoholic fatty liver disease fibrosis score, fibrosis-4 score, and liver stiffness measurement to identify patients with advanced fibrosis. Clin Gastroenterol Hepatol 2019;17:2570–2580 e2537.
[28] Eddowes PJ, Sasso M, Allison M, et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. Gastroenterology 2019;156:1717–1730.
[29] de Franchis R, Bosch J, Garcia-Tsao G, et al. Baveno VII - renewing consensus in portal hypertension. J Hepatol 2022;76:959–974.
[30] Pennisi G, Enea M, Pandolfo A, et al. AGILE 3+ score for the diagnosis of advanced fibrosis and for predicting liver-related events in NAFLD. Clin Gastroenterol Hepatol 2023;21:1293–1302 e1295.
[31] Miura K, Hayashi H, Kamada Y, et al. Agile 3+ and Agile 4, noninvasive tests for liver fibrosis, are excellent formulae to predict liver-related events in nonalcoholic fatty liver disease. Hepatol Res 2023;53:978–988.
[32] Nakatsuka T, Tateishi R, Sato M, et al. Agile scores are a good predictor of liver-related events in patients with NAFLD. J Hepatol 2023;79:e126–e127.
[33] Papatheodoridi M, De Ledinghen V, Lupsor-Platon M, et al. Agile scores in MASLD and ALD: external validation and their utility in clinical algorithms. J Hepatol 2024;81:590–599.
[34] Vali Y, Lee J, Boursier J, et al. Biomarkers for staging fibrosis and non-alcoholic steatohepatitis in non-alcoholic fatty liver disease (the LITMUS project): a comparative diagnostic accuracy study. Lancet Gastroenterol Hepatol 2023;8:714–725.

Keywords: MASLD; FIB-4; VCTE; hepatic decompensation; Hepatocellular carcinoma.
Received 3 September 2024; received in revised form 12 January 2025; accepted 20 January 2025; available online 23 January 2025