

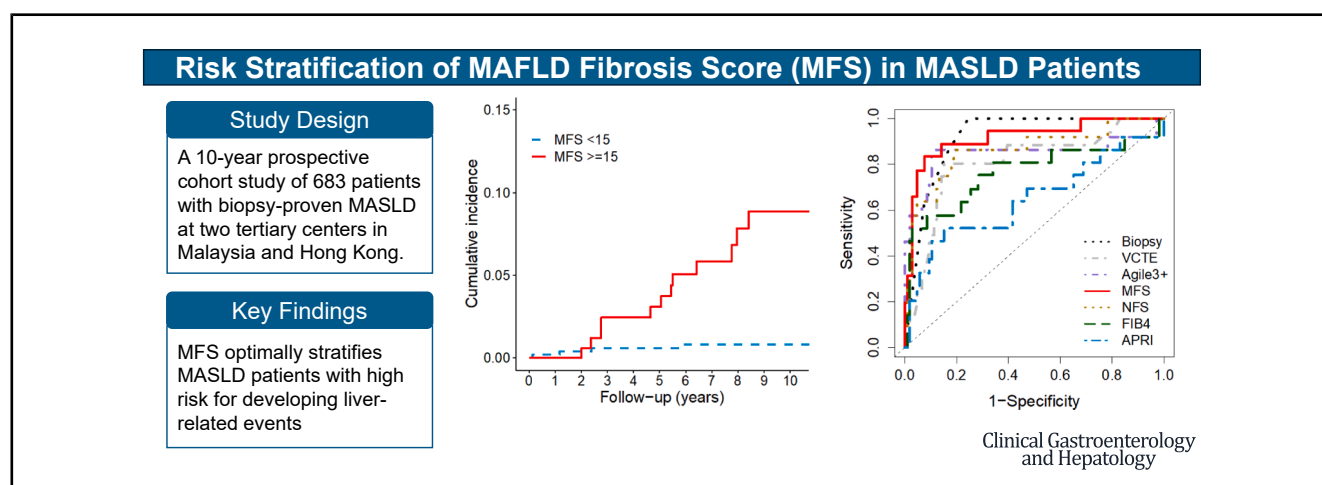
Risk Stratification of a Novel Noninvasive Fibrosis Score in Metabolic Dysfunction-Associated Steatotic Liver Disease



Johnny T. K. Cheung,¹ Xinrong Zhang,² Grace Lai-Hung Wong,^{3,4} Terry Cheuk-Fung Yip,^{3,4} Huapeng Lin,^{3,4} Guanlin Li,^{3,4} Howard Ho-Wai Leung,⁵ Jimmy Che-To Lai,^{3,4} Sanjiv Mahadeva,⁶ Nik Raihan Nik Mustapha,⁷ Xiao-Dong Wang,⁸ Wen-Yue Liu,⁹ Wah-Kheong Chan,⁶ Ming-Hua Zheng,^{8,10} and Vincent Wai-Sun Wong^{3,4}

¹Department of Family Medicine and Primary Care, New Territories East Cluster, Hospital Authority, Hong Kong, China; ²Division of Gastroenterology and Hepatology, School of Medicine, Stanford University, Palo Alto, California; ³Department of Medicine and Therapeutics, Medical Data Analytics Centre, The Chinese University of Hong Kong, Hong Kong, China; ⁴State Key Laboratory of Digestive Disease, Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China; ⁵Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Hong Kong, China; ⁶Gastroenterology and Hepatology Unit, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; ⁷Department of Pathology, Hospital Sultanah Bahiyah, Alor Setar, Kedah, Malaysia; ⁸Key Laboratory of Diagnosis and Treatment for the Development of Chronic Liver Disease in Zhejiang Province, Wenzhou, China; ⁹Department of Endocrinology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China; and ¹⁰Department of Hepatology, MAFLD Research Centre, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

This article has an accompanying educational activity on page e3. Upon completion of this exam, successful learners will be able to: 1) understand the central role of liver fibrosis in determining the prognosis of people with metabolic dysfunction-associated steatotic liver disease; 2) understand how to conduct studies to validate prognostic biomarkers; and 3) learn the prognostic performance of the metabolic dysfunction-associated fatty liver disease fibrosis score.



BACKGROUND & AIMS:

Risk stratification in metabolic dysfunction-associated steatotic liver disease (MASLD) may identify high-risk groups for early intervention to reduce liver-related complications. We examined the prognostic accuracy of the MAFLD fibrosis score (MFS) in predicting risks of liver-related events and mortality among patients with MASLD.

Abbreviations used in this paper: ALP, alkaline phosphatase; ALT, alanine aminotransferase; APRI, aspartate aminotransferase-to-platelet ratio index; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; BMI, body mass index; CI, confidence interval; CIF, cumulative incidence function; FIB-4, Fibrosis-4 index; GGT, gamma-glutamyl transpeptidase; H&E, hematoxylin and eosin; HbA1c, glycated hemoglobin; HCC, hepatocellular carcinoma; HR, hazard ratio; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; INR, international normalized ratio; IQR, interquartile range; LSM, liver stiffness measurement; MAE, mean absolute error; MAFLD, metabolic-associated fatty liver disease; MASH, metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; MFS, MAFLD fibrosis score; MICE, multivariate imputation by chained equations; missForest, Random

Forest algorithm; NAFLD, nonalcoholic fatty liver disease; NFS, NAFLD fibrosis score; NPV, negative predictive value; PPV, positive predictive value; SD, standard deviation; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis; VCTE, vibration-controlled transient elastography.

Most current article

© 2026 American Gastroenterological Association Institute. Published by Elsevier Inc.
1542-3565/\$36.00
<https://doi.org/10.1016/j.cgh.2026.01.035>

METHODS:

This prospective cohort study with a median follow-up period of 9.7 years included 683 patients with biopsy-proven MASLD at 2 tertiary centers in Malaysia and Hong Kong. Noninvasive scores, including MFS, were computed using routine clinical data. Cumulative incidences of any liver-related events and all-cause mortality were compared between patients at low (MFS <15) and high risk (MFS ≥15) with Kaplan-Meier estimation and Cox regression.

RESULTS:

With mean age 52.0 ± 11.8 years and 53.0% male, the high-risk group had an incidence of liver-related events of 844 (95% confidence interval [CI], 436–1475) and all-cause mortality of 1532 (95% CI, 948–2341), compared with 75 (95% CI, 21–193) and 383 (95% CI, 234–592) per 100,000 person-years in the low-risk group, respectively. The hazard ratios were 9.90 (95% CI, 3.19–30.76) for liver-related events and 4.13 (95% CI, 2.22–7.69) for all-cause mortality. MFS outperformed other noninvasive scores, including vibration-controlled transient elastography, in predicting liver-related events and all-cause mortality at 15-year follow-up (area under receiver operating characteristic curve 0.916 and 0.698). A single cutoff of MFS ≥15 demonstrated sensitivity of 77.2% and specificity of 94.3% in predicting liver-related events and 41.4% and 94.8% in predicting mortality.

CONCLUSIONS:

MFS optimally stratifies patients with MASLD with high risk for developing long-term liver-related events and mortality.

Keywords: MAFLD; MASH; NAFLD; Steatohepatitis.

Metabolic dysfunction-associated steatotic liver disease (MASLD), previously known as metabolic-associated fatty liver disease (MAFLD) or nonalcoholic fatty liver disease (NAFLD),¹ is the leading cause of chronic liver disease worldwide.² MASLD encompasses a spectrum of disorders ranging from simple steatosis to metabolic dysfunction-associated steatohepatitis (MASH). Liver fibrosis is a predictor of decompensated cirrhosis and its associated liver-related events such as hepatocellular carcinoma (HCC) and esophageal varices.^{3,4} The financial burden associated with MASLD is substantial, with an estimated annual cost exceeding \$100 billion in the United States and €35 billion in Germany, France, Italy, and the United Kingdom.⁵ Identifying patients at high risk of liver-related events and mortality at an early stage may facilitate timely treatment and prevention.

Given the invasive and labor-intensive nature of liver biopsy, researchers have developed noninvasive tests for fibrosis screening in patients with MASLD. Our group recently developed and validated a new score called the MAFLD fibrosis score (MFS), which incorporates routine clinical and laboratory variables.⁶ In our validation study, MFS showed a significantly higher accuracy in identifying advanced fibrosis (≥F3) compared with other conventional noninvasive scores including aspartate aminotransferase (AST)-to-platelet ratio index (APRI), Fibrosis-4 index (FIB-4), and NAFLD fibrosis score (NFS).⁶ In our cohort with a prevalence of advanced fibrosis 17.1%, a 2-step MFS-vibration-controlled transient elastography (VCTE) screening algorithm demonstrated positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy of 93.4%, 89.5%, and 93.2%, respectively, with only 4.0% of patients classified into the gray zone.⁶

Meanwhile, the prognostic performance of MFS has not been examined. It remains unclear if the risk stratification by MFS can accurately predict the liver-related events in the long term.

The objective of this study is to examine the prognostic accuracy of MFS in stratifying the risks of liver-related events and mortality, compared with other scoring systems.

Methods

Study Design and Participants

This is a multicenter, prospective cohort study using data of patients with MASLD aged 18 and above undergoing liver biopsy at 2 tertiary centers in Malaysia (from 2012 to 2017) and Hong Kong (from 2006 to 2021) with follow-up until the end of 2024. Eligible subjects had: (1) liver biopsy-proved steatosis in at least 5% of hepatocytes; (2) absence of other etiologies including autoimmune liver diseases, hepatitis B or C, secondary fatty liver (the use of systemic steroids or methotrexate, for example), and history of malignancy in the past 2 years; (3) daily alcohol consumption <30 g in men and 20 g in women; and (4) no prior decompensating events.

This study was approved by the respective ethical committee at each center and conformed to the Declaration of Helsinki. Written informed consent was obtained from all patients. This study was reported as per the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline and Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Statement.

Histological Assessment

Ultrasound-guided percutaneous liver biopsy was performed using 16G or 18G needles. Skilled hepatopathologists at each center, who were blinded to clinical and biochemical information, examined formalin-fixed paraffin-embedded biopsy tissue stained with hematoxylin and eosin (H&E) and Masson's trichrome. The diagnosis of MASLD was established if steatosis was present in $\geq 5\%$ of hepatocytes. Liver fibrosis was classified into 5 stages: absence of fibrosis (F0); perisinusoidal or portal fibrosis (F1); perisinusoidal and portal/periportal fibrosis (F2); septal or bridging fibrosis (F3); and cirrhosis (F4).⁷ A F3 or F4 indicated advanced fibrosis.

Clinical and Laboratory Parameters

We measured baseline anthropometric parameters including body mass index (BMI) and blood pressure at enrollment. We took venous blood samples 1 day before liver biopsy after overnight fasting for at least 8 hours. Blood tests related to liver function including platelet count, alanine aminotransferase (ALT), AST, albumin, alkaline phosphatase (ALP), gamma-glutamyl transpeptidase (GGT), and international normalized ratio (INR) were ordered. Lipid and glucose profiles including total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total triglycerides, fasting blood glucose, and glycated hemoglobin (HbA1c), as well as renal function (creatinine and urea), were measured.

We obtained the history of hypertension and type 2 diabetes from clinical assessment. Hypertension was defined as blood pressure $\geq 140/90$ mm Hg or the use of anti-hypertensives. Diabetes was diagnosed by fasting plasma glucose ≥ 7 mmol/L (126 mg/dL), HbA1c $\geq 6.5\%$ (48 mmol/mol), or use of anti-diabetic drugs. Impaired fasting glycemia was defined as fasting plasma glucose level from 5.6 mmol/L (101 mg/dL) to 6.9 mmol/L (124 mg/dL). Background of cardiovascular and cerebrovascular events were obtained from self-reported medical history.

The outcomes of interest included liver-related events and all-cause mortality. Liver-related events were defined as a composite endpoint of HCC and cirrhotic complications (ascites, spontaneous bacterial peritonitis, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome) based on International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes and procedure codes (Supplementary Table 1). The accuracy of ICD-9 codes in identifying liver-related complications was assessed in a Hong Kong-based cohort using a random sample of 125 patients (28 with liver-related complications), demonstrating a PPV of 85.7% and a NPV of 95.9%.⁸ Mortality was identified via access to death registry.

What You Need to Know

Background

Risk stratification in metabolic dysfunction-associated steatotic liver disease (MASLD) may help identify high-risk patients for early intervention and reduce complications. MAFLD fibrosis score (MFS) is one of the noninvasive scores computed using routine clinical data.

Findings

The MFS ≥ 15 effectively distinguishes patients with MASLD at high risk for liver-related events and mortality in Hong Kong and Malaysian cohorts, possibly outperforming other noninvasive scores.

Implications for patient care

MFS can be used to identify patients with MASLD who may benefit from closer monitoring and early interventions to reduce adverse events.

FibroScan

Experienced operators utilized VCTE (FibroScan, Echosens) to obtain liver stiffness measurements (LSMs) to assess the fibrosis stage. The LSM value was determined by calculating the median of 10 successful measurements. To ensure reliability, a minimum of 10 successful acquisitions were required, with an interquartile range (IQR)-to-median ratio of ≤ 0.3 . In the Hong Kong cohort, patients with a BMI < 30 kg/m² were assessed using the M probe, whereas those with a BMI ≥ 30 kg/m² were assessed using the XL probe. A simple LSM cutoff of < 10 kPa was used to exclude advanced fibrosis, whereas a cutoff of ≥ 15 kPa was used to diagnose advanced fibrosis for both M and XL probes in keeping with the Baveno consensus on the assessment of compensated advanced chronic liver disease.⁹

MFS and Other Noninvasive Scores

In our prior validation study on MFS, we employed a machine learning R package, namely Boruta, to select variables with significant predictive value for advanced fibrosis ($\geq F3$).⁶ This package utilized the random forest classification algorithm to identify variables deemed of 'high importance,' which were subsequently used in a backward stepwise regression to develop an equation for calculating MFS in the training cohort. This equation enables the prediction of advanced fibrosis and is as follows: $MFS = 0.078 \times \text{age (year)} - 0.007363 \times \text{platelet count (10}^9\text{/L)} + 0.0146 \times \text{AST (U/L)} + 0.007618 \times \text{GGT (U/L)} + 6.673 \times \text{INR} + 0.09833 \times \text{BMI (kg/m}^2\text{)} + 1.425 \times \text{type 2 diabetes (yes = 1, no = 0)}$.⁶ In our validation cohort, MFS demonstrated an area under the receiver operating characteristic curve (AUROC) of 0.823 for correctly identifying the presence

or absence of liver fibrosis.⁶ MFS demonstrated a correct prediction rate of 90.4%, resulting in a larger number of patients with correct diagnosis compared with APRI, FIB-4, and NFS.⁶ We computed NFS, FIB-4, APRI, Agile 3+, and VCTE LSM and adopted cutoffs proposed from the previous studies to identify high- and low-risk groups.^{9–13} Agile 3+ was a novel score consisting of both LSM and routinely measured parameters.¹⁴

Data Analysis

We present descriptive statistics of the sociodemographic, clinical, and laboratory profiles of the 2 cohorts as mean $\pm$ standard deviation (SD) for continuous variables or number (percentages) for categorical variables. The Kaplan-Meier method was used to estimate the cumulative incidence of liver-related events and all-cause mortality, with log-rank test comparing the incidence of the positive events across patients with MFS <15 vs ≥ 15 . Hazard ratios (HRs) and their 95% confidence intervals (95% CIs) were computed. In addition, we used Fine-Gray models to estimate subdistribution HRs, accounting for the competing risk of liver-related events and mortality. Cumulative incidence function (CIF) curves were plotted to visualize the absolute risk of events over time. The performance of the noninvasive scores, VCTE, and biopsy-proven fibrosis stage in predicting the events at 5, 10, and 15 years of follow-up was measured and compared by the time-dependent AUROC. Furthermore, we calculated the sensitivity, specificity, PPV, and NPV using predefined cutoffs for risk stratification. We compared the number of patients with MASLD and correct identification of absence or presence of liver-related events and mortality in each scoring system.

For handling missing data, we adopted the Random Forest algorithm (missForest), which was regarded as the best imputation technique for mixed parametric and nonparametric data as well as outlier handling.¹⁵ Among 4 imputation methods (missForest, mean imputation, nearest neighbor imputation, and multivariate imputation by chained equations [MICE]), missForest has demonstrated the least imputation error for both continuous and categorical variables at each frequency of missingness, along with the smallest prediction difference when models used imputed laboratory values.

To examine the robustness of the study findings, we performed sensitivity analyses by repeating all the statistical tests in the patients retained after listwise deletion. This technique excluded patients with any missing data of variables necessary for calculating any noninvasive scores. A total of 9 and 281 patients from the Malaysia and Hong Kong cohorts were excluded in the sensitivity analyses, respectively. The substantial amount of missing data in the Hong Kong cohort arose due to the non-routine inclusion of AST as a laboratory parameter. [Supplementary Table 2](#) displayed the

profiles of patients selected for sensitivity analyses. A 2-tailed P -value $<.05$ indicated statistical significance. All analyses were performed using R version 4.0.2.

Results

Patient Characteristics

[Table 1](#) describes the profile of 683 patients with biopsy-proven MASLD recruited from Malaysia and Hong Kong. The patients underwent liver biopsy, VCTE, and clinical and laboratory assessment within 1 week. The median follow-up period was 9.7 years (IQR, 6.6–11.4 years). No major adverse events were reported from the noninvasive tests or liver biopsy. The overall mean age was 52.0 ± 11.8 years, with approximately one-half being male (53.0%). About 31.6% had advanced fibrosis (fibrosis stage F3–4). During the follow-up period, 2.3% developed at least one liver-related event, and 6.0% died. Compared with the Hong Kong cohort, the Malaysian cohort was older, had a higher BMI, fasting glucose and HbA_{1c}, and proportion of ischemic heart disease. Meanwhile, the Hong Kong cohort tended to have a more advanced liver fibrosis stage.

Prognostic Performance of MFS

[Figure 1](#) shows the cumulative incidence of liver-related events (panel A) and all-cause mortality (panel B). The Kaplan-Meier analyses demonstrated significantly higher incidences of these outcomes in patients with MFS ≥ 15 (high-risk group) compared with their counterparts with MFS <15 (low-risk group). The incidence rate of liver-related events was 844 (95% CI, 436–1475) in the high-risk group, compared with 75 (95% CI, 21–193) per 100,000 person-years in the low-risk group. The incidence rate of all-cause mortality was 1532 (95% CI, 948–2341) in the high-risk group, compared with 383 (95% CI, 234–592) per 100,000 person-years in the low-risk group. Cox regression analysis showed that the high-risk group had significantly higher risks for liver-related events (HR, 9.90; 95% CI, 3.19–30.76) and mortality (HR, 4.13; 95% CI, 2.22–7.69) than the low-risk group. Fine-Gray subdistribution hazard models demonstrated similar HRs for liver-related events (HR, 9.88; 95% CI, 3.14–31.1) and mortality (HR, 4.04; 95% CI, 2.20–7.41), with the CIF curve plotted in [Supplementary Figure 1](#). The high-risk group contributed to an additional 81 liver-related events per 1000 patients (95% CI, 48–113 per 1000) and 116 additional mortality cases per 1000 patients (95% CI, 40–193 per 1000) compared with the low-risk group.

[Figure 2](#) and [Supplementary Table 2](#) shows the performance of MFS against other assessment tools, including VCTE and histological stage, in predicting

Table 1. Characteristics of Patients With Biopsy-Proven MASLD in Hong Kong and Malaysia Cohorts

	Overall (N = 683)	Hong Kong (n = 481)	Malaysia (n = 202)	P
Age, y	52.0 ± 11.8	51.3 ± 12.1	53.7 ± 11.0	.014
Male sex	362 (53.0)	266 (55.3)	96 (47.5)	.076
BMI, kg/m ²	29.1 ± 4.8	28.7 ± 4.8	30.2 ± 4.7	<.001
Systolic blood pressure, mm Hg	137.3 ± 17.7	136.6 ± 17.8	139.0 ± 17.2	.099
Diastolic blood pressure, mm Hg	80.7 ± 11.1	80.2 ± 11.3	81.9 ± 10.7	.061
Hypertension	409 (59.9)	279 (58.0)	130 (64.4)	.144
Type 2 diabetes	428 (62.7)	291 (60.5)	137 (67.8)	.086
History of ischemic heart disease	45 (6.6)	24 (5.0)	21 (10.4)	.015
History of cerebrovascular event	15 (2.2)	11 (2.3)	4 (2.0)	1.000
Platelet count, ×10 ⁹ /L	245.0 ± 70.2	235.5 ± 69.2	267.8 ± 67.3	<.001
INR	1.0 ± 0.1	1.0 ± 0.1	1.0 ± 0.1	<.001
ALT, U/L	72.4 ± 88.0	75.3 ± 100.9	65.6 ± 43.6	.188
ALP, U/L	81.1 ± 34.0	81.8 ± 36.0	79.5 ± 28.8	.429
GGT, U/L	92.3 ± 108.4	91.0 ± 113.8	95.4 ± 94.6	.625
AST, U/L	43.6 ± 32.9	43.0 ± 35.0	44.8 ± 27.2	.527
Albumin, g/L	42.9 ± 3.8	43.6 ± 3.6	41.2 ± 3.6	<.001
Creatinine, μmol/L	76.1 ± 22.3	76.2 ± 22.0	75.8 ± 22.9	.809
Urea, mmol/L	5.0 ± 1.8	5.2 ± 1.9	4.7 ± 1.7	<.001
Triglycerides, mmol/L	1.8 ± 1.0	1.8 ± 1.1	1.8 ± 0.9	.539
Total cholesterol, mmol/L	4.8 ± 1.1	4.8 ± 1.1	4.7 ± 1.1	.115
LDL-cholesterol, mmol/L	2.7 ± 0.9	2.7 ± 0.9	2.7 ± 0.9	.605
HDL-cholesterol, mmol/L	1.4 ± 2.8	1.4 ± 0.7	1.5 ± 5.0	.582
Fasting glucose, mmol/L	6.8 ± 2.4	6.7 ± 2.1	7.2 ± 2.9	.009
HbA1c, %	6.8 ± 1.4	6.6 ± 1.2	7.2 ± 1.6	<.001
MFS	13.9 ± 1.9	13.7 ± 1.9	14.2 ± 1.8	.008
NFS	−1.4 ± 1.7	−1.4 ± 1.7	−1.4 ± 1.4	.680
APRI	0.5 ± 0.4	0.5 ± 0.4	0.5 ± 0.3	.143
FIB4	1.3 ± 1.2	1.3 ± 1.3	1.2 ± 0.8	.395
VCTE LSM, kPa	10.3 ± 6.9	10.2 ± 7.2	10.5 ± 6.0	.575
Fibrosis stage ^a				<.001
0	151 (22.1)	100 (20.8)	51 (25.2)	
1	236 (34.6)	154 (32.0)	82 (40.6)	
2	80 (11.7)	66 (13.8)	14 (6.9)	
3	131 (19.2)	81 (16.8)	50 (24.8)	
4	85 (12.4)	80 (16.6)	5 (2.5)	
Liver-related events ^b	16 (2.3)	11 (2.3)	5 (2.5)	1.000
Ascites	6 (0.8)	5 (1.0)	1 (0.5)	
Variceal bleeding	5 (0.7)	2 (0.4)	3 (1.5)	
Hepatic encephalopathy	3 (0.4)	3 (0.6)	0 (0.0)	
Hepatorenal syndrome	1 (0.1)	1 (0.2)	0 (0.0)	
Spontaneous bacterial peritonitis	0 (0.0)	0 (0.0)	0 (0.0)	
Hepatocellular carcinoma	5 (0.7)	4 (0.8)	1 (0.5)	1.000
All-cause mortality	41 (6.0)	34 (7.1)	7 (3.5)	.103

NOTE. Data are presented as number (%) or mean ± standard deviation.

ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; GGT, gamma-glutamyl transpeptidase; HbA1c, Hemoglobin A1c; HDL, high-density lipoprotein; INR, international normalized ratio; LDL, low-density lipoprotein; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; VCTE, vibration-controlled transient elastography (FibroScan).

^aFibrosis stage was classified by the histological scoring systems validated by the Pathology Committee of the NASH Clinical Research Network.

^bLiver-related events include ascites, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, and hepatocellular carcinoma. The total count of liver-related events did not equal to the sum of the individual events because they were not mutually exclusive.

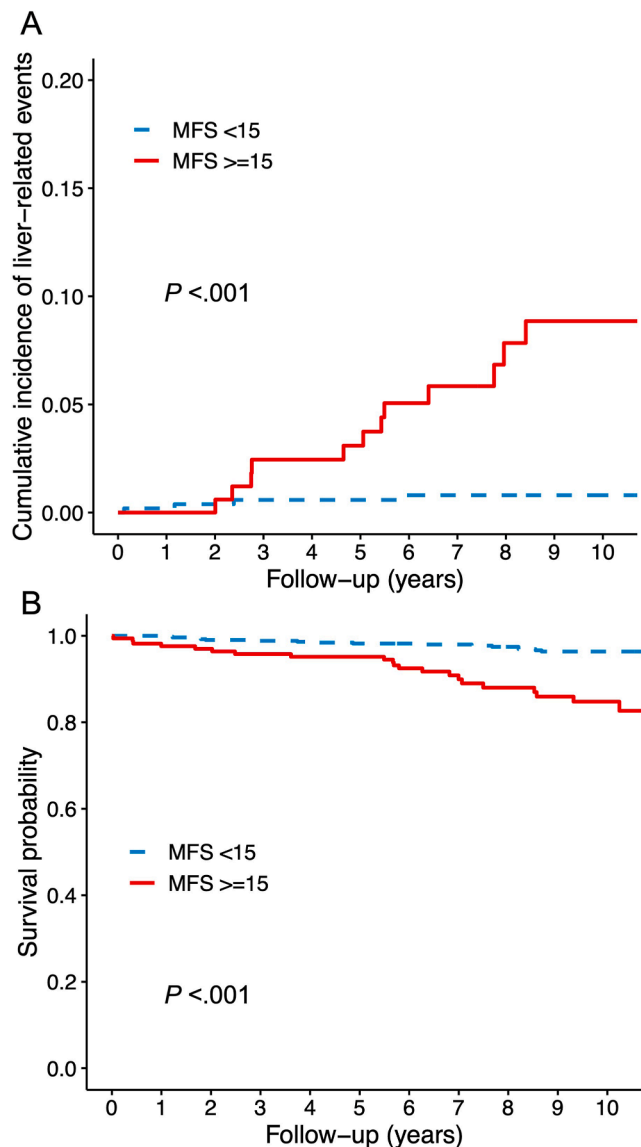


Figure 1. Kaplan-Meier curves for liver-related events (A) and all-cause mortality (B) in patients with MASLD stratified by MFS. MFS <15 and ≥15 indicate low and high likelihood of advanced liver fibrosis, respectively. Log-ranked *P* values are presented within the figure panels. Liver-related events include ascites, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, and hepatocellular carcinoma.

liver-related events and mortality. MFS outperformed FIB-4 and APRI in predicting 15-year liver-related events (AUROC of 0.916 for MFS vs 0.765 for FIB-4 and 0.648 for APRI). Similarly, MFS demonstrated a higher accuracy in predicting all-cause mortality than APRI (AUROC 0.698 vs 0.535), but no statistical significance was achieved in comparison with NFS and FIB-4. The accuracy of MFS in predicting both outcomes at 15-year follow-up was comparable to those of biopsy, VCTE, and Agile 3+.

Table 2 presents the prognostic accuracy of MFS using 15 as a single cutoff, in comparison to NFS, FIB-4, APRI, VCTE, and Agile 3+ using the predefined dual

cutoffs and liver biopsy using F3 fibrosis stage as cutoff. For liver-related events, MFS ≥15 demonstrated a sensitivity of 77.2% and a specificity of 94.3%. For all-cause mortality, MFS ≥15 demonstrated a sensitivity of 41.4% and a specificity of 94.8%.

Table 3 illustrates the prognostic performance of the assessment tools in our cohort. A single cutoff of MFS ≥15 could correctly predict 76.9% of liver-related events and 75.8% of all-cause mortality in the overall cohort, compared with 70.1% and 67.3% by histological stage ≥3. Other tools using dual cutoff could include 56.4% to 81.0% of subjects, with correct prediction rate of 48.9% to 66.2% and 47.3% to 64.9% for any liver-related events and mortality, respectively, in the overall cohort.

Nomograms mapped total points to predicted probabilities of liver-related events (Supplementary Figure 2A) and mortality (Supplementary Figure 3A). The apparent and bias-corrected lines (adjusted for overfitting via 1000-repetition bootstrap validation) followed the ideal line. Calibration was excellent overall, with mean absolute errors (MAEs) of 0.006 (liver-related events) (Supplementary Figure 2B) and 0.007 (mortality) (Supplementary Figure 3B). Decision curve analysis demonstrated net benefit of MFS over default strategies for threshold probabilities of 2% to 13% (liver-related events) (Supplementary Figure 2C) and 4% to 15% (mortality) (Supplementary Figure 3C).

Sensitivity Analysis

The sensitivity analyses, which involved repeating the statistical tests on subjects retained after listwise deletion, yielded consistent results. The observed differences in the clinical profiles between the Hong Kong and Malaysia cohorts aligned with those found in the main analysis (Supplementary Table 3). Similarly, the MFS-stratified high-risk group exhibited the greatest risk for developing liver-related events (HR, 10.38; 95% CI, 2.87–37.53) and mortality (HR, 4.45; 95% CI, 1.67–11.88) (Supplementary Figure 4). MFS surpassed FIB-4 and APRI in prediction in terms of AUROCs (Supplementary Figure 5 and Supplementary Table 4). The prognostic accuracy and performance are presented in Supplementary Tables 5 and 6.

Discussion

In this multicenter prospective cohort study, with a median follow-up of 10 years, MFS effectively stratified patients into low- or high-risk groups for developing liver-related events and mortality. MFS outperformed other noninvasive scores in accurately predicting the presence or absence of the adverse events. Furthermore, the prognostic accuracy of MFS increased with longer follow-up durations.

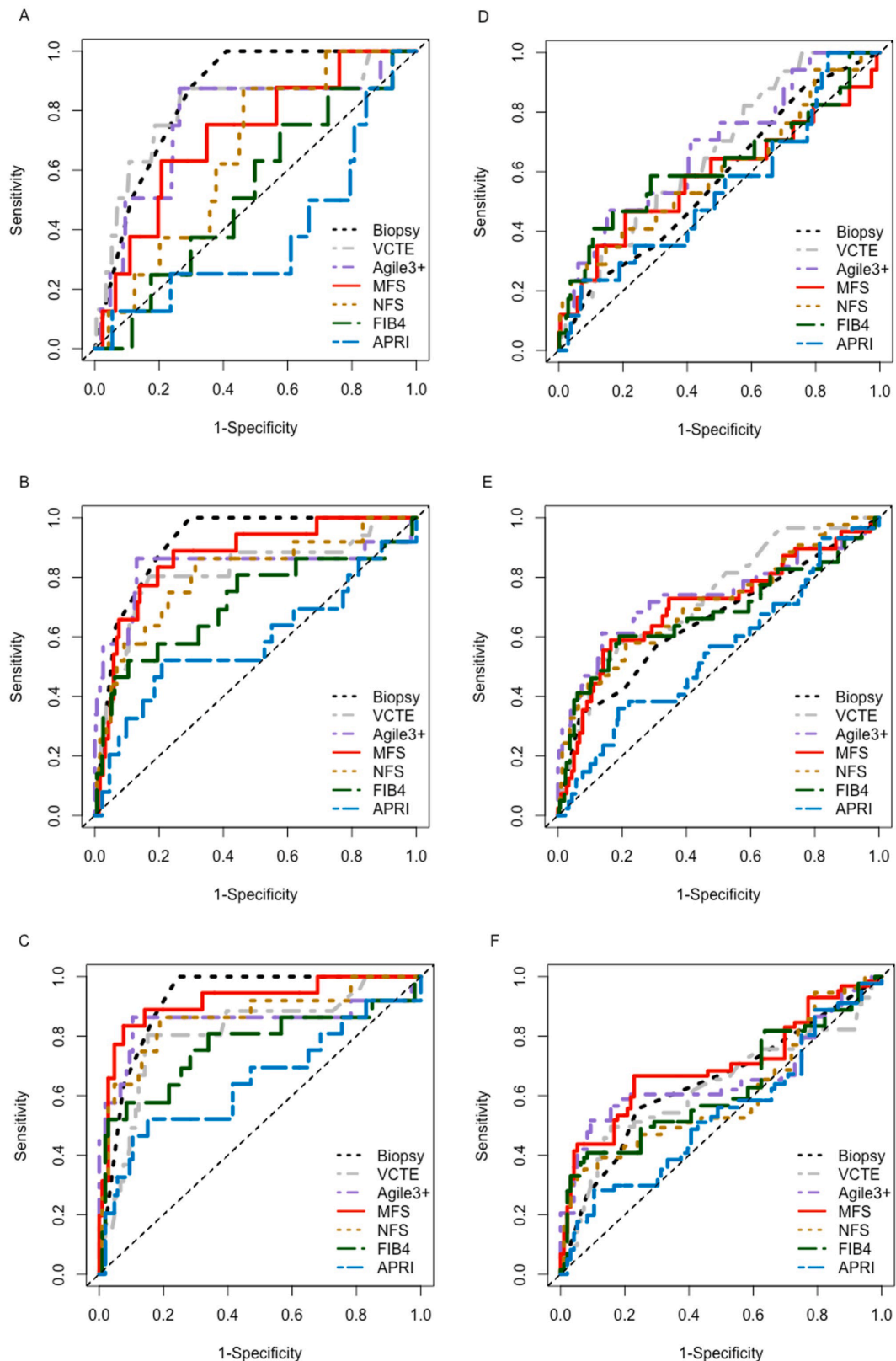


Figure 2. Receiver-operating characteristic curves of biopsy, VCTE, Agile3+, MFS, NFS, FIB-4, and APRI for predicting liver-related events at 5 years (A), 10 years (B), and 15 years (C), and all-cause mortality at 5 years (D), 10 years (E), and 15 years (F). Liver-related events include ascites, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, and hepatocellular carcinoma.

Table 2. Accuracy of Assessment Tools in Predicting Risk for Liver-Related Events and All-Cause Mortality in Patients With MASLD

	Cutoff	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Liver-related events					
MFS	≥ 15	77.2	94.3	27.2	99.3
NFS	< -1.455	86.3	78.3	9.8	99.5
FIB-4	≥ 0.675	25.6	98.1	27.1	98.0
	≤ 1.3	57.6	90.6	14.3	98.7
APRI	≥ 2.67	32.9	98.1	32.3	98.2
	≤ 0.5	52.1	79.3	6.4	98.4
Agile 3+	> 1.5	0.0	98.1	0.0	97.3
	≤ 0.251	86.3	76.4	9.1	99.5
VCTE LSM	> 0.565	86.3	88.6	17.3	99.6
	< 10 kPa	80.4	81.1	10.4	99.3
Fibrosis stage	≥ 15 kPa	49.1	88.7	10.6	98.5
	≥ 3	86.3	83.0	12.2	99.6
All-cause mortality					
MFS	≥ 15	41.4	94.8	44.9	94.0
NFS	< -1.455	43.0	77.1	16.2	93.0
FIB-4	≥ 0.675	16.5	97.9	44.8	92.0
	≤ 1.3	40.8	89.6	28.7	93.7
APRI	≥ 2.67	14.1	97.9	41.0	91.8
	≤ 0.5	29.8	77.1	11.8	91.5
Agile 3+	> 1.5	0.0	97.9	0.0	90.5
	≤ 0.251	60.4	76.0	20.5	94.9
VCTE LSM	> 0.565	51.6	89.6	33.7	94.8
	< 10 kPa	49.5	82.3	22.3	94.1
Fibrosis stage	≥ 15 kPa	35.0	88.5	23.9	93.0
	≥ 3	38.8	82.3	18.4	92.9

NOTE. Liver-related events include ascites, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, and hepatocellular carcinoma.

NOTE. A biopsy-proven fibrosis stage 3 or above is regarded as advanced liver fibrosis.

APRI, aspartate aminotransferase-to-platelet ratio index; FIB-4, Fibrosis-4; LSM, liver stiffness measurement; MFS, MAFLD fibrosis score; NFS, non-alcoholic fatty liver disease (NAFLD) fibrosis score; NPV, negative predictive value; PPV, positive predictive value; VCTE, vibration controlled transient elastography.

A systematic review of 13 studies conducted in secondary or tertiary care settings evaluated the prognostic capabilities of noninvasive markers.¹⁶ The review found that these markers had satisfactory AUROC values for predicting liver-related events and mortality by FIB-4 (AUROCs of 0.71–0.89 and 0.67–0.82), NFS (AUROCs of 0.77–0.92 and 0.70–0.83), and APRI (AUROCs of 0.69–0.89 and 0.54–0.73).¹⁶ On the other hand, the review suggested the need for more refined models to improve the prognostic performance and clinical utility of these markers. In our study, MFS demonstrated an AUROC of 0.92 and 0.70 in predicting 15-year liver-related events and mortality, respectively, surpassing FIB-4, NFS, and APRI.

Considering the high prevalence of MASLD, the use of operator-dependent imaging including VCTE or ultrasound as the initial screening for fibrosis may be costly and impractical.¹⁷ There is growing promise in risk stratification using noninvasive markers for identifying patients more likely to develop severe liver events.

Therefore, the European Association for the Study of the Liver,¹⁸ the American Association for the Study of Liver Diseases,¹⁹ and the Asian Pacific Association for the Study of the Liver²⁰ recommend the use of simple noninvasive panels as the first step in the screening among patients with MASLD. Data shown in our previous report support the screening for advanced liver disease starting in patients aged 50 years or above with MASLD and type 2 diabetes.²¹ Additionally, these scores are recommended as prognostic markers to rule out disease progression.²² However, APRI and FIB-4 were originally developed and validated among patients with chronic hepatitis C to diagnose advanced liver fibrosis. Although their accuracy has been demonstrated in patients with MASLD, our study shows that there is still room for improvement in these conventional scores. In contrast, MFS, developed through machine learning, represents a novel score that exhibits improved diagnostic and prognostic value compared with conventional scores.

Table 3. Performance of Assessment Tools in Predicting Risk for Liver-Related Events and All-Cause Mortality in the Cohort Patients

	Within the range ^a	Within-range correct prediction ^b	Overall correct prediction ^c
Liver-related events			
MFS <15 or ≥15	683 (100.0)	525 (76.9)	525 (76.9)
NFS <-1.455 or >0.675	385 (56.4)	334 (86.8)	334 (48.9)
FIB-4 ≤1.3 or ≥2.67	497 (72.8)	452 (90.9)	452 (66.2)
APRI ≤0.5 or >1.5	477 (69.8)	452 (94.8)	452 (66.2)
Agile 3+ <0.251 or ≥0.565	548 (80.2)	444 (81.0)	444 (65.0)
VCTE LSM <10 or ≥15 kPa	553 (81.0)	444 (80.3)	444 (65.0)
Fibrosis stage <3 or ≥3	683 (100.0)	479 (70.1)	479 (70.1)
All-cause mortality			
MFS <15 or ≥15	683 (100.0)	518 (75.8)	518 (75.8)
NFS <-1.455 or >0.675	385 (56.4)	323 (83.9)	323 (47.3)
FIB-4 ≤1.3 or ≥2.67	497 (72.8)	443 (89.1)	443 (64.9)
APRI ≤0.5 or >1.5	477 (69.8)	434 (91.0)	434 (63.5)
Agile 3+ <0.251 or ≥0.565	548 (80.2)	338 (61.7)	338 (49.5)
VCTE LSM <10 or ≥15 kPa	553 (81.0)	436 (78.8)	436 (63.8)
Fibrosis stage <3 or ≥3	683 (100.0)	460 (67.3)	460 (67.3)

NOTE. Liver-related events include ascites, variceal bleeding, hepatic encephalopathy, hepatorenal syndrome, spontaneous bacterial peritonitis, and hepatocellular carcinoma.

NOTE. A biopsy-proven fibrosis stage 3 or above is regarded as advanced liver fibrosis.

APRI, aspartate aminotransferase-to-platelet ratio index; FIB-4, Fibrosis-4; LSM, liver stiffness measurement; MFS, MAFLD fibrosis score; NFS, non-alcoholic fatty liver disease (NAFLD) fibrosis score; VCTE, vibration controlled transient elastography.

^aThe number and proportion of patients included within the range specified by corresponding scores.

^bThe number and proportion of patients with correct identification of the presence or absence of events among those included within the range specified by corresponding scores.

^cThe number and proportion of patients with correct identification of the presence or absence of events among the overall cohort.

Recent clinical trials have explored the safety and efficacy profile of pharmacological treatments for MASH. Dual glucagon and glucagon-like peptide-1 receptor agonists, exemplified by survodutide, has shown improvements in MASH and fibrosis.²³ The MAESTRO-NASH trial demonstrated that resmetirom, an oral, liver-directed, thyroid hormone receptor beta-selective agonist, may lead to MASH resolution and liver fibrosis improvement by at least one stage.²⁴ Given the large population of patients with MASLD and limited resources, noninvasive markers could be utilized to identify high-risk patients for MASH, enabling targeted treatment in clinical practice where liver biopsy is rarely feasible.²⁴ MFS is one such prognostic marker that can assist clinicians in identifying patients at high risk of liver decompensation and its associated clinical manifestations.

The present study has several limitations. First, patients requiring liver biopsy at tertiary care were expected to have a higher prevalence of liver fibrosis than patients in a primary care setting,²⁵ limiting the generalizability of the findings. Nonetheless, liver biopsy is the gold standard for MASLD and fibrosis diagnosis, ensuring the robustness and validity of our study findings. It is often unethical or impractical to perform liver biopsy in low-risk patients without sound justification. Further studies are needed to assess the generalizability of the MFS in primary care settings, non-Asian populations, and larger cohorts. Second, our cohort data did not record the change of the clinical and laboratory

parameters that construct MFS and other fibrosis scores. Our analyses could not capture the decrease in liver stiffness and serum biomarker levels, which reflects the regression of liver fibrosis and inflammation.²⁶ In addition, we recognize the possibility of underreported alcohol consumption, as self-reported data may be subject to recall or social desirability bias. Phosphatidylethanol was not measured given its relative novelty. At the same time, all patients were prospectively recruited, and excessive alcohol use was systematically excluded through structured history-taking and validated questionnaires.

In conclusion, MFS optimally stratifies patients with MASLD into low- and high-risk groups for liver-related events and mortality over the long term. Further validation and prognostic studies of MFS should be conducted in primary and secondary care settings.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at <https://doi.org/10.1016/j.cgh.2026.01.035>.

References

1. Rinella ME, Lazarus JV, Ratzliff V, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *J Hepatol* 2023;79:1542–1556.

2. Zhang X, Goh GBB, Chan WK, et al. Unhealthy lifestyle habits and physical inactivity among Asian patients with non-alcoholic fatty liver disease. *Liver Int* 2020;40:2719–2731.
3. Powell EE, Wong VWS, Rinella M. Non-alcoholic fatty liver disease. *Lancet* 2021;397:2212–2224.
4. Zhang X, Wong GLH, Wong VWS. Application of transient elastography in nonalcoholic fatty liver disease. *Clin Mol Hepatol* 2020;26:128–141.
5. Younossi ZM, Blissett D, Blissett R, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. *Hepatology* 2016;64:1577–1586.
6. Cheung JTK, Zhang X, Wong GLH, et al. MAFLD fibrosis score: using routine measures to identify advanced fibrosis in metabolic-associated fatty liver disease. *Aliment Pharmacol Ther* 2023;58:1194–1204.
7. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999;94:2467–2474.
8. Yip T, Li G, Wong G, et al. Accuracy of administrative coding for nonalcoholic fatty liver disease and related hepatic complications: a single-center validation study. *Hepatology* 2021;74:928A–929A.
9. Wong VWS, Irls M, Wong GLH, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. *Gut* 2019;68:2057–2064.
10. 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.
11. 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.
12. Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology* 2007;45:846–854.
13. Wai CT, Greenon JK, Fontana RJ, et al. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003;38:518–526.
14. 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.
15. Stekhoven DJ, Bühlmann P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012;28:112–118.
16. Lee J, Vali Y, Boursier J, et al. Prognostic accuracy of FIB-4, NAFLD fibrosis score and APRI for NAFLD-related events: a systematic review. *Liver Int* 2021;41:261–270.
17. Lin H, Lee HW, Yip TCF, et al. Vibration-controlled transient elastography scores to predict liver-related events in steatotic liver disease. *JAMA* 2024;331:1287–1297.
18. European Association for the Study of the Liver (EASL), European Association for the Study of diabetes (EASD), European Association for the Study of obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). *J Hepatol* 2024;81:492–542.
19. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 2018;67:328–357.
20. Eslam M, Sarin SK, Wong VWS, 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. *Hepatol Int* 2020;14:889–919.
21. Zhang X, Wong GL, Yip TC, et al. Risk of liver-related events by age and diabetes duration in patients with diabetes and nonalcoholic fatty liver disease. *Hepatology* 2022;76:1409–1422.
22. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol* 2016;64:1388–1402.
23. SA J, Pierre B, Mandy F, NG W, Eric L, Elisabetta B. A phase 2 randomized trial of survodutide in MASH and fibrosis. *N Engl J Med* 2024;391:311–319.
24. Harrison SA, Bedossa P, Guy CD, et al. A phase 3, randomized, controlled trial of resmetrom in NASH with liver fibrosis. *N Engl J Med* 2024;390:497–509.
25. Wong VWS, Zelber-Sagi S, Cusi K, et al. Management of NAFLD in primary care settings. *Liver Int* 2022;42:2377–2389.
26. Sun YM, Chen SY, You H. Regression of liver fibrosis: evidence and challenges. *Chin Med J (Engl)* 2020;133:1696–1702.

Correspondence

Address correspondence to: Ming-Hua Zheng, MD, PhD, Department of Hepatology, MAFLD Research Centre, The First Affiliated Hospital of Wenzhou Medical University, No. 2 Fuxue Lane, Wenzhou 325000, China. e-mail: zhengmh@wmu.edu.cn; or Wah-Kheong Chan, MD, PhD, Gastroenterology and Hepatology Unit, Department of Medicine, Faculty of Medicine, University of Malaya, Lembah Pantai, Kuala Lumpur 50603, Malaysia. e-mail: wahkheong2003@hotmail.com; or 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.

CRedit Authorship Contributions

Johnny T.K. Cheung (Study design; Data analysis; Writing the paper)
Xinrong Zhang (Data interpretation; Critical revision of the paper)
Grace Lai-Hung Wong (Data interpretation; Critical revision of the paper)
Terry Cheuk-Fung Yip (Data interpretation; Critical revision of the paper)
Huapeng Lin (Data interpretation; Critical revision of the paper)
Guanlin Li (Data interpretation; Critical revision of the paper)
Jimmy Che-To Lai (Data interpretation; Critical revision of the paper)
Howard Ho-Wai Leung (Data collection; Critical revision of the paper)
Sanjiv Mahadeva (Data collection; Critical revision of the paper)
Nik Raihan Nik Mustapha (Data collection; Critical revision of the paper)
Xiao-Dong Wang (Data collection; Critical revision of the paper)
Wen-Yue Liu (Data collection; Critical revision of the paper)
Vincent Wai-Sun Wong (Study conceptualization; Critical revision of the paper)
Wah-Kheong Chan (Study conceptualization; Critical revision of the paper)
Ming-Hua Zheng (Study conceptualization; Critical revision of the paper)

Conflicts of interest

These authors disclose the following: Grace Lai-Hung Wong has served as a speaker for Echosens. Vincent Wai-Sun Wong has served as a consultant and speaker for Echosens and an independent director for Furui. Wah-Kheong Chan has served as a speaker for Echosens. The remaining authors disclose no conflicts.

Funding

This study was supported by a direct grant from The Chinese University of Hong Kong (2022.031).

Data Availability

Data are available upon reasonable request.