scientific reports

Check for updates

OPEN

CLEIA-based type IV collagen 7S for second-step assessment of MRE-defined clinically significant fibrosis in MASLD with indeterminate FIB-4

Shiho Murakami[1], Miwa Kawanaka[1]✉, Kento Imajo[2], Syohei Shiota[1], Nozomi Miyake[1], Masahiko Sue[1], Takuya Adachi[1], Yasuto Takeuchi[1], Hideki Onishi[1], Akinobu Takaki[1], Noriyo Urata[3], Hirofumi Kawamoto[3] & Motoyuki Otsuka[1]

Stepwise MASLD pathways based on the fibrosis-4 (FIB-4) index leave a large indeterminate subgroup requiring second-step assessment. We evaluated automated chemiluminescent enzyme immunoassay (CLEIA)-based type IV collagen 7S (COL4-7S) as a candidate second-line test for MRE-defined clinically significant fibrosis (≥F2). In this retrospective derivation-validation study, adults with MASLD underwent clinically indicated magnetic resonance elastography (MRE) with serum biomarkers measured within 30 days in a two-center derivation cohort (n=347) and an independent external validation cohort (n=126). COL4-7S was compared with ELF, M2BPGi, and FIB-4 using MRE-defined fibrosis stage (F0-F4). In the derivation cohort, MRE liver stiffness correlated with COL4-7S (r=0.59), ELF (r=0.58), M2BPGi (r=0.62), and FIB-4 (r=0.49; all p<0.001), with similar ordinal discrimination across F0-F4. Within the indeterminate FIB-4 range, second-step COL4-7S correctly classified 90/136 patients (66%) for MRE-defined ≥F2; ELF and M2BPGi correctly classified 90/136 (66%) and 73/136 (54%), respectively. In external validation, COL4-7S and ELF correctly classified 32/50 (64%) and 34/50 (68%) patients, respectively. These findings indicate numerically similar second-step performance of CLEIA-based COL4-7S and ELF for MRE-defined ≥F2 fibrosis in this retrospective cohort. Prospective studies are needed to determine whether COL4-7S improves workflow, cost-effectiveness, and patient outcomes in routine MASLD pathways.

Keywords Liver stiffness, MR elastography, ELF score

Metabolic dysfunction-associated steatotic liver disease (MASLD) is the most common chronic liver disease worldwide, and fibrosis stage is the principal determinant of liver-related and overall outcomes[1–4]. Accordingly, noninvasive identification of clinically meaningful fibrosis has become a key requirement in routine care. This includes not only advanced fibrosis (≥F3), which typically prompts specialist referral and intensified surveillance[5,6], but also clinically significant fibrosis (≥F2), which is increasingly relevant to treatment decisions in non-cirrhotic MASH[7,8].

Current clinical pathways generally employ a stepwise noninvasive testing strategy anchored by the fibrosis-4 (FIB-4) index[9,10] because of its simplicity, low cost, and broad availability. However, its major limitation is the large indeterminate ("gray-zone") range, in which first-line assessment does not provide sufficiently actionable triage and a reliable second-step test is required[5,6,11]. Although elastography-based modalities can address this gap, limited access and throughput continue to support the need for serum-based second-line tests[12–14]. The Enhanced Liver Fibrosis (ELF) test[15–17] is an established option, particularly for prioritizing referral for advanced

[1]Department of Gastroenterology and Hepatology, Okayama University Graduate School of Medicine, 2-5-1 Shikata-cho, Kita-ku, Okayama City, Okayama 700-8558, Japan. [2]Department of Gastroenterology, Shin-Yurigaoka General Hospital, 255 Kanagawa,Kosawamiyako, Kawasaki City, Kanagawa 215-0026, Japan. [3]Department of General Internal Medicine 2, General Medical Center, Kawasaki Medical School, 2-6-1 Nakasange, Kita-ku, Okayama City, Okayama 700-8505, Japan. ✉email: m.kawanaka@okayama-u.ac.jp

fibrosis, but its reliance on three analytes may constrain implementation depending on laboratory infrastructure, workflow, and cost[18].

At the same time, the role of second-step testing in patients with indeterminate FIB-4 is changing. As emerging care pathways and pharmacologic strategies increasingly focus on non-cirrhotic MASH with moderate-to-advanced fibrosis, particularly F2-F3 disease[7,8], second-line testing may be useful not only for referral-oriented triage of advanced fibrosis but also for identifying patients in the gray zone who are likely to have treatment-relevant fibrosis and may require confirmatory assessment or therapeutic consideration. Thus, the optimal positioning of second-line tests may depend on the clinical endpoint.

Type IV collagen 7S (COL4-7S) is a serum fibrosis marker reflecting basement membrane collagen turnover and extracellular matrix remodeling. Although historically measured by radioimmunoassay (RIA)[19–24], it can now be quantified by chemiluminescent enzyme immunoassay (CLEIA) on routine automated platforms, improving suitability for high-throughput clinical practice[25,26]. However, evidence for CLEIA-based COL4-7S in MASLD remains limited, particularly in direct comparison with established noninvasive tests and in two-step strategies focused on the indeterminate FIB-4 population.

In this study, we compared CLEIA-measured COL4-7S with ELF, M2BPGi[27–29], and FIB-4 using an MRE-defined, rather than histologic, ordinal fibrosis reference and ordinal discrimination metrics, including the distance-weighted Obuchowski measure[30,31]. We further assessed its performance in patients with indeterminate FIB-4 values to explore whether CLEIA-based COL4-7S could serve as a second-line test for identifying MRE-defined treatment-relevant fibrosis (≥F2, particularly F2-F3), while benchmarking its performance for advanced fibrosis against established second-line markers. No formal equivalence or non-inferiority hypothesis was prespecified; therefore, between-marker findings were interpreted descriptively.

Materials and methods

Study design and participants

This retrospective derivation-validation study included a two-center derivation cohort and an independent external validation cohort. For the derivation cohort, we retrospectively identified adults with MASLD at Kawasaki Medical School General Medical Center and Okayama University Hospital who underwent clinically indicated liver MRE between March 2022 and October 2025 and had serum fibrosis markers measured within 30 days of MRE. After application of the eligibility criteria, 347 patients were included in the final derivation cohort. Patient enrollment was completed before manuscript preparation and analysis finalization.

For the external validation cohort, we retrospectively identified 126 adults with MASLD at Shin-Yurigaoka General Hospital who underwent clinically indicated MRE between 2024 and 2025 and had serum fibrosis markers measured within 30 days of MRE. The same clinical definition of MASLD and the same eligibility framework were applied to both cohorts. The manuscript was prepared after completion of the study period; guidance published after patient enrollment was cited only to contextualize treatment-relevant F2-F3 disease and was not used for patient selection, staging, or statistical analysis.

MASLD was defined as imaging evidence of hepatic steatosis plus at least one cardiometabolic risk factor, after exclusion of other causes of chronic liver disease, in accordance with contemporary consensus criteria[32]. Key exclusion criteria included hepatitis B or C infection, autoimmune liver disease, drug-induced liver injury, and inherited or metabolic liver diseases. Alcohol intake was required to be low (<30 g/day for men and <20 g/day for women), as assessed by structured clinical interview and chart review.

Ethics

This retrospective multicenter study was conducted under centralized ethics review and approved by the Ethics Committee of Okayama University, which served as the central institutional review board for all participating institutions, including Kawasaki Medical Center and Shin-Yurigaoka General Hospital (approval No. 2508-029, and 2508-030). The study was conducted in accordance with the Declaration of Helsinki. Because of the retrospective nature of the study, informed consent was obtained using an opt-out approach, as approved by the ethics committee.

MRE acquisition and fibrosis staging

Liver MRE was performed on 3.0-T MRI systems (Siemens) using standard clinical sequences with a vibration frequency of 60 Hz. Liver stiffness maps were reconstructed using vendor-provided software. Regions of interest were placed in the right hepatic lobe while avoiding large vessels, bile ducts, and artifact-prone areas, and mean liver stiffness was recorded in kilopascals (kPa). Because MRE was used as the reference standard, all diagnostic endpoints in this study are reported as MRE-defined fibrosis categories and should not be interpreted as histologic diagnoses.

MRE-defined fibrosis stage (F0-F4) was assigned using established MRE thresholds for MASLD derived from an individual participant data meta-analysis: <2.65 kPa (F0), 2.65-<3.14 kPa (F1), 3.14-<3.53 kPa (F2), 3.53-<4.45 kPa (F3), and ≥4.45 kPa (F4)[12]. These fixed thresholds were used for analytic consistency, but they may not fully capture histologic heterogeneity or biopsy-defined fibrosis stage.

Serum biomarkers and clinical variables

Serum COL4-7S was measured using a chemiluminescent enzyme immunoassay (CLEIA; Fujirebio, Tokyo, Japan)[25,26]. The ELF score was calculated from hyaluronic acid, procollagen III N-terminal peptide, and tissue inhibitor of metalloproteinases-1 measured on the Siemens platform[15–17]. M2BPGi was measured using a lectin-antibody immunoassay and reported as a cut-off index (Sysmex, Kobe, Japan)[27,29]. COL4-7S and ELF were measured using the same commercial assay systems in the derivation and external validation cohorts; M2BPGi was available for the derivation analysis but was not included in the external validation analysis. Blood sampling

was performed as part of routine clinical care within the 30-day MRE window, and assays were conducted according to manufacturer-recommended procedures at the participating institutions. Preanalytical handling was based on local routine laboratory workflow rather than a centrally standardized research protocol.

The FIB-4 index was calculated as: FIB-4 = [age (years) × AST (U/L)] / [platelet count (10^9/L) × √ALT (U/L)][9,10]. Demographic characteristics, comorbidities, available medication information, and routine laboratory data were extracted from electronic medical records.

Statistical analysis

Continuous variables are presented as median (range) and categorical variables as number (%). Associations between MRE liver stiffness and each serum marker were assessed using Spearman's rank correlation coefficient. For visualization and regression-based summaries, liver stiffness and/or biomarker values were log-transformed as appropriate to reduce skewness and improve interpretability. To characterize transportability, baseline characteristics of the derivation and external validation cohorts were summarized side by side, including the prevalence of MRE-defined ≥F2 and ≥F3 fibrosis (Supplementary Table S1).

Overall ordinal discrimination across MRE-defined fibrosis stages (F0-F4) was summarized using the distance-weighted Obuchowski measure[30,31] with inverse-distance weights ($w = 1/|i-j|$), which integrates pairwise discrimination across all stage combinations while penalizing larger stage misclassifications more heavily than adjacent-stage errors. Stage-stratified bootstrap resampling (2,000 iterations) was used to estimate 95% confidence intervals (CIs). Because no equivalence or non-inferiority margin was prespecified, overlapping CIs were not interpreted as evidence of equivalence, and between-marker results were described as numerical differences rather than proof of comparability. Where applicable, resampling was stratified by center.

Diagnostic performance for MRE-defined clinically significant fibrosis (≥F2) and MRE-defined advanced fibrosis (≥F3) was evaluated using receiver operating characteristic (ROC) curves, and AUROCs were reported. In the derivation cohort, data-driven operating points were additionally explored using the adjusted Youden index (AYI)[33] as a sensitivity analysis, and corresponding sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with 95% CIs were calculated. To reduce overinterpretation of data-derived cutoffs, these thresholds were considered cohort-specific exploratory operating points rather than universal clinical cutoffs.

For evaluation of the two-step algorithm, analyses were restricted to patients with indeterminate FIB-4 values (1.30-2.66). In the derivation cohort, second-line markers (COL4-7S, ELF, and M2BPGi) were assessed using the ≥F2 operating points identified in the sensitivity analysis. In the external validation cohort, the prespecified ≥F2 cutoffs derived in the derivation cohort were applied without re-estimation. Classification performance against MRE-defined ≥F2 was summarized using overall correct classification and related measures, and concordant and discordant classifications were visualized using quadrant plots.

All analyses were performed in R (version 4.5.2). All statistical tests were two-sided, and P values <0.05 were considered statistically significant. Complete-case analysis was used; there were no missing data for the primary variables, including MRE liver stiffness and serum fibrosis markers.

Results

Study population and baseline characteristics

A total of 347 adults with MASLD in whom liver fibrosis stage was noninvasively estimated by MRE were included in the derivation analysis (Supplementary Table S2). The median age was 62 years (range 20–86), and 159 (46%) were male. The median body mass index was 26.9 kg/m^2 (range 18–40.3). Type 2 diabetes mellitus, hypertension, and dyslipidemia were present in 62%, 47%, and 85% of patients, respectively. Median liver enzymes were ALT 37 U/L (11–298), AST 31 U/L (11–154), and γ-GTP 39 U/L (11–538), with a median platelet count of 22 ×10^4/μL (8.6–28.6). Median serum fibrosis markers were COL4-7S 3.8 ng/mL (1.4–12), ELF 9.4 (7.2–13.1), M2BPGi 0.75 C.O.I. (0.24–9.1), and FIB-4 1.3 (0.2–5.9), and the median MRE liver stiffness was 2.6 kPa (1.18–12.57) (Supplementary Table S2). Using the predefined MRE thresholds (<2.65, 2.65–<3.14, 3.14–<3.53, 3.53–<4.45, and ≥4.45 kPa)[12], fibrosis stage distribution was F0: 173 (49.9%), F1: 43 (12.4%), F2: 27 (7.8%), F3: 43 (12.4%), and F4: 61 (17.6%) (Supplementary Table S3). A side-by-side comparison of the derivation and external validation cohorts is provided in Supplementary Table S1. The prevalence of MRE-defined ≥F2 fibrosis was 37.8% in the derivation cohort and 35.7% in the validation cohort, and the prevalence of MRE-defined ≥F3 fibrosis was 30.0% and 30.2%, respectively.

Correlation of serum fibrosis markers with MRE liver stiffness and stage-wise trends

In scatter plots using log-transformed values, all four serum fibrosis markers were positively associated with MRE-derived liver stiffness (Fig. 1). Correlations were moderate for COL4-7S (r=0.59), ELF (r=0.58), and M2BPGi (r=0.62), whereas FIB-4 showed a weaker correlation (r=0.49). These findings suggest that each marker captures, to varying degrees, fibrosis-related variation reflected by MRE, while also highlighting substantial inter-individual dispersion across the stiffness spectrum.

When stratified by MRE-based fibrosis stage (F0–F4), all markers showed an overall upward shift with advancing stage (Fig. 2). Nonetheless, the distributions exhibited considerable overlap, particularly across lower-to-intermediate stages, whereas higher values were increasingly concentrated in advanced fibrosis (F3–F4). Taken together, these results indicate that serum markers track MRE both as continuous measures and in stage-wise trends, but with overlap that can limit clear separation at the individual level, supporting their evaluation as second-line tests to refine risk among clinically ambiguous groups, particularly patients with indeterminate FIB-4, where a single index often fails to yield an actionable decision.

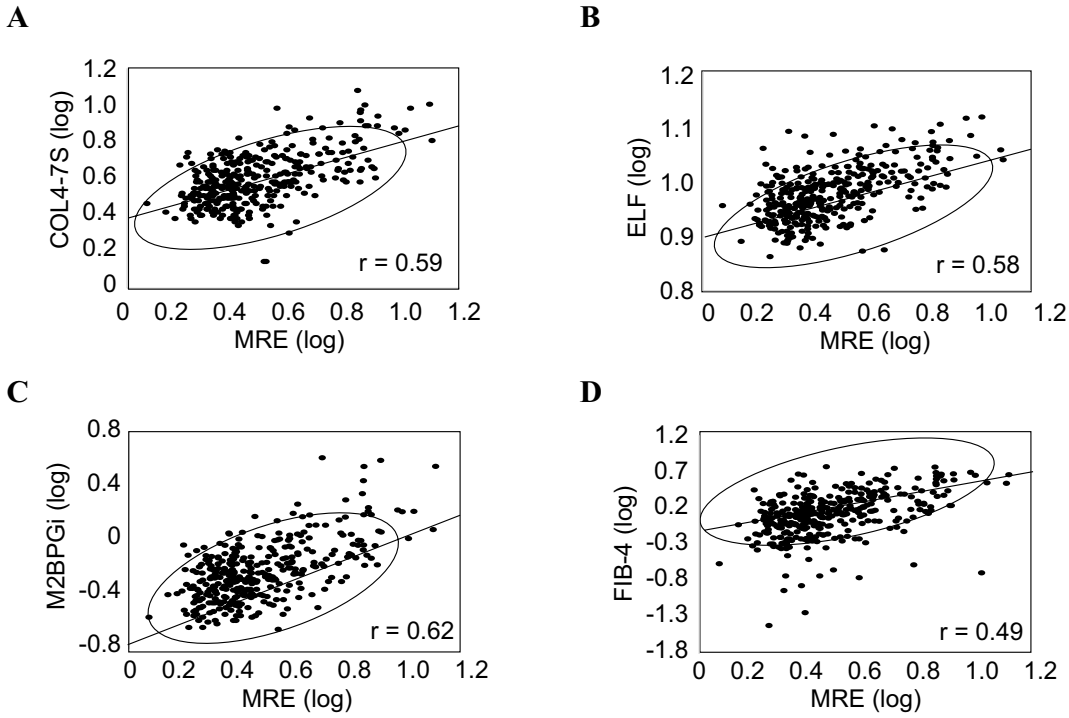

Fig. 1. Associations between MRE liver stiffness and serum fibrosis markers Scatterplots show log-transformed magnetic resonance elastography (MRE) liver stiffness (kPa; x-axis) versus log-transformed values of (**A**) type IV collagen 7S (COL4-7S; ng/mL, CLEIA), (**B**) ELF score, (**C**) Mac-2 binding protein glycosylation isomer (M2BPGi; cutoff index [C.O.I.]), and (**D**) fibrosis-4 (FIB-4) index (y-axes). Each point represents one participant. Solid lines indicate fitted linear regression, and ellipses depict the overall data distribution as a visual guide. Spearman rank correlation coefficients are shown in each panel (COL4-7S r=0.59, ELF r=0.58, M2BPGi r=0.62, FIB-4 r=0.49; all p<0.001).

Diagnostic accuracy across ordinal MRE fibrosis stages

Next, we summarized the overall ordinal diagnostic accuracy of each marker against MRE-defined fibrosis stages (F0–F4) using the distance-weighted Obuchowski measure (Fig. 3)[30,31]. The Obuchowski measure is an AUC-like metric designed for ordered, multi-category reference standards; it summarizes discrimination across the full staging spectrum while penalizing larger stage misclassifications more heavily than adjacent-stage misclassifications. Point estimates were numerically close across COL4-7S, ELF, M2BPGi, and FIB-4, and the 95% confidence intervals overlapped. Because no equivalence or non-inferiority margin was prespecified, these results should be interpreted as descriptive numerical similarity rather than statistical equivalence.

AYI-derived operating points for clinically relevant fibrosis thresholds

To provide clinically interpretable operating points, we evaluated threshold-based classification for significant fibrosis (≥F2) and advanced fibrosis (≥F3) using adjusted Youden index (AYI)–derived operating points as data-driven sensitivity analyses (Table 1)[33]. For ≥F2, AYI operating points were COL4-7S 3.6 ng/mL, ELF 9.3, M2BPGi 0.67 C.O.I., and FIB-4 1.28, with AUROCs ranging from 0.78 to 0.85. For ≥F3, AYI operating points were COL4-7S 5.3 ng/mL, ELF 10.1, M2BPGi 1.04 C.O.I., and FIB-4 2.1, with AUROCs ranging from 0.80 to 0.83 (Table 1).

Notably, these AYI-derived operating points highlighted distinct trade-offs across markers, underscoring that a single optimal threshold is inherently endpoint- and purpose-dependent. For the ≥F2 endpoint, COL4-7S (CLEIA) provided an operating point that favored sensitivity/NPV (cutoff 3.6 ng/mL; sensitivity 0.86; specificity 0.53; PPV 0.53; NPV 0.86; AUROC 0.78), whereas ELF showed a more balanced sensitivity–specificity profile (cutoff 9.3; sensitivity 0.83; specificity 0.66; PPV 0.58; NPV 0.86; AUROC 0.79). M2BPGi yielded a similar sensitivity-dominant profile with modest specificity (cutoff 0.67; sensitivity 0.84; specificity 0.54; PPV 0.52; NPV 0.85; AUROC 0.79), and for context, the FIB-4 operating point for ≥F2 (cutoff 1.28; sensitivity 0.82; specificity 0.60; PPV 0.55; NPV 0.85; AUROC 0.85) reflects its role as a first-line screen rather than a dedicated second-step rule-in test. For ≥F3, COL4-7S yielded a high-specificity operating point (sensitivity 0.42; specificity 0.93; PPV 0.70; NPV 0.80), whereas ELF and M2BPGi maintained a more balanced sensitivity–specificity profile (ELF: sensitivity 0.62, specificity 0.87; M2BPGi: sensitivity 0.63, specificity 0.83).

These AYI-derived thresholds were used only as data-driven operating points in sensitivity analyses to illustrate marker-specific trade-offs for clinically relevant endpoints.

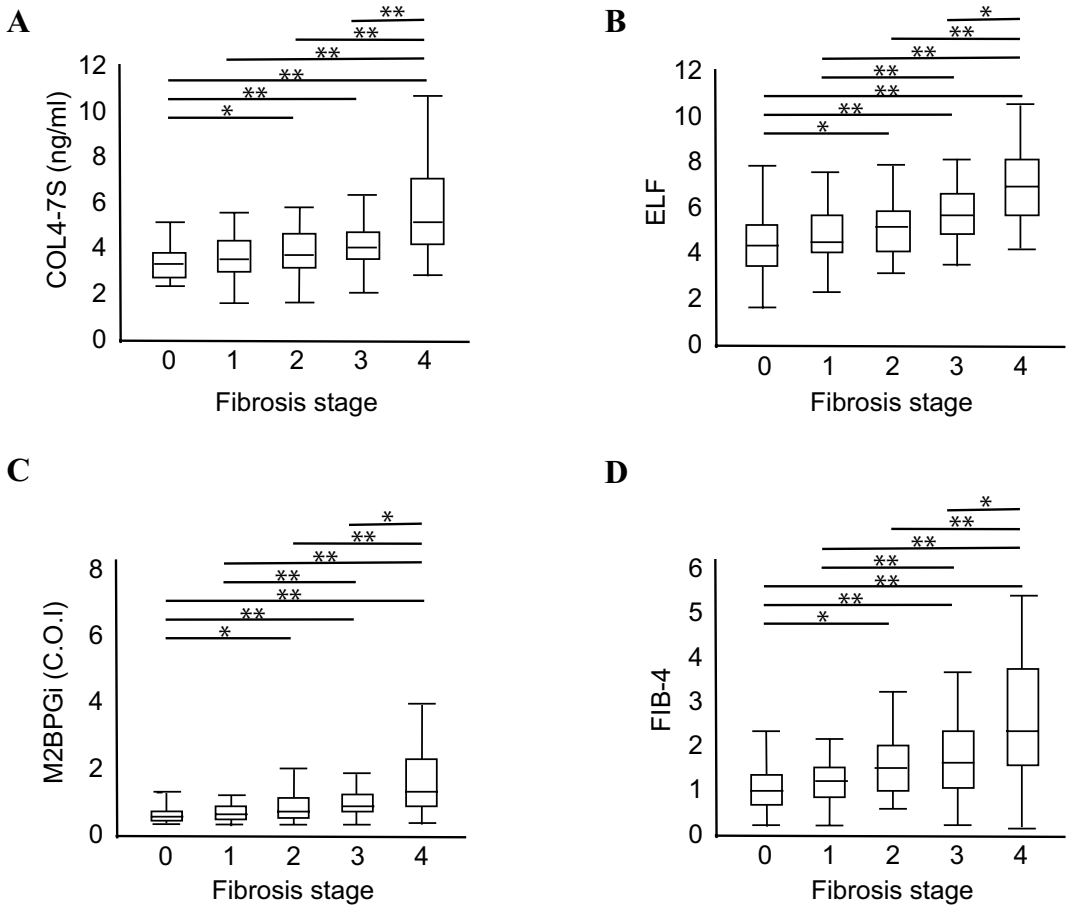

Fig. 2. Distributions of serum fibrosis markers across MRE-defined fibrosis stages in MASLD Box-and-whisker plots show (**A**) COL4-7S (ng/mL), (**B**) ELF score, (**C**) M2BPGi (cut-off index [C.O.I.]), and (**D**) FIB-4 index, across MRE-defined fibrosis stages (F0–F4; shown as stages 0–4 on the x-axis). Boxes indicate the median and interquartile range (IQR); whiskers extend to 1.5×IQR; points denote outliers. Horizontal brackets indicate statistically significant pairwise differences between fibrosis stages based on a Kruskal–Wallis test with post hoc pairwise comparisons (*P<0.05; **P<0.01).

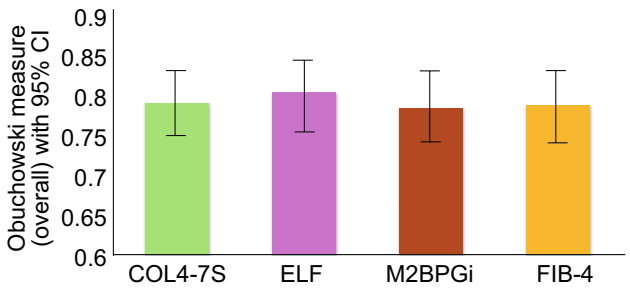

Fig. 3. Overall ordinal diagnostic accuracy of serum fibrosis markers against MRE-defined fibrosis stage Bars show point estimates of the Obuchowski measure (ordinal AUC) for COL4-7S, ELF, M2BPGi, and FIB-4 using MRE-defined fibrosis stage (F0–F4) as the ordinal reference standard. Error bars indicate 95% confidence intervals obtained by stratified bootstrap resampling (2,000 replicates). For overall performance across F0–F4, ELF (0.803; 95% CI 0.763–0.841), COL4-7S (0.789; 95% CI 0.755–0.825), FIB-4 (0.787; 95% CI 0.747–0.827), and M2BPGi (0.783; 95% CI 0.740–0.822) showed numerically close estimates. These overlapping confidence intervals should not be interpreted as evidence of statistical equivalence.

Second-line stratification within the indeterminate FIB-4 range

We next focused on patients only in the indeterminate FIB-4 range (1.30–2.66; n=136), a clinically important bottleneck in stepwise MASLD pathways, to evaluate a pragmatic two-step strategy in which a second-line serum marker refines risk against the MRE reference (Fig. 4). Within this gray zone, MRE classified 72 patients

Biomarker	Unit	AYI cutoff (≥F2)	Sensitivity	Specificity	PPV	NPV	AUROC
COL4-7S (CLEIA)	ng/mL	3.6	0.86	0.53	0.53	0.86	0.78
ELF	unitless	9.3	0.83	0.66	0.58	0.86	0.79
M2BPGi	C.O.I.	0.67	0.84	0.54	0.52	0.85	0.79
FIB-4	unitless	1.28	0.82	0.60	0.55	0.85	0.85
Biomarker	**Unit**	**AYI cutoff (≥F3)**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC**
COL4-7S (CLEIA)	ng/mL	5.3	0.42	0.93	0.70	0.80	0.81
ELF	unitless	10.1	0.62	0.87	0.67	0.84	0.83
M2BPGi	C.O.I.	1.04	0.63	0.83	0.62	0.84	0.82
FIB-4	unitless	2.1	0.55	0.87	0.65	0.82	0.80

Table 1. Sensitivity analysis: AYI-derived operating points and diagnostic performance of noninvasive biomarkers for MRE-defined ≥F2 and ≥F3 fibrosis in MASLD in the derivation cohort. AYI-derived thresholds were evaluated as data-driven operating points in a sensitivity analysis and should be interpreted as cohort-specific rather than universal cutoffs. AYI, adjusted Youden index; AUROC, area under the receiver operating characteristic curve; CLEIA, chemiluminescent enzyme immunoassay; C.O.I., cut-off index; COL4-7S, type IV collagen 7S; ELF, Enhanced Liver Fibrosis; FIB-4, fibrosis-4; MASLD, metabolic dysfunction-associated steatotic liver disease; M2BPGi, Mac-2 binding protein glycosylation isomer; MRE, magnetic resonance elastography; NPV, negative predictive value; PPV, positive predictive value.

as ≤F1 and 64 patients as ≥F2. Using the ≥F2 operating points (AYI-derived; Table 1), quadrant-based plots showed that COL4-7S and ELF yielded the same overall correct classification for MRE-defined ≥F2: COL4-7S correctly classified 90/136 patients (66%; TP=53, TN=37), and ELF likewise correctly classified 90/136 (66%; TP=52, TN=38). In contrast, M2BPGi showed lower overall accuracy (73/136, 54%; TP=49, TN=24), driven by a greater number of discordant classifications (Fig. 4).

Collectively, these findings suggest that, among patients left "uncertain" after first-line screening, COL4-7S provided second-step enrichment for MRE-defined treatment-relevant fibrosis (≥F2) with a sensitivity-oriented profile and an overall correct classification rate numerically similar to ELF. However, the modest overall accuracy and the presence of false-positive and false-negative classifications indicate that COL4-7S should not be interpreted as a definitive diagnostic test.

External validation of second-step serum marker stratification within the indeterminate FIB-4 range

Finally, we externally evaluated COL4-7S and ELF as second-step serum markers for identifying MRE-defined clinically significant fibrosis (≥F2), a treatment-relevant stage, among patients with indeterminate FIB-4 values. An independent MASLD cohort comprising 126 patients with paired MRE and serum biomarker data was analyzed. The median age was 59 years (range, 19–83), 78 patients (61%) were male, and the median body mass index was 26.8 kg/m^2 (range, 18–43.5). Median values were 3.6 ng/mL (range, 0.7–12.0) for COL4-7S, 9.5 (range, 6.9–13.1) for ELF, 1.5 (range, 0.3–6.7) for FIB-4, and 2.6 kPa (range, 1.67–12.0) for MRE (Supplementary Table S4). Compared with the derivation cohort, the external validation cohort had a higher proportion of men and lower prevalences of type 2 diabetes, hypertension, and dyslipidemia, whereas BMI, median MRE liver stiffness, and the prevalences of MRE-defined ≥F2 and ≥F3 fibrosis were broadly similar (Supplementary Table S1). Based on the predefined MRE thresholds (<2.65, 2.65–<3.14, 3.14–<3.53, 3.53–<4.45, and ≥4.45 kPa), fibrosis stages were distributed as F0 in 64 patients (50.8%), F1 in 17 (13.5%), F2 in 7 (5.6%), F3 in 17 (13.5%), and F4 in 21 (16.7%) (Supplementary Table S5).

Among these 126 patients, 50 (39.7%) had indeterminate FIB-4 values (1.30–2.66) and underwent second-step stratification (Fig. 5). Within this subgroup, MRE classified 26 patients as ≤F1 and 24 as ≥F2. Using the prespecified ≥F2 cutoffs derived in the derivation cohort (AYI-derived; Table 1), COL4-7S and ELF showed numerically similar second-step classification performance in the external validation cohort. COL4-7S correctly classified 32 of 50 patients (64%; 23 true positives and 9 true negatives), whereas ELF correctly classified 34 of 50 patients (68%; 22 true positives and 12 true negatives) (Fig. 5). COL4-7S yielded fewer false negatives than ELF (1 vs. 2), whereas ELF yielded fewer false positives (14 vs. 17), indicating a modest trade-off between sensitivity and specificity in this gray-zone population.

These findings suggest that the numerical pattern observed in the derivation cohort was reproduced externally: CLEIA-based COL4-7S provided second-step enrichment for MRE-defined clinically significant fibrosis within the indeterminate FIB-4 range, with performance close to ELF but without evidence of statistical equivalence. Prospective studies in routine care pathways are required to confirm transportability and clinical impact.

Discussion

In this retrospective derivation-validation study of adults with MASLD staged by MRE, CLEIA-measured COL4-7S showed numerically similar overall performance to ELF, M2BPGi, and FIB-4. Notably, within the clinically important indeterminate FIB-4 range[5,6], COL4-7S provided second-step classification for MRE-defined clinically significant fibrosis (≥F2) with correct classification rates close to ELF, while retaining the

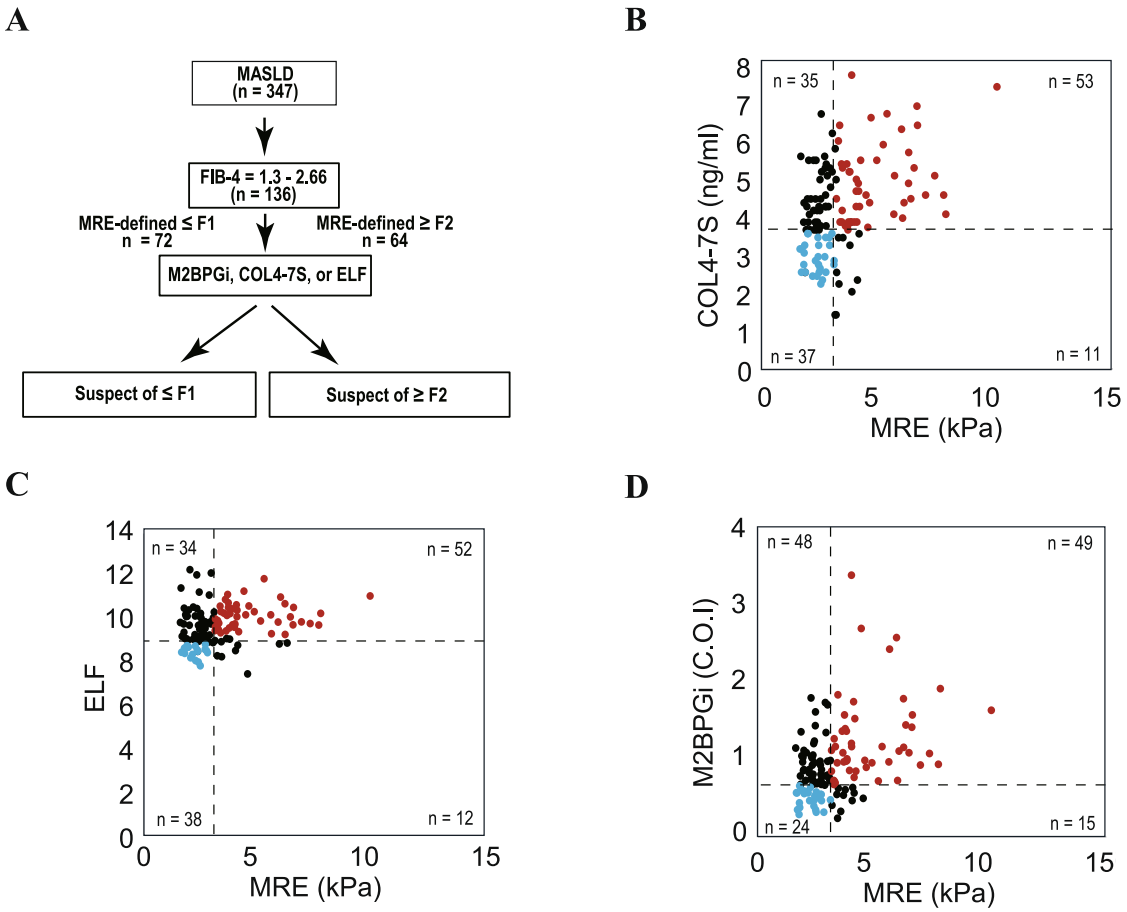

Fig. 4. Second-step stratification within the indeterminate FIB-4 range for MRE-defined clinically significant fibrosis (≥F2) using serum fibrosis markers against the MRE reference (**A**) Study workflow for second-step testing among patients with indeterminate FIB-4 (1.30–2.66; n=136), with MRE as the reference standard dichotomized as ≤F1 (n=72) vs ≥F2 (n=64). (**B**–**D**) Scatterplots restricted to the indeterminate FIB-4 subgroup showing MRE liver stiffness (kPa; x-axis) versus (**B**) COL4-7S (ng/mL, CLEIA), (**C**) ELF score, and (**D**) M2BPGi (cutoff index [C.O.I.]) (y-axes). The vertical dashed line indicates the MRE threshold for ≥F2 (3.14 kPa), and horizontal dashed lines indicate AYI-derived biomarker cutoffs for ≥F2 (sensitivity analysis; COL4-7S 3.6 ng/mL, ELF 9.3, M2BPGi 0.67 C.O.I.). Each dot represents one participant; red points indicate concordant true positives (MRE ≥F2 and biomarker above cutoff), blue points indicate concordant true negatives (MRE ≤F1 and biomarker below cutoff), and black points indicate discordant classifications (false positives/false negatives). Numbers within each quadrant denote participant counts.

practical feature of being a single-analyte automated assay. These findings should be interpreted as exploratory and descriptive rather than as evidence of equivalence.

Across the full cohort, all markers correlated with MRE liver stiffness and increased in a stage-dependent manner, supporting their relevance as noninvasive indicators of fibrosis burden. At the same time, the substantial overlap observed between adjacent fibrosis stages, particularly in the lower-to-intermediate range, underscores an important limitation of serum markers and simple indices: although they reflect fibrosis severity on a continuum, they do not fully resolve early-stage disease at the individual level. This observation reinforces the rationale for stepwise clinical pathways[5,6], in which first-line screening identifies patients who require second-step refinement rather than definitive staging by any single test.

To account for the ordinal structure of the MRE reference, we used the distance-weighted Obuchowski measure[30,31] to summarize discrimination across F0-F4 while assigning greater penalty to wider stage misclassification. By this metric, COL4-7S, ELF, M2BPGi, and FIB-4 showed close point estimates with overlapping 95% confidence intervals. However, overlapping confidence intervals do not establish equivalence, and no formal non-inferiority or equivalence margin was prespecified. Therefore, the clinical value of these tests should be considered in relation to specific endpoints and intended use rather than in terms of a single universally superior marker.

From a clinical standpoint, the most informative analysis was that restricted to patients within the indeterminate FIB-4 range (1.30–2.66), which remains a major decision bottleneck in contemporary MASLD care pathways[5,6]. In the derivation cohort, COL4-7S and ELF yielded the same overall correct classification for MRE-defined ≥F2, whereas M2BPGi performed less favorably. These findings suggest that COL4-7S may be a

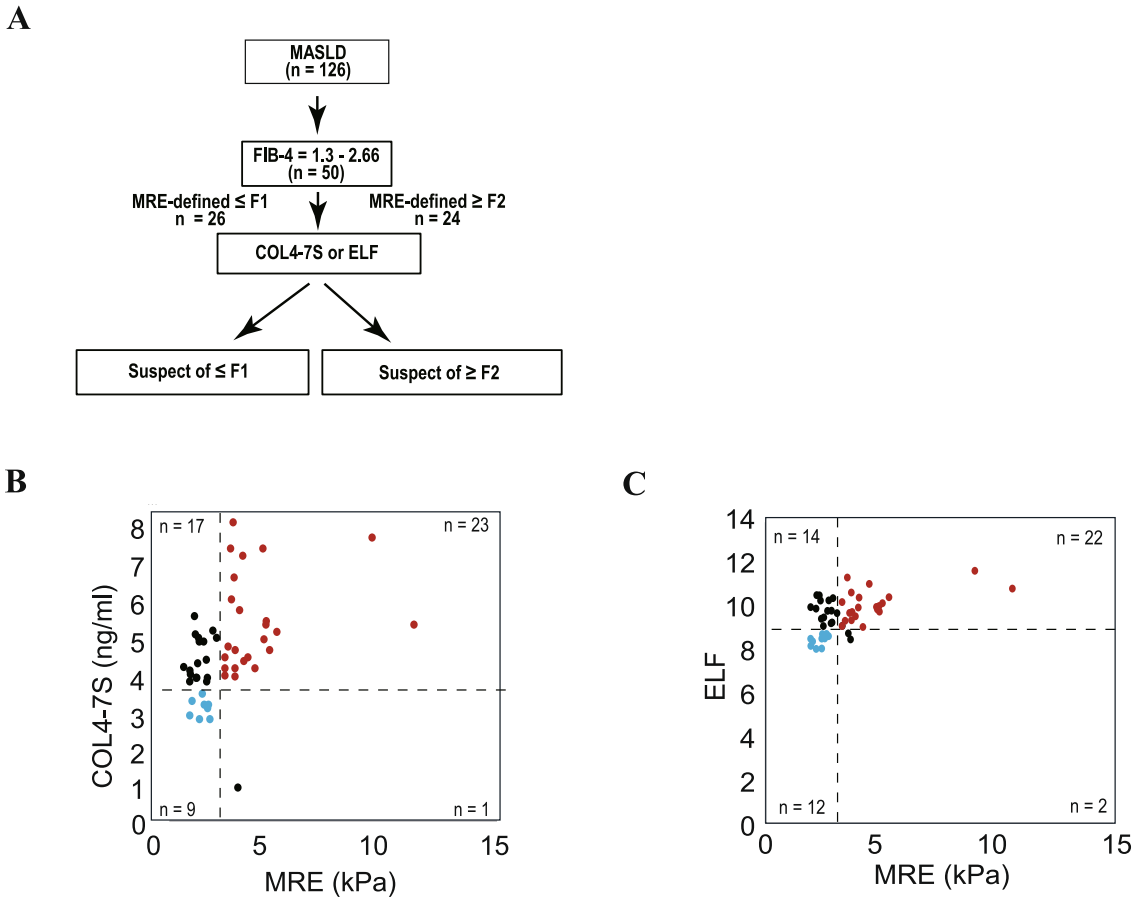

Fig. 5. External validation of second-step stratification within the indeterminate FIB-4 range for MRE-defined clinically significant fibrosis (≥F2) using serum fibrosis markers against the MRE reference (**A**) Study workflow for second-step testing among patients in the external validation cohort. Of 126 patients with MASLD, 50 had indeterminate FIB-4 values (1.30-2.66) and were included in the second-step analysis. Using MRE as the reference standard, this subgroup was dichotomized as ≤F1 (n=26) or ≥F2 (n=24). (B, C) Scatterplots restricted to the indeterminate FIB-4 subgroup showing MRE liver stiffness (kPa; x-axis) versus (**B**) COL4-7S (ng/mL, CLEIA) and (**C**) ELF score (y-axes). The vertical dashed line indicates the MRE threshold for ≥F2 (3.14 kPa), and the horizontal dashed lines indicate the AYI-derived biomarker cutoffs for ≥F2 established in the derivation cohort (COL4-7S, 3.6 ng/mL; ELF, 9.3). Each dot represents one participant; red points indicate concordant true positives (MRE ≥F2 and biomarker above cutoff), blue points indicate concordant true negatives (MRE ≤F1 and biomarker below cutoff), and black points indicate discordant classifications (false positives or false negatives). Numbers within each quadrant denote participant counts. For COL4-7S, the quadrant counts were 23 true positives, 9 true negatives, 17 false positives, and 1 false negative; for ELF, the corresponding counts were 22, 12, 14, and 2.

useful second-line option for refining risk among patients who remain unclassified after first-line screening, particularly when the goal is to identify individuals likely to harbor MRE-defined treatment-relevant fibrosis and therefore warrant confirmatory assessment and therapeutic consideration.

This numerical pattern was also observed in the external validation cohort. Among patients with indeterminate FIB-4 values, second-step classification for MRE-defined ≥F2 remained close between COL4-7S and ELF, with correct classification in 64% and 68% of patients, respectively. Although the balance between false-positive and false-negative classifications differed modestly between the two markers, the external dataset supports the reproducibility of the two-step signal beyond the derivation cohort, while still requiring prospective confirmation in routine clinical pathways.

The appeal of COL4-7S is supported by practical assay features but should not be overstated. Whereas ELF is a multi-analyte composite score[16,17], COL4-7S can be measured as a single analyte on routine automated CLEIA platforms[25,26]. In this context, the transition from historical RIA-based assays to CLEIA may facilitate integration into high-throughput laboratory workflows. However, the present study did not directly evaluate turnaround time, cost, adoption barriers, reproducibility, or workflow advantages over ELF. Therefore, COL4-7S has practical features that may facilitate implementation, but formal evaluation of feasibility, cost-effectiveness, and workflow impact is still needed.

Several limitations merit consideration. First, PPV and NPV are inherently prevalence-dependent, and both cohorts comprised patients referred for clinically indicated MRE; accordingly, operating characteristics may differ in lower-prevalence primary care settings. Second, although MRE is well validated and provides whole-liver assessment, it is not a histologic gold standard[13,14,34]. Outcomes in this study therefore represent MRE-defined fibrosis categories based on fixed thresholds, not biopsy-proven fibrosis stage, and these thresholds may misclassify some patients relative to histology. Third, despite external validation, the study was retrospective and the validation cohort was modest in size; AYI-derived thresholds[33] should therefore be interpreted as cohort-specific exploratory operating points rather than universal cutoffs. Fourth, preanalytical handling followed routine local laboratory workflows rather than a centrally standardized research protocol, and medication-use data were not uniformly structured across centers for a reliable comparative analysis. Finally, this study did not assess treatment decisions, workflow implementation, cost, reproducibility, or prospective clinical outcomes. Prospective validation in routine stepwise pathways, ideally incorporating feasibility and health-economic considerations, will be important.

In conclusion, CLEIA-based COL4-7S demonstrated ordinal discrimination in the same numerical range as established noninvasive tests and, within the indeterminate FIB-4 gray zone, provided second-step classification for MRE-defined ≥F2 with performance close to ELF in both derivation and validation cohorts. As a single-analyte automated assay, COL4-7S may be a candidate second-line option for MASLD risk stratification, but its workflow, cost-effectiveness, and impact on treatment decisions and patient outcomes require prospective evaluation.

Data availability

The de-identified data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to institutional and ethical restrictions**.**

Received: 6 May 2026; Accepted: 11 June 2026
Published online: 18 June 2026

References

1. Younossi, Z. M. et al. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. *Hepatology* **77**, 1335–1347 (2023).
2. Estes, C. et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016–2030. *J. Hepatol.* **69**, 896–904 (2018).
3. Fujii, H. et al. Clinical outcomes in biopsy-proven nonalcoholic fatty liver disease patients: a multicenter registry-based cohort study. *Clin. Gastroenterol. Hepatol.* **21**, 370–379 (2023).
4. Taylor, R. S. et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Gastroenterology.* **158**, 1611-1625.e12 (2020).
5. Rinella, M. E. et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* **77**, 1797–1835 (2023).
6. EASL-EASD-EASO Clinical Practice guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). Obes Facts **17** ,374-444 (2024).
7. Bansal, M. B. et al. Semaglutide therapy for metabolic dysfunction-associated steatohepatitis: november 2025 updates to AASLD practice guidance. *Hepatology* **83**, 1326-1340 (2026)
8. Harrison, S. A. et al. A Phase 3, randomized, controlled trial of resmetirom in NASH with liver fibrosis. *N Engl J Med.* **390**, 497–509 (2024).
9. Shah, A. G. et al. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* **7**, 1104–1112 (2009).
10. Sterling, R. K. et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* **43**, 1317–1325 (2006).
11. Srivastava, A. et al. Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. *J. Hepatol.* **71**, 371–378 (2019).
12. Liang, J. X. et al. An individual patient data meta-analysis to determine cut-offs for and confounders of NAFLD-fibrosis staging with magnetic resonance elastography. *J. Hepatol.* **79**, 592–604 (2023).
13. Imajo, K. et al. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than Transient Elastography. *Gastroenterology* **150**, 626-637.e7 (2016).
14. Huwart, L. et al. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. *Gastroenterology* **135**, 32–40 (2008).
15. Rosenberg, W. M. et al. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* **127**, 1704–1713 (2004).
16. Inadomi, C. et al. Accuracy of the enhanced liver fibrosis test, and combination of the enhanced liver fibrosis and non-invasive tests for the diagnosis of advanced liver fibrosis in patients with non-alcoholic fatty liver disease. *Hepatol. Res.* **50**, 682–692 (2020).
17. Younossi, Z. M. et al. Performance of the enhanced liver fibrosis test to estimate advanced fibrosis among patients with nonalcoholic fatty liver disease. *JAMA Netw. Open* **4**, e2123923 (2021).
18. Kamada, Y. et al. SWOT analysis of noninvasive tests for diagnosing NAFLD with severe fibrosis: an expert review by the JANIT Forum. *J Gastroenterol.* **58**, 79–97 (2023).
19. Nishizono, I. et al. Rapid and sensitive chemiluminescent enzyme immunoassay for measuring tumor markers. *Clin. Chem.* **37**, 1639–1644 (1991).
20. Ishiba, H. et al. Accuracy of type IV collagen 7S versus enhanced liver fibrosis score for diagnosing fibrosis in patients with metabolic dysfunction-associated steatotic liver disease. *Hepatol. Commun.* **9**, e0563 (2025).
21. Murawaki, Y. et al. Serum type III procollagen peptide, type IV collagen 7S domain, central triple-helix of type IV collagen and tissue inhibitor of metalloproteinases in patients with chronic viral liver disease: relationship to liver histology. *Hepatology* **20**, 780–787 (1994).
22. Ishiba, H. et al. Prediction of advanced fibrosis in metabolic dysfunction-associated steatotic liver disease by Type IV collagen 7S. *Gastro. Hep. Adv.* **4**, 100668 (2025).
23. Honda, Y. et al. Meta-analysis of the diagnostic accuracy of serum type IV collagen 7S concentration for the staging of liver fibrosis in nonalcoholic fatty liver disease. *Hepatol. Res.* **53**, 219–27 (2023).
24. Yoneda, M. et al. Type IV collagen 7s domain is an independent clinical marker of the severity of fibrosis in patients with nonalcoholic steatohepatitis before the cirrhotic stage. *J. Gastroenterol.* **42**, 375–381 (2007).

25. Kawanaka, M. et al. The new assay type IV collagen 7S (CLEIA) is a useful test to identify progressive fibrosis in chronic liver disease: a study based on magnetic resonance elastography. *Hepatol .Res.* **55**, 994–1004 (2025).
26. Shima, T. et al. A novel rapid immunoassay of serum type IV collagen 7S for the diagnosis of fibrosis stage of nonalcoholic fatty liver diseases. *Hepatol. Res.* **51**, 263–276 (2021).
27. Kuno, A. et al. Reconstruction of a robust glycodiagnostic agent supported by multiple lectin-assisted glycan profiling. *Proteom. Clin. Appl.* **7**, 642–647 (2013).
28. Toshima, T. et al. A novel serum marker, glycosylated Wisteria floribunda agglutinin-positive Mac-2 binding protein (WFA(+)-M2BP), for assessing liver fibrosis. *J. Gastroenterol.* **50**, 76–84 (2015).
29. Shirabe, K. et al. Mac-2 binding protein glycan isomer (M2BPGi) is a new serum biomarker for assessing liver fibrosis: more than a biomarker of liver fibrosis. *J. Gastroenterol.* **53**, 819–826 (2018).
30. Obuchowski, N. A. Comparing the accuracy of diagnostic tests when disease is characterized by an ordinal scale. *Am. J. Epidemiol.* **192**, 632–643 (2023).
31. Obuchowski, N. A. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary. *Acad. Radiol.* **12**, 1198–1204 (2005).
32. Rinella, M. E. et al. A multisociety delphi consensus statement on new fatty liver disease nomenclature. *Ann. Hepatol.* **29**, 101133 (2024).
33. Hua, J. et al. A comprehensive and comparative review of optimal cut-points selection methods for diseases with multiple ordinal stages. *J. Biopharm. Stat.* **30**, 46–68 (2020).
34. Ratziu, V. et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* **128**, 1898–1906 (2005).

Acknowledgments

The authors thank Prof. Takashi Kumada, Department of Nursing, Gifu Kyoritsu University, for his valuable guidance and insightful comments throughout this study and during the preparation of the manuscript.

Author contributions

Miwa Kawanaka and Motoyuki Otsuka designed the study. All authors collected and interpreted the data. Miwa Kawanaka and Noriyo Urata analyzed the data. Shiho Murakami, Miwa Kawanaka, and Motoyuki Otsuka drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology, Japan (#25K22578, #25K03028, and #24K11153 to M.O.), and by grants from the Japan Agency for Medical Research and Development (JP25fk0310540, JP25fk0310541, JPck0106947, JPym0126810 and JP-25fk0210158 to M.O. and JP26fk0210210 to M.K.).

Declarations

Competing interests

Miwa Kawanaka has received consulting fees from Fujirebio, AbbVie Inc, and Souwakai. Akinobu Takaki has received consulting fees from Souwakai. All other authors declare no conflicts of interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-58113-w.

Correspondence and requests for materials should be addressed to M.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026