pRFPS: A Subgroup-Specific Interpretable Risk Score for Predicting Rapid Fibrosis Progression in Metabolic Dysfunction-Associated Steatotic Liver Disease

Tahmina Sultana Priya[1,2], Huihuang Yan[3,4], Tara L. Haynes[5], Kirk J. Wangensteen[6], Andres J. Acosta[6,†], Konstantinos N. Lazaridis[6], Eric W. Klee[3,4], Shulan Tian[3,4*], Danfeng (Daphne) Yao[1,2,7*]

[1]Department of Computer Science, [2]Sanghani Center for Artificial Intelligence and Data Analytics, Virginia Tech, Blacksburg, VA, 24061, USA.
[3]Division of Computational Biology, Department of Quantitative Health Sciences, [4]Center for Individualized Medicine, [5]Center for Digital Health, [6]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Mayo Clinic, Rochester, MN, 55902, USA.
[7]Department of Computer Science, North Carolina State University, Raleigh, NC, 27695, USA.
Email: tpriya@vt.edu; tian.shulan@mayo.edu; danfeng@vt.edu

Metabolic dysfunction-associated steatotic liver disease (MASLD) is highly heterogeneous, and predicting rapid fibrosis progression remains a major clinical challenge, particularly among patients with low baseline fibrosis risk. We developed a partitioned framework that integrates longitudinal laboratory measurements, diagnosis histories, and genetic data to predict rapid fibrosis progression. The framework first applied unsupervised phenotypic partitioning to identify clinically meaningful subgroups. Within each subgroup, multiple feature selection strategies and machine learning models were combined to optimize predictive performance. To enhance interpretability, we developed the partitioned Rapid Fibrosis Progression Score (pRFPS), derived from Shapley Additive Explanations and translated into a weighted clinical risk score. The framework was developed in the Mayo Clinic Biobank and validated in the Tapestry Study cohort. Subgroup-specific modeling improved predictive performance in the MCB test set, with gains of up to 13.5% points in F1 score and 9.5% points in AUROC, while maintaining robust performance in the Tapestry Study cohort. High pRFPS was associated with increased odds of hepatocellular carcinoma and liver transplantation. Notably, pRFPS identified 22%-25% of rapid progressors who were classified as low risk by conventional clinical tools. These findings demonstrate the potential of interpretable, phenotype-aware modeling to improve personalized risk stratification and support earlier intervention for MASLD progression.

Keywords: Metabolic dysfunction-associated steatotic liver disease; rapid fibrosis progression; subgroup-specific modeling; interpretable risk score; machine learning.

1. Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) has emerged as the most prevalent chronic liver disease worldwide, largely driven by the global epidemics of obesity, type 2 diabetes (T2D), and metabolic syndrome[1-3]. Current estimates indicate that MASLD affects approximately 38% of adults globally, with projections exceeding 55% by 2040 in the absence of effective interventions[4]. This rising prevalence is expected to substantially elevate the burden of

[†] Conflict of interests declared by A.J.A. are provided in the Supplementary Materials (https://drive.google.com/PSB 2027) [*] Co-corresponding authors.

© 2026 The Authors. Open Access chapter published by World Scientific Publishing Company and distributed under the terms of the Creative Commons Attribution Non-Commercial (CC BY-NC) 4.0 License.

advanced fibrosis, cirrhosis, hepatocellular carcinoma (HCC), and liver-related mortality[5], positioning MASLD as a major and growing public health challenge.

Fibrosis stage remains the strongest predictor of liver-related outcomes in MASLD[6]. However, the rate of fibrosis progression is highly heterogeneous. While the average time to fibrosis progression is 7-14 years[7], approximately 18% of patients experience rapid progression, advancing to advanced fibrosis or cirrhosis within a few years[7-9]. Current diagnostic tools, including the gold standard liver biopsy[10] and non-invasive tests[11] such as the fibrosis-4 index (FIB-4), nonalcoholic fatty liver disease fibrosis score (NFS) and vibration-controlled transient elastography, primarily assess current fibrosis stage rather than future risk of progression[11]. As a result, patients at risk of rapid deterioration may be labeled as low risk at baseline, limiting opportunities for early interventions and increasing the risk of mortality and liver-related complications. Although various clinical and genetic factors are associated with disease severity, they do not reliably predict the rate of fibrosis progression[7,12].

Recent advances in machine learning (ML) have enabled integration of complex clinical, laboratory, and genomic information for improved characterization of MASLD. ML-based approaches have demonstrated improved performance over conventional fibrosis scores for fibrosis staging and disease severity prediction[13-17]. More recently, longitudinal ML models incorporating temporal electronic health record (EHR) information have been explored to predict future fibrosis progression[18,19]. However, most existing approaches model MASLD as a homogeneous condition, despite increasing evidence that patients exhibit heterogeneous clinical phenotypes and disease trajectories. Data-driven clustering studies using clinical, genomic, and multi-omic features have identified reproducible MASLD subtypes, including liver-specific (LS) and cardiometabolic (CM) phenotypes[20-26]. In our prior analysis of the Mayo Clinic Biobank (MCB)[27] and Tapestry Study[28] cohorts, we identified and replicated five clinically and genetically distinct MASLD subtypes using a different analytic sample and indicator set[22]. This study focused on subtype characterization rather than individual-level prediction of rapid fibrosis progression. Nevertheless, these phenotypic differences have rarely been incorporated into predictive frameworks for rapid fibrosis progression. Furthermore, limited interpretability of many ML models remains an important barrier to clinical implementation[8,13,15]. Although explainable artificial intelligence approaches such as Shapley Additive Explanations (SHAP) have improved model interpretation[29], their application in subgroup-specific MASLD progression prediction remains limited.

To address these limitations, we developed an interpretable partitioned framework for predicting rapid fibrosis progression in MASLD using two independent Mayo Clinic cohorts: MCB and the Tapestry Study. Patients were partitioned into clinically meaningful phenotypes, and rapid progression was defined using longitudinal age-adjusted FIB-4 trajectories[30] and International Classification of Diseases (ICD) diagnosis codes. We then developed subgroup-specific ML models using multimodal clinical, laboratory, and genetic risk scores and compared their performance with conventional population-level models. Finally, SHAP-based feature attribution[29] was used to derive the Partitioned Rapid Fibrosis Progression Score (pRFPS), an interpretable additive risk score for individualized prediction. This framework demonstrates the potential of subgroup-aware ML to improve rapid fibrosis progression prediction while maintaining clinical interpretability.

2. Methods

Our framework is motivated by four observations: (1) fibrosis progression in MASLD is heterogeneous, with a subset of patients experiencing rapid progression despite comparable baseline risk; (2) longitudinal EHR data capture complementary signals from clinical events and laboratory trajectories; (3) progression patterns differ across patient subgroups; and (4) interpretable models are needed for clinical translation. Guided by these observations, our objective is to identify and predict rapid fibrosis progression in MASLD using an interpretable framework.

2.1. *Problem Setup*

Let $D = \{(X_i, Y_i)\}_{i=1}^{N}$ represent a cohort of N patients. For patient i, X_i represents a longitudinal clinical record consisting of diagnosis events and laboratory measurements: $X_i = (C_i, L_i)$ where $C_i = [c_{i,1}, \dots, c_{i,m}]$ denotes the sequence of diagnosis code events and $L_i = \left[l_{i,1}, \dots, l_{i,n}\right]$ denotes longitudinal laboratory observations. Diagnosis events and laboratory measurements are represented by their corresponding timestamps t and clinical information. The objective is to learn a predictive function: $\phi: X \to Y$ where $Y_i \in \{0,1\}$ indicates whether patient i experiences rapid fibrosis progression. Rapid-progression thresholds were defined separately for each subgroup g to account for differences in progression timelines. Longitudinal timestamps were used to determine progression timing and category but were not included as predictive features. Continuous clinical variables were summarized as the mean of available measurements after MASLD diagnosis to accommodate irregular EHR measurement frequency while maximizing cohort size.

2.2. *Cohort Selection*

This study utilized data from two large-scale, retrospective, EHR-linked genomic cohorts at Mayo Clinic. The MCB cohort[27,31] (IRB #08-007049) and the Tapestry Study[28] (IRB #19-000001) both received institutional review board approval, and all participants provided written informed consent for the use of their data in genetic and health research. The Tapestry Study is a decentralized, clinical-grade genomic sequencing initiative that enrolled over 98,000 participants between July 2020 and May 2024[28,32]. Exome+® sequencing[33,34] was performed by Helix, Inc. (San Mateo, CA). CDC Tier 1 genomic results were returned to participants, and raw sequencing data were stored at Mayo Clinic for research use.

In this study, we identified 5,747 individuals with MASLD in the MCB and 7,480 individuals in the Tapestry cohort. MASLD cases were identified from EHR problem lists and billing records using ICD-9/10 codes (Supplemental Table 1). The EHR-derived MASLD cohort was supplemented with additional cases identified through abdominal imaging and detailed medical record review using a previously described approach by Ahmed et al.[35]. Inclusion required the presence of at least one cardiometabolic risk factor[36], including obesity, T2D or prediabetes (hemoglobin A1c [HbA1c] ≥5.7%), hypertension, elevated triglycerides ≥150 mg/dL, or low high-density lipoprotein (HDL) cholesterol (≤40 mg/dL in males and ≤50 mg/dL in females). Patients with competing etiologies of liver disease (viral hepatitis, alcohol-related liver disease etc.) were excluded. To ensure independent cohorts, any overlapping participants were removed from the Tapestry cohort.

2.3. *Data Extraction*

Clinical variables included demographic characteristics (age and sex), body mass index (BMI), and routine laboratory measurements including alanine aminotransferase (ALT), aspartate aminotransferase (AST), platelet count, albumin, alkaline phosphatase (ALP), HDL, low-density lipoprotein (LDL), triglycerides, HbA1c, and blood urea nitrogen (BUN). In addition, established non-invasive fibrosis indices, including FIB-4, the AST-to-platelet ratio index (APRI), and the NFS[37], were calculated from available laboratory measurements.

Comorbidities, including hypertension, dyslipidemia, T2D, chronic kidney disease (CKD), cardiovascular disease, depression, and liver-related outcomes, including advanced fibrosis/cirrhosis, HCC, and liver transplantation, were identified using ICD codes or clinical documentation (Supplemental Table 1). Variables were selected based on prior evidence linking them to MASLD subgroup heterogeneity, disease severity, and progression[22-24,38-40], as well as their availability in the dataset.

2.4. *Genetic Risk Score*

Our genetic analysis focused on five well-established MASLD-associated variants, including *PNPLA3* rs738409 C>G, *TM6SF2* rs58542926 C>T, and *HSD17B13* rs13118664, which have been validated across multiple cohorts[21,23,41-43]. Two genetic risk scores were constructed following previously published approaches: an unweighted 3-single-nucleotide polymorphism (3-SNP) polygenic risk score (PRS-3)[43] and a weighted 5-SNP MASLD polygenic risk score (PRS-5)[41]. PRS-3 summed risk-allele dosages across PNPLA3 (GG=2, CG=1, CC=0), TM6SF2 (TT=2, CT=1, CC=0), and HSD17B13 (non-A/non-A=2, A/non-A=1, A/A=0), yielding scores from 0 to 6. PRS-5 was calculated using published variant weights from the Polygenic Score Catalog (PGS000872) with pgsc_calc[44,45] and categorized as low (<40th percentile), intermediate (40th-89th percentile), and high (≥90th percentile) genetic risk groups, as previously described[22].

2.5. *Model Architecture*

Fig. 1 illustrates the overall framework, which consists of mainly three components: (i) phenotypic partitioning of MASLD patients, (ii) derivation of progression categories using longitudinal data, and (iii) subgroup-specific ML models with an interpretable rapid fibrosis progression score.

2.5.1. *Phenotypic Partitioning*

To incorporate MASLD heterogeneity into predictive modeling, patients were partitioned into clinically distinct subgroups $g \in G$ using latent class analysis (LCA). Phenotypic partitioning was performed among 2,422 MCB and 3,093 Tapestry patients with complete clinical variables (Supplemental Tables 2-3)[22,23,26]. In this work, LCA was performed using six routinely available clinical variables (ALT, age, HbA1c, BMI, triglycerides, and LDL) previously used for MASLD subgrouping[23]. Continuous variables were discretized according to clinical thresholds[22], and each patient was assigned to the subgroup with the highest posterior probability. LCA-based subgroup

assignment and subgroup-specific progression thresholds were established using the full eligible MCB cohort before applying the predictive-model eligibility criteria.

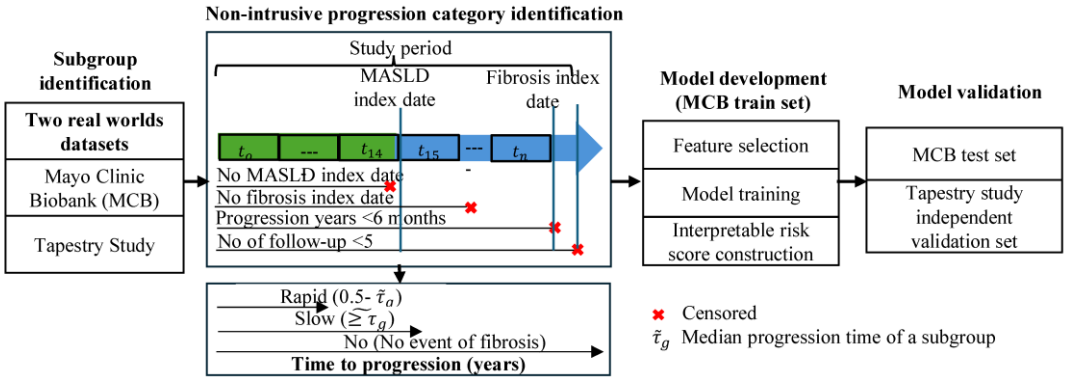


Fig. 1. Overview of the partitioned framework for rapid fibrosis-progression prediction.

2.5.2. *Non-intrusive Progression Category Identification from Longitudinal Data*

Fibrosis progression category was defined using a hybrid strategy integrating ICD-based diagnoses and longitudinal laboratory trajectories. Age, AST, ALT, and platelet count were used to calculate longitudinal FIB-4 values (Eq. 1), which were categorized as low, intermediate, or high fibrosis risk stages, denoted $S_i(t)$ using age-adjusted thresholds:(<65 years: <1.3, 1.3-2.67, >2.67; ≥65 years: <2.0, 2.0-2.67, >2.67)[46,47].

$$\text{FIB-4} = \frac{\text{Age (years)} \times \text{AST (U/L)}}{\text{Platelet count } (10^9/\text{L}) \times \sqrt{\text{ALT (U/L)}}} \tag{1}$$

For each patient i, the MASLD index date $t_i^{(0)}$was defined as the earliest occurrence of a MASLD diagnosis, a metabolic dysfunction-associated steatohepatitis (MASH) diagnosis, or a metabolic proxy condition (T2D, obesity, dyslipidemia, or hypertension)[8]. The first ICD-based fibrosis or cirrhosis diagnosis during follow-up was denoted as $t_i^{(fib)}$. For patients with ICD-confirmed fibrosis progression, progression time was calculated as:

$$\Delta_{\text{ICD},i} = t_i^{(\text{fib})} - t_i^{(0)}, \ \ \Delta_{\text{ICD},i} \geq 6 \ months \tag{2}$$

For patients without sufficient ICD-based temporal separation i.e., $\Delta_{ICD,i} < 6$ months, progression time was estimated from the interval between the first observed FIB-4 stage and the subsequent recorded stage transition: $\Delta_{\text{FIB4},i} = t_i\left(S_i^{(\text{last})}\right) - t_i\left(S_i^{(\text{first})}\right)$. When no FIB-4 stage transition was observed progression time was instead defined as the interval from the MASLD index date to the first transition into an intermediate or high-risk FIB-4 category:

$$\Delta_{stage,i} = t_i^{(k)} - t_i^{(0)}, \text{ where } S_i\left(t_i^{(k)}\right) \in \{\text{Intermediate,High}\} \tag{3}$$

To account for heterogeneity in progression patterns, subgroup-specific thresholds were determined using the median progression time among patients with ICD-confirmed progression in the MCB cohort (n = 2,422), thereby providing clinically documented reference timelines for each subgroup: $\tilde{\tau}_g = \text{median}\{\Delta_{ICD,i} : i \in g\}$. The binary outcome was then defined as:

$$Y_i = \begin{cases} 1, & \text{if } \Delta_i \leq \tilde{\tau}_{g_i} \text{ (Rapid Progression)} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

These MCB-derived subgroup-specific thresholds were applied to classify rapid progression in the Tapestry Study validation cohort. Patients without evidence of progression but with adequate longitudinal follow-up (≥5 observations within a consistent FIB-4 risk category) were classified as slow progressors. Patients lacking sufficient evidence for progression assessment were excluded from downstream analyses.

2.5.3. *Partitioned Modeling Framework*

To predict rapid fibrosis progression, we developed a partitioned modeling framework. Within each subgroup, feature selection strategies were applied to identify the most informative predictors from a multimodal feature space. Multiple supervised ML models, including linear, tree-based, instance-based, kernel-based, ensemble, boosting, and neural network approaches were evaluated to capture both linear and nonlinear relationships among predictors.

To obtain clinically interpretable prediction scores while preserving the structure learned by subgroup-specific ML models, we constructed subgroup-wise scoring systems using SHAP values[29]. Feature importance was quantified using the mean absolute SHAP value irrespective of direction and the top five predictors were selected from the best-performing model. SHAP values were calculated using LinearExplainer for linear models, TreeExplainer for tree-based models, and KernelExplainer for non-tree models with a reduced background set[29,48]. Selected predictors were converted into an additive scoring system by normalizing their SHAP contributions and assigning integer weights from 0 to 10 according to: $\text{Points}_j = \text{round}\left(\frac{\text{mean}(|\text{SHAP}_j|)}{\max_j \text{mean}(|\text{SHAP}_j|)} \times 10\right)$. For each feature, optimal thresholds were determined using the Youden Index[49], $J = \text{Sensitivity} + \text{Specificity} - 1$. This yielded a subgroup-specific additive scoring function: $pRFPS_g(x_i) = \sum_{j \in F_g} w_{j,g}\,(I_{j,g}(x_{i,j}))$, where F_g denotes the selected feature set for subgroup g, $w_{j,g}$ denotes the assigned feature weight, and $I_{j,g}(x_{i,j}) \in \{0,1\}$ denotes the subgroup-specific threshold-based indicator function for feature j. We termed this score pRFPS, which provides a transparent and clinically actionable tool for identifying patients at high risk of rapid fibrosis progression.

2.6. *Experimental Setup*

This section describes learning model implementation, feature selection strategies, and performance evaluation protocols.

2.6.1. *Implementation Details*

We evaluated a comprehensive suite of supervised machine learning algorithms including Logistic Regression (LR), k-Nearest Neighbors (kNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), eXtreme Gradient Boosting (XGB), Light Gradient Boosting (LGBM) and Multi-Layer Perceptron (MLP). These models were selected based on their widespread use in clinical risk prediction tasks[8,17,50,51]. The MLP neural-network model was included to evaluate potentially complex nonlinear predictor–outcome relationships and higher-

order interactions. After excluding patients with insufficient longitudinal data or missing variables (Supplemental Table 2), the final cohorts included 752 MCB and 1,240 Tapestry participants. The MCB cohort served as the primary development cohort and was randomly divided into an 80:20 training-test split. The training set was used for model development, including feature selection, hyperparameter optimization, and pRFPS construction, while the held-out test set was reserved for evaluation. The Tapestry Study cohort was used as an independent validation cohort. Feature selection, hyperparameter optimization, training, and threshold selection were performed globally and within MASLD subgroup to capture heterogeneity in predictive determinants.

To identify informative predictors, we applied a multi-method feature selection framework including ANOVA F-test, mutual information, recursive feature elimination, Random Forest importance, LassoCV, and SHAP-based importance using only the training data[52-55]. Predictors identified by at least three methods were retained for model development. Hyperparameter optimization was performed exclusively on the MCB training set using exhaustive grid search with stratified 5-fold cross-validation. F1-score was used as the optimization criterion to account for class imbalance, and the hyperparameter configuration achieving the best cross-validated performance was selected. Details of the hyperparameter search space are provided in Supplemental Table 4. Classification thresholds were subsequently optimized separately for each model on the MCB training set by maximizing the F1-score. The resulting fixed thresholds were subsequently applied to the held-out MCB test set and the Tapestry Study validation cohort.

2.6.2. *Performance Evaluation Metrics*

The optimal number of latent classes was determined by evaluating LCA models with 1–10 classes. Model selection was based on Bayesian Information Criterion (BIC), entropy, average posterior class membership probability (≥0.5), and a minimum class size threshold of 5% of the cohort[56,57]. Models with a class size <5% or mean posterior membership probability <0.5 were excluded. Among the remaining models, entropy was the primary criterion for class separation, while BIC and mean posterior membership probability were used to assess model fit and assignment confidence. Local independence was assessed using bivariate residuals. Although additional classes reduced residual dependence, they produced smaller and less clinically interpretable classes[58]. Therefore, final model selection balanced statistical fit, entropy, adequate class size, and clinically meaningful phenotype separation, with a relaxation of strict local independence to accommodate biologically correlated indicators[57]. Clinical and demographic characteristics across phenotypes and progression categories were compared using the Wilcoxon rank-sum test for continuous variables and the chi-square test or Fisher's exact test for categorical variables, as appropriate. Reported two-sided *p*-values were unadjusted, with $p < 0.05$ considered statistically significant.

Predictive performance was evaluated using F1 and AUROC. Accuracy, precision, recall, AUPRC, and Brier score were also reported. 95% confidence intervals (CI) were estimated using 1,000 stratified bootstrap resamples. Sensitivity analyses evaluated with a fixed 3-year definition of rapid progression and a min–max normalized continuous pRFPS[#] using the same SHAP-derived predictors and weights. To quantify the contribution of subgroup-specific modeling, we performed

[#] Details are provided in Supplementary Materials (https://drive.google.com/PSB 2027)

an ablation analysis comparing four configurations: (1) the default baseline, defined as the best-performing overall model trained on the full MCB training cohort using all candidate features and evaluated within each subgroup; and the best-performing subgroup-specific models using (2) all candidate features, (3) overall selected features, or (4) subgroup-specific selected features.

In addition, we assessed the interpretability and clinical utility of the proposed pRFPS by benchmarking it against FIB-4. Although FIB-4 is designed to stage fibrosis[40] rather than predict progression, it is commonly used in clinical practice as a risk screening tool, where lower scores indicate lower risk of advanced fibrosis and higher scores indicate increased risk requiring closer monitoring[59]. To enable a consistent comparison, patients with baseline low and high age-adjusted FIB-4 risk were classified as slow and rapid progressors, respectively. This analysis was restricted to clearly defined risk strata to ensure higher confidence in risk stratification. Comparative performance was assessed by comparing AUROCs using the DeLong test[60].

3. Results

In this section, we present a comprehensive evaluation of our partitioned modeling framework for predicting rapid fibrosis progression in MASLD.

Table 1. Progression time and sources in the final predictive cohorts.

Dataset	Subgroup	N	Median progression (years)	Progression (%)	
				ICD based	FIB-4 based
MCB	Overall	752	4.75	25.3	74.7
	LS	427	4.55	23	77
	CM	325	4.75	28.3	71.7
Tapestry	Overall	1240	3.1	23.1	76.9
	LS	831	2.9	21.8	78.2
	CM	409	4.15	25.9	74.1

3.1. *Subgroup Characteristics*

The two-class solution was selected as optimal in the MCB cohort, achieving the highest entropy (0.749), highest mean posterior membership probability (0.907), second-lowest BIC (25,366.04) and ~70% of patients assigned with probability ≥0.90 (Supplemental Tables 5–6). Based on their clinical characteristics, the identified classes corresponded to previously described liver-specific and cardiometabolic specific MASLD phenotypes[22,23]. LS exhibited lower BMI (31.3 vs. 34.5 kg/m²), better glycemic control (HbA1c: 5.51 vs. 7.18%), lower triglyceride levels (128 vs 189 mg/dL), higher LDL cholesterol (93.3 vs 80.5 mg/dL), and lower diabetes prevalence than CM (all $p < 0.001$), whereas CM showed a greater cardiometabolic burden (4 vs. 2.6 in LS; $p < 0.001$). These phenotypes were consistently reproduced in the independent Tapestry cohort (Supplemental Tables 7-8).

3.2. *Fibrosis Progression Category Across Subgroups*

After applying progression assessment criteria, median progression times among patients included in ML model development and evaluation were comparable between LS and CM in the MCB cohort, while LS showed faster progression in the Tapestry cohort (Table 1). Longitudinal FIB-4 trajectories identified substantially more progression events (>70%) than ICD-based definitions alone. After integrating all progression criteria, more than 50% of patients in both subgroups were classified as slow progressors with low baseline FIB-4 risk (Fig. 2). Notably, 37.5% (n=160) of LS and 38.8% (n=126) of CM patients were rapid progressors. Despite low baseline FIB-4 risk, 24.6% of LS and 23.7% of CM patients were classified as rapid progressors, including 17.8% and 15.1%,

respectively, identified by longitudinal FIB-4 progression. Conversely, 7.3% of patients with high FIB-4 showed slow progression.

Within each subgroup, rapid progressors demonstrated distinct clinical profiles compared with slow progressors (Table 2 and Supplemental Tables 9-10). Across both subgroups, rapid progressors were younger and showed evidence of greater hepatic injury, including higher AST, ALT, and ALP levels, lower platelet counts, and higher MASH prevalence (all $p < 0.05$), suggesting a consistent association between active liver injury and accelerated fibrosis progression. In LS, rapid progression was additionally associated with male sex, higher T2D prevalence, the *PNPLA3* G/G genotype, and higher PRS-5, suggesting a stronger genetic contribution. In CM, where diabetes was highly prevalent across patients along with lower albumin levels and higher CKD burden. No significant differences in *PNPLA3*, genetic risk score, or sex distribution were observed between progression groups in CM, highlighting the heterogeneous factors associated with rapid fibrosis progression across MASLD phenotypes.

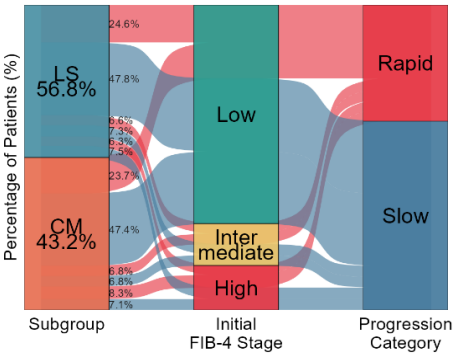


Fig. 2. Alluvial plot of MASLD phenotypes, initial FIB-4 risk, and progression categories in MCB cohort.

Table 2. Characteristics of rapid vs. slow progressors in MCB cohort.

Characteristic	LS		*p*	CM		*p*
	Rapid	Slow		Rapid	Slow	
Age	57.3 (12.6)	59.8 (14.5)	0.021	60.3 (10.8)	62.8 (12.8)	0.024
Female	73 (46%)	166 (62%)	0.001	63 (50%)	101 (51%)	>0.9
ALT	47.0 (41.3)	39.8 (35.9)	0.023	51.9 (46.6)	46.6 (58.6)	0.023
AST	54.3 (57.1)	38.1 (25.8)	<0.001	62.8 (91.7)	44.6 (51.2)	<0.001
ALP	115.6 (65.1)	93.8 (44.0)	<0.001	128.7 (96.4)	102.4 (54.5)	<0.001
Platelet Count	179.6 (88.9)	234.8 (91.7)	<0.001	162.1 (89.2)	250.7 (112.3)	<0.001
Albumin	6.2 (13.2)	8.4 (32.5)	0.12	8.7 (18.7)	19.9 (51.7)	<0.001
Triglycerides	121.6 (64.9)	129.0 (63.2)	0.03	169.6 (95.5)	185.1 (76.4)	0.019
BUN	21.3 (10.3)	16.8 (6.0)	<0.001	22.7 (10.0)	20.2 (9.0)	0.021
Diabetes	67 (42%)	80 (30%)	0.016	120 (95%)	177 (89%)	0.077
MASH	60 (38%)	62 (23%)	0.002	77 (61%)	51 (26%)	<0.001
CKD	81 (51%)	119 (45%)	0.3	79 (63%)	96 (48%)	0.015
Depression	63 (39%)	145 (54%)	0.004	64 (51%)	103 (52%)	>0.9
PNPLA3 (G/G)	33 (21%)	16 (6%)	<0.001	23 (18%)	24 (12%)	0.3
PRS-5	0.2 (0.4)	0.1 (0.3)	0.003	0.2 (0.3)	0.1 (0.3)	0.3

Continuous variables are presented as mean (SD) and categorical variables as n (%). Unadjusted *p*-values were calculated using Wilcoxon rank-sum, chi-square, or Fisher's exact tests, as appropriate.

3.3. *Subgroup-specific Prediction Performance*

Table 3 summarizes the performance of models selected based on the MCB test set and their

Table 3. Performance of the selected model in MCB test set and Tapestry Study validation set.

Setting	Model	MCB test set		Tapestry Study validation set	
		F1 [95% CI]	AUROC [95% CI]	F1 [95% CI]	AUROC [95% CI]
Overall	GB	0.700 [0.625–0.771]	0.749 [0.663–0.822]	0.737 [0.713-0.760]	0.812 [0.786-0.836]
LS	RF	0.720 [0.622–0.811]	0.809 [0.718–0.894]	0.721 [0.687-0.750]	0.808 [0.774-0.837]
CM	MLP	0.834 [0.740–0.923]	0.844 [0.721–0.949]	0.739 [0.694-0.785]	0.798 [0.748-0.839]

LS = Liver-specific; CM = Cardiometabolic; GB = Gradient Boosting; RF = Random Forest; MLP = Multi-Layer Perceptron; CI = Confidence Interval

corresponding performance in the Tapestry validation cohort, with complete results provided in Supplemental Tables 11-12. In the MCB test set, the best performing models achieved AUROCs of 0.809 for LS (RF) and 0.844 for CM (MLP), with F1-scores of 0.720 and 0.834, respectively, outperforming the best overall model (GB; AUROC: 0.749, F1-score: 0.70). In the Tapestry validation cohort, these models achieved AUROCs of 0.808 in LS and 0.798 in CM and F1-scores of 0.721 and 0.739, respectively, demonstrating consistent performance, although the overall model achieved the highest AUROC. Ablation analysis further demonstrated the contribution of the partitioned framework (Table 4). Compared with the baseline model, partitioned modeling with all features increased the F1-score from 0.48 to 0.62 in LS and 0.57 to 0.62 in CM, with corresponding AUROCs increasing from 0.46 to 0.84 and from 0.45 to 0.82, respectively. Subgroup-adapted feature selection further increased performance, achieving F1-scores of 0.72 (LS) and 0.83 (CM) with AUROCs of 0.81 and 0.84, respectively. These findings demonstrated that subgroup-aware modeling improved prediction of rapid fibrosis progression.

Table 4. Ablation analysis of subgroup-specific modeling and feature selection for rapid fibrosis progression prediction in the MCB test set.

Subgroup	Configuration	F1	AUROC
LS	Baseline[a]	0.48	0.46
	All features	0.62	0.84
	Overall selected features[b]	0.67	0.79
	Subgroup selected features[c]	**0.72**	**0.81**
CM	Baseline[a]	0.57	0.45
	All features	0.62	0.82
	Overall selected features[b]	0.62	0.79
	Subgroup selected features[d]	**0.83**	**0.84**

[a]Best-performing overall-cohort model, evaluated within each subgroup; [b]ALP, Albumin, AST, APRI, BMI, BUN, LDL, platelet count, triglycerides, age; [c]ALP, AST, APRI, BMI, BUN, HDL, LDL, platelet count, sex, age; [d]ALP, Albumin, AST, APRI, BUN, diabetes, HDL, LDL, platelet count, age.

3.4. *Subgroup-specific Interpretable Partitioned Rapid Fibrosis Progression Score*

SHAP analysis of the best-performing models identified distinct predictor profiles across MASLD phenotypes. The derived pRFPS (Eqs. 5 and 6) incorporated the most influential predictors from each phenotype-specific model, with LS weighted by age, BUN, ALP, platelet count, and HDL, and CM weighted by ALP, platelet count, LDL, albumin, and AST. As summarized in Table 5, pRFPS outperformed FIB-4 for identifying rapid fibrosis progression in both the MCB test set and Tapestry validation cohort. In LS, pRFPS showed greater discrimination than FIB-4 in the MCB test set (ΔAUROC=0.23, p = 0.028) and Tapestry cohort (ΔAUROC=0.11, $p < 0.001$), with modest improvements observed in CM. Using the phenotype-specific pRFPS cutoffs, high-risk individuals had increased odds of HCC (LS: OR=1.95, 95% CI: 0.99-3.85; CM: OR=4.84, 95% CI: 2.17-10.76;) and liver transplantation (LS: OR=2.87, 95% CI: 1.70-4.83; CM: OR=3.52, 95% CI: 1.91-6.48) compared with low-risk individuals. In contrast, advanced baseline fibrosis showed no consistent association with HCC or liver transplantation across the two phenotypes (Supplemental Table 13). Together, these findings demonstrated that pRFPS provided phenotype-specific risk stratification for rapid fibrosis progression and identified patients at increased risk of adverse liver-related outcomes.

Table 5. Classification performance of pRFPS and FIB-4 for rapid fibrosis progression prediction

Subgroup	pRFPS Cutoff	Test Set	AUROC			p
			pRFPS	FIB-4	Δ	
LS	15	MCB	**0.68**	0.45	0.23	0.028
		Tapestry	**0.71**	0.59	0.11	<0.001
CM	18	MCB	**0.74**	0.69	0.05	0.638
		Tapestry	**0.72**	0.64	0.08	0.085

$$pRFPS_{LS} = 10 * I(\text{Age} > 45) + 5 * I(\text{BUN} > 21) + 5 * I(\text{ALP} > 109) + 3 * I(\text{Platelet Count} < 125) + 3 * I(\text{HDL} < 61) \quad (5)$$

$$pRFPS_{CM} = 10 * I(\text{ALP} > 97) + 6 * I(\text{Platelet Count} < 313) + 1 * I(\text{LDL} > 64) + 1 * I(\text{Albumin} < 13) + 1 * I(\text{AST} > 39) \quad (6)$$

4. Discussion

In this study, we developed a partitioned modeling framework to predict rapid fibrosis progression in patients with MASLD by integrating longitudinal data with subgroup-specific disease patterns. Because fibrosis progression is variable and advanced fibrosis carries substantial morbidity[6,61,62], early identification of individuals at increased risk is essential for timely intervention[19]. Unlike conventional approaches that primarily assess current fibrosis severity[11], our framework leverages longitudinal laboratory trends and diagnosis histories to capture dynamic changes associated with future disease progression. By incorporating MASLD phenotypes and developing subgroup-specific predictive models, we account for disease heterogeneity beyond a single population-level approach. The pRFPS further translates these predictions into an interpretable tool for individualized risk stratification and intervention.

Our findings identified clinically distinct MASLD phenotype patterns consistent with previous studies, supporting MASLD as a heterogeneous disease spectrum with distinct contributions from hepatic injury, metabolic dysfunction, and comorbidities[20,23,24,26]. The liver-specific subgroup was characterized by hepatic injury markers, genetic susceptibility (high PRS-5: 17.7% vs. 15.7%), and lipid-related features consistent with altered hepatic lipid regulation[24] (high LDL), whereas the cardiometabolic subgroup showed stronger contributions from metabolic and systemic risk profiles. These findings showed the importance of phenotype-aware prediction and risk stratification[23,63].

Accounting for clinical heterogeneity through partitioned modeling improved prediction of rapid fibrosis progression compared with globally trained models across most evaluated models. In the MCB test set, the best subgroup-specific models improved F1-scores by up to 13% points compared with the best full-cohort model and retained good performance in the Tapestry Study validation cohort. Clinical characterization of rapid progressors identified prominent hepatic injury and fibrosis-related features, including ALT, AST, ALP, and platelet count, consistent with established fibrosis-risk markers.[11,18]. Metabolic and systemic characteristics, including BMI, diabetes, and lipid parameters, further differentiated progression patterns, reflecting the contribution of metabolic dysfunction and extrahepatic burden to MASLD progression[19,64,65]. Genetic characterization revealed higher PRS values among rapid progressors[43,66], particularly in LS with enrichment of the *PNPLA3* GG genotype. The protective *HSD17B13* variant was more frequent among slow progressors (6.4%–9.0%), particularly in CM. Ablation analysis further showed that subgroup-specific feature selection improved performance beyond stratified training alone, supporting phenotype-aware modeling of heterogeneous disease progression.

Beyond predictive performance, SHAP-based feature attribution enabled translation of complex model outputs into the interpretable, subgroup-specific pRFPS. The score can be calculated from routinely available clinical variables and was applicable to 2,272 MCB participants, including 1,520 excluded from model development because of insufficient follow-up, suggesting potential utility for earlier risk stratification. Its associations with HCC and liver transplantation further support its potential clinical relevance. Direct comparison with previously published MASLD prediction

models[8,18,19] was not performed because existing studies differ substantially in cohort characteristics, outcome definitions, and available feature sets.

This study has several limitations and opportunities for future improvement. First, rapid fibrosis progression was defined using longitudinal FIB-4 trajectories and ICD-based clinical events rather than biopsy-confirmed fibrosis progression. A sensitivity analysis using a fixed 3-year progression window yielded fewer rapid progressors, greater class imbalance, and lower predictive performance than the primary subgroup-specific definition (Supplemental Table 14), highlighting the trade-off between fixed time windows and adequate outcome representation for model development. Second, routinely collected EHR data introduces variability in follow-up duration, missingness, and documentation, and the recorded MASLD diagnosis date may not reflect true disease onset. Finally, pRFPS uses integer weights and binary thresholds to provide a simple additive representation of subgroup-specific predictors. This approach may not retain all information, particularly near feature cutoffs, and the selected thresholds may be sample-dependent. However, in a sensitivity analysis, a continuous min–max-normalized score without feature thresholding, using the same predictors and weights, showed no consistent advantage: AUROC was lower in LS (0.633 vs. 0.705) but higher in CM (0.785 vs. 0.703) (Supplemental Table 15). Although an independent non-overlapping validation cohort was included, both cohorts were derived from the same healthcare system. Further cross-institutional, and cross-population validation is needed to assess the stability and generalizability of the phenotype-specific framework and refine the pRFPS.

Future work will focus on incorporating additional data modalities, including imaging-derived biomarkers, more comprehensive genomic risk scores, and additional clinical features, while evaluating performance against serial biopsy-defined fibrosis progression, histological outcomes, and long-term liver-related endpoints to define clinically grounded rapid, slow, moderate, and non-progression trajectories. Rapid progression was less frequent in both cohorts, representing a modest class imbalance. Although F1-score was used for model and threshold selection, additional sampling strategies warrant evaluation. Sex- and age-stratified analyses showed comparable performance within subgroup-specific models (Supplemental Table 16). However, further validation in other underrepresented populations remains needed. As this was a retrospective study, prospective validation across diverse populations will be essential to establish the clinical utility of pRFPS.

In conclusion, this study provides a phenotype-aware predictive framework for identifying patients at increased risk of rapid fibrosis progression in MASLD. Beyond prediction, pRFPS may support personalized surveillance, earlier risk-based intervention, and improved identification of patients for future clinical trials. As MASLD management moves toward precision medicine and emerging therapies[67] target distinct metabolic and fibrotic pathways, improved risk stratification may enable more targeted patient selection and clinical decision-making.

Acknowledgements

The Mayo Clinic Biobank was supported in part by the Mayo Clinic Center for Individualized Medicine, with genetic data provided by the Regeneron Genetics Center. Tapestry Study genomic data were supported by the Mayo Clinic Center for Individualized Medicine. We thank the Mayo Clinic Biobank and Tapestry Study research teams and all participants for their contributions.

Data and Code Availability

Individual-level data are not publicly available because of participant privacy and institutional data-use restrictions. Access to MCB data can be requested by contacting Biobank@mayo.edu. Tapestry Study data were used under license, and reasonable requests for access from qualified researchers may be directed to Dr. Konstantinos Lazaridis, subject to institutional approvals and data-use agreements with Helix, Inc. In accordance with informed consent provisions, projects using these data must include a Mayo Clinic investigator as a collaborator. The code used for this study is available at: https://github.com/tpriya14/MASLD-pRFPS. Conflict-of-interest disclosures and Supplemental materials are available at: https://drive.google.com/PSB 2027.

References

1. Byrne, C.D. & Targher, G. NAFLD: a multisystem disease. *J Hepatol* **62**, S47-64 (2015).
2. Stefan, N. & Cusi, K. A global view of the interplay between non-alcoholic fatty liver disease and diabetes. *Lancet Diabetes Endocrinol* **10**, 284-296 (2022).
3. Koliaki, C., Dalamaga, M., Kakounis, K. & Liatis, S. Metabolically Healthy Obesity and Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD): Navigating the Controversies in Disease Development and Progression. *Curr Obes Rep* **14**, 46 (2025).
4. Younossi, Z.M., Kalligeros, M. & Henry, L. Epidemiology of metabolic dysfunction-associated steatotic liver disease. *Clin Mol Hepatol* **31**, S32-s50 (2025).
5. Le, P., *et al.* Estimated Burden of Metabolic Dysfunction-Associated Steatotic Liver Disease in US Adults, 2020 to 2050. *JAMA Netw Open* **8**, e2454707 (2025).
6. Lekakis, V. & Papatheodoridis, G.V. Natural history of metabolic dysfunction-associated steatotic liver disease. *Eur J Intern Med* **122**, 3-10 (2024).
7. Singh, S., *et al.* Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. *Clin Gastroenterol Hepatol* **13**, 643-654 e641-649; quiz e639-640 (2015).
8. Schattenberg, J.M., *et al.* Identification of Fast Progressors Among Patients With Nonalcoholic Steatohepatitis Using Machine Learning. *Gastro Hep Adv* **3**, 101-108 (2024).
9. Sanyal, A.J., *et al.* The Natural History of Advanced Fibrosis Due to Nonalcoholic Steatohepatitis: Data From the Simtuzumab Trials. *Hepatology* **70**, 1913-1927 (2019).
10. Pandey, N., Hoilat, G.J. & John, S. Liver Biopsy. in *StatPearls* (Treasure Island (FL), 2025).
11. Kaswala, D.H., Lai, M. & Afdhal, N.H. Fibrosis Assessment in Nonalcoholic Fatty Liver Disease (NAFLD) in 2016. *Dig Dis Sci* **61**, 1356-1364 (2016).
12. Jonas, W. & Schurmann, A. Genetic and epigenetic factors determining NAFLD risk. *Mol Metab* **50**, 101111 (2021).
13. Canbay, A., *et al.* Non-invasive assessment of NAFLD as systemic disease-A machine learning perspective. *PLoS One* **14**, e0214436 (2019).
14. Docherty, M., *et al.* Development of a novel machine learning model to predict presence of nonalcoholic steatohepatitis. *J Am Med Inform Assoc* **28**, 1235-1241 (2021).
15. Feng, G., *et al.* Machine learning algorithm outperforms fibrosis markers in predicting significant fibrosis in biopsy-confirmed NAFLD. *J Hepatobiliary Pancreat Sci* **28**, 593-603 (2021).
16. Eslam, M., *et al.* FibroGENE: A gene-based model for staging liver fibrosis. *J Hepatol* **64**, 390-398 (2016).

17. Alkhouri, N.*, et al.* ALADDIN: A Machine Learning Approach to Enhance the Prediction of Significant Fibrosis or Higher in Metabolic Dysfunction-Associated Steatotic Liver Disease. *Am J Gastroenterol* **121**, 362-374 (2026).
18. Ghandian, S.*, et al.* Machine learning to predict progression of non-alcoholic fatty liver to non-alcoholic steatohepatitis or fibrosis. *JGH Open* **6**, 196-204 (2022).
19. Dai, R., Sun, M., Lu, M. & Deng, L. Deep learning for predicting fibrotic progression risk in diabetic individuals with metabolic dysfunction-associated steatotic liver disease initially free of hepatic fibrosis. *Heliyon* **10**, e34150 (2024).
20. Hong, C.*, et al.* Novel subtypes of metabolic associated steatotic liver disease linked to clinical outcomes: implications for precision medicine. *J Transl Med* **23**, 769 (2025).
21. Priya, T.S.*, et al.* Deep Phenotyping of Non-Alcoholic Fatty Liver Disease Patients with Genetic Factors for Insights into the Complex Disease. *Machine Learning for Health (ML4H) Symposium, Findings Track* (2023).
22. Priya, T.S.*, et al.* Subtyping metabolic dysfunction-associated steatotic liver disease using electronic health record-linked genomic cohorts reveals diverse etiologies and progression. *Nature Communications* (2026).
23. Raverdy, V.*, et al.* Data-driven cluster analysis identifies distinct types of metabolic dysfunction-associated steatotic liver disease. *Nat Med* **30**, 3624-3633 (2024).
24. Jamialahmadi, O.*, et al.* Partitioned polygenic risk scores identify distinct types of metabolic dysfunction-associated steatotic liver disease. *Nat Med* **30**, 3614-3623 (2024).
25. Ding, J.*, et al.* Integrative multiomic analysis identifies distinct molecular subtypes of NAFLD in a Chinese population. *Sci Transl Med* **16**, eadh9940 (2024).
26. Zhou, X.D.*, et al.* Validation of a data-driven clustering model for MASLD: Evidence from three large-scale Asian cohorts. *JHEP Rep* **8**, 101645 (2026).
27. Olson, J.E.*, et al.* Characteristics and utilisation of the Mayo Clinic Biobank, a clinic-based prospective collection in the USA: cohort profile. *BMJ Open* **9**, e032707 (2019).
28. Bandel, L.A.*, et al.* Mayo Clinic Tapestry Study: A Large-Scale Decentralized Whole Exome Sequencing Study for Clinical Practice, Research Discovery, and Genomic Education. *Mayo Clin Proc* (2024).
29. Lundberg, S.M. & Lee, S.-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems* **30**(2017).
30. Ishiba, H.*, et al.* The novel cutoff points for the FIB4 index categorized by age increase the diagnostic accuracy in NAFLD: a multi-center study. *J Gastroenterol* **53**, 1216-1224 (2018).
31. Hathcock, M.A.*, et al.* Characteristics Associated With Recruitment and Re-contact in Mayo Clinic Biobank. *Front Public Health* **8**, 9 (2020).
32. Bharucha, A.E.*, et al.* Factors Associated With Enrollment to a Decentralized Study. *Mayo Clin Proc* **100**, 52-67 (2025).
33. Helix, I. Performance and Applications of Helix's Exome+® Assay. (2026).
34. Grzymski, J.J.*, et al.* Population genetic screening efficiently identifies carriers of autosomal dominant diseases. *Nat Med* **26**, 1235-1239 (2020).
35. Ahmed, O.T.*, et al.* Natural History of Nonalcoholic Fatty Liver Disease With Normal Body Mass Index: A Population-Based Study. *Clin Gastroenterol Hepatol* **20**, 1374-1381.e1376 (2022).
36. Rinella, M.E.*, et al.* A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology* **78**, 1966-1986 (2023).

37. Ruffillo, G.*, et al.* Comparison of NAFLD fibrosis score and BARD score in predicting fibrosis in nonalcoholic fatty liver disease. *J Hepatol* **54**, 160-163 (2011).
38. Su, Y.*, et al.* Major Depression as a Catalyst for Nonalcoholic Fatty Liver Disease: Insights From a Comprehensive Mendelian Randomization Study. *Can J Gastroenterol Hepatol* **2025**, 6632867 (2025).
39. Qiao, Y., Wang, Y., Chen, C., Huang, Y. & Zhao, C. Association between triglyceride-glucose (TyG) related indices and cardiovascular diseases and mortality among individuals with metabolic dysfunction-associated steatotic liver disease: a cohort study of UK Biobank. *Cardiovasc Diabetol* **24**, 12 (2025).
40. 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* **64**, 1388-1402 (2016).
41. Bianco, C.*, et al.* Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. *J Hepatol* **74**, 775-782 (2021).
42. Krawczyk, M., Liebe, R. & Lammert, F. Toward Genetic Prediction of Nonalcoholic Fatty Liver Disease Trajectories: PNPLA3 and Beyond. *Gastroenterology* **158**, 1865-1880 e1861 (2020).
43. Vilar-Gomez, E.*, et al.* Genetic and non-genetic drivers of histological progression and regression in MASLD. *J Hepatol* **84**, 502-516 (2026).
44. Lambert, S.A.*, et al.* Enhancing the Polygenic Score Catalog with tools for score calculation and ancestry normalization. *Nat Genet* **56**, 1989-1994 (2024).
45. Lambert, S.A.*, et al.* The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat Genet* **53**, 420-425 (2021).
46. Rinella, M.E.*, et al.* AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* **77**, 1797-1835 (2023).
47. Sung, S.*, et al.* Age-dependent differences in FIB-4 predictions of fibrosis in patients with MASLD referred from primary care. *Hepatol Commun* **9**(2025).
48. Lundberg, S.M., Erion, G.G. & Lee, S.-I. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888* (2018).
49. Martínez-Camblor, P. & Pardo-Fernández, J.C. The Youden index in the generalized receiver operating characteristic curve context. *The international journal of biostatistics* **15**, 20180060 (2019).
50. Tanabe, H.*, et al.* Machine learning-based reproducible prediction of type 2 diabetes subtypes. *Diabetologia* (2024).
51. Mani, S., Chen, Y., Elasy, T., Clayton, W. & Denny, J. Type 2 diabetes risk forecasting from EMR data using machine learning. *AMIA Annu Symp Proc* **2012**, 606-615 (2012).
52. Grissa, D.*, et al.* Feature Selection Methods for Early Predictive Biomarker Discovery Using Untargeted Metabolomic Data. *Front Mol Biosci* **3**, 30 (2016).
53. Yamada, M., Jitkrittum, W., Sigal, L., Xing, E.P. & Sugiyama, M. High-dimensional feature selection by feature-wise kernelized lasso. *Neural computation* **26**, 185-207 (2014).
54. Vergara, J.R. & Estévez, P.A. A review of feature selection methods based on mutual information. *Neural computing and applications* **24**, 175-186 (2014).

55. Xiao, L.*, et al.* Development and Validation of Machine Learning-Based Marker for Early Detection and Prognosis Stratification of Nonalcoholic Fatty Liver Disease. *Adv Sci (Weinh)*, e10527 (2025).
56. Zhou, W.*, et al.* Latent class analysis-derived classification improves the cancer-specific death stratification of molecular subtyping in colorectal cancer. *NPJ Precis Oncol* **7**, 60 (2023).
57. Sinha, P., Calfee, C.S. & Delucchi, K.L. Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls. *Crit Care Med* **49**, e63-e79 (2021).
58. Reboussin, B.A., Ip, E.H. & Wolfson, M. Locally dependent latent class models with covariates: an application to under-age drinking in the USA. *J R Stat Soc Ser A Stat Soc* **171**, 877-897 (2008).
59. European Association for the Study of the, L. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. *J Hepatol* **75**, 659-689 (2021).
60. DeLong, E.R., DeLong, D.M. & Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845 (1988).
61. Tsamos, G.*, et al.* Therapeutic Potentials of Reducing Liver Fat in Non-Alcoholic Fatty Liver Disease: Close Association with Type 2 Diabetes. *Metabolites* **13**(2023).
62. 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 e1612 (2020).
63. Jamialahmadi, O.*, et al.* Exome-Wide Association Study on Alanine Aminotransferase Identifies Sequence Variants in the GPAM and APOE Associated With Fatty Liver Disease. *Gastroenterology* **160**, 1634-1646.e1637 (2021).
64. Cusi, K.*, et al.* Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) in People With Diabetes: The Need for Screening and Early Intervention. A Consensus Report of the American Diabetes Association. *Diabetes Care* **48**, 1057-1082 (2025).
65. Lee, C.W. Albumin-bilirubin score as a practical tool for prognostication in metabolic-associated steatotic liver disease. *World J Hepatol* **17**, 113844 (2025).
66. Gellert-Kristensen, H.*, et al.* Combined Effect of PNPLA3, TM6SF2, and HSD17B13 Variants on Risk of Cirrhosis and Hepatocellular Carcinoma in the General Population. *Hepatology* **72**, 845-856 (2020).
67. Leoni, S.*, et al.* Current guidelines for the management of non-alcoholic fatty liver disease: A systematic review with comparative analysis. *World J Gastroenterol* **24**, 3361-3373 (2018).