JHEP|Reports

Integration of *PNPLA3* and *TM6SF2* genotypes provides incremental improvement in advanced fibrosis prediction among MASLD patients with type 2 diabetes mellitus

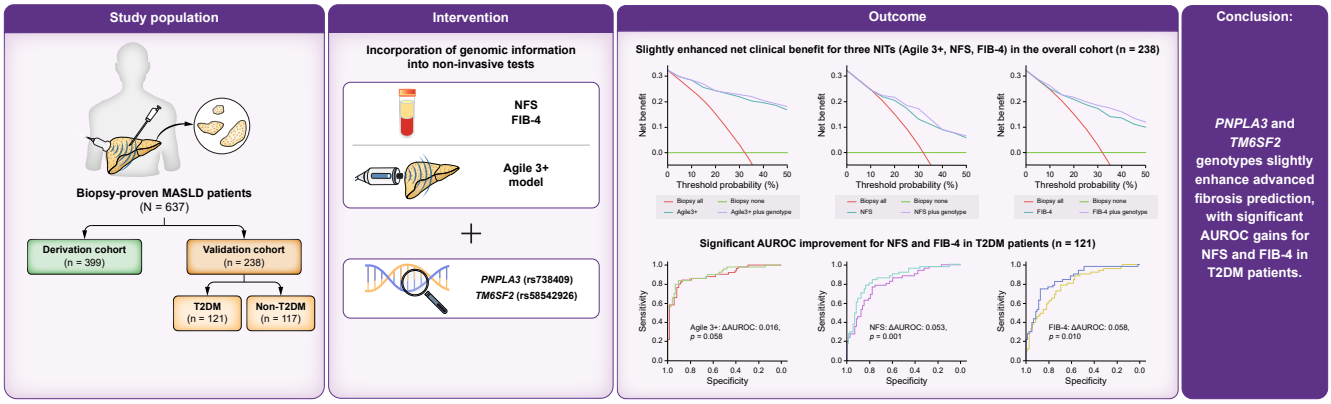
Authors

Dong Yun Kim, Hyun-Soo Zhang, Jae Seung Lee, …, Heon Yung Gee, Jung Il Lee, **Jun Yong Park,**

Correspondence

mdflorence@yuhs.ac (J.I. Lee), drpjy@yuhs.ac (J.Y. Park).

Graphical abstract



Highlights:

- Integration of *PNPLA3* and *TM6SF2* genotypes modestly improves advanced fibrosis prediction.
- Net clinical benefit was consistently improved across Agile 3+, NFS, and FIB-4.
- Adding genotype data to NFS and FIB-4 led to significant AUROC gains in patients with T2DM.
- Genetic data may support clinical decision making in high-risk metabolic populations.

Impact and implications:

Our research demonstrates that incorporating *PNPLA3* and *TM6SF2* genetic information into non-invasive fibrosis tests provides modest but measurable improvements in clinical utility for patients with metabolic dysfunction-associated steatotic liver disease, particularly those with type 2 diabetes mellitus – a high-risk population prone to accelerated fibrosis progression and liver-related complications. These findings are relevant for clinicians managing these patients, as genotype-enhanced models (particularly NAFLD fibrosis score and Fibrosis-4 index) showed statistically significant improvements in diagnostic accuracy, enabling better identification of patients requiring closer monitoring. In practice, genetic information could be integrated into routine risk stratification, especially in settings where elastography is unavailable, allowing serum-based tests to more accurately identify high-risk patients who warrant referral for biopsy or specialist evaluation. However, given the modest absolute improvements and tertiary care derivation, further validation in diverse populations is essential before routine clinical implementation can be recommended.

https://doi.org/10.1016/j.jhepr.2025.101713

© 2025 The Authors. Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Research article

JHEP|Reports

Integration of *PNPLA3* and *TM6SF2* genotypes provides incremental improvement in advanced fibrosis prediction among MASLD patients with type 2 diabetes mellitus

Dong Yun Kim[1], Hyun-Soo Zhang[2], Jae Seung Lee[1], Hye Won Lee[1], Mi Na Kim[1], Beom Kyung Kim[1], Seung Up Kim[1], Do Young Kim[1], Sang Hoon Ahn[1], Hyun Woong Lee[3], Heon Yung Gee[4], Jung Il Lee[3,*,‡], Jun Yong Park[1,*,‡]

JHEP Reports **2026**. vol. 8 | 1–12

Check for updates

Background & Aims: Genetic information is not yet used for the clinical diagnosis of advanced fibrosis in patients with metabolic dysfunction-associated steatotic liver disease (MASLD). Herein, we investigated whether incorporating genetic information regarding *PNPLA3* and *TM6SF2* into existing non-invasive fibrosis scoring systems could improve their predictive accuracy, particularly in patients with type 2 diabetes mellitus (T2DM), a high-risk population for MASLD-related complications.

Methods: Data were collected from a cohort of 637 patients with biopsy-proven MASLD. All participants underwent liver stiffness measurement (LSM), serum marker analysis, and genotyping for *PNPLA3* (rs738409), *TM6SF2* (rs58542926), and other relevant single nucleotide polymorphisms. We evaluated the benefit of adding genetic information to existing non-invasive tests (NITs) – including the Agile 3+, Fibrosis-4 (FIB-4) index, and NAFLD fibrosis score (NFS).

Results: Decision curve analysis in the validation cohort (n = 238) demonstrated that incorporating *PNPLA3* and *TM6SF2* genetic information marginally enhanced net clinical benefit across all three models over a range of threshold probabilities (10-50%). At a 30% threshold probability, the net benefit of genotype-enhanced models increased from 22.0 to 22.8 per 100 patients for Agile 3+, from 17.0 to 18.4 for NFS, and from 13.0 to 16.9 for FIB-4. In the T2DM subgroup (n = 121), genotype incorporation led to small but statistically significant improvements in discrimination for NFS (AUROC increase: 0.053, $p = 0.001$) and FIB-4 (AUROC increase: 0.058, $p = 0.010$), while Agile 3+ showed a favorable trend (AUROC increase: 0.016, $p = 0.058$).

Conclusions: Incorporating *PNPLA3* and *TM6SF2* genetic information into non-invasive fibrosis scoring systems for MASLD provides limited but measurable improvements, with statistically significant AUROC gains for NFS and FIB-4, particularly among patients with T2DM. Further validation is required before routine clinical implementation can be recommended.

© 2025 The Authors. Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) is a prevalent condition, affecting approximately 30% of the adult population worldwide.[1] MASLD often progresses to metabolic dysfunction-associated steatohepatitis, which can further advance to liver fibrosis and ultimately cirrhosis.[2] Among these stages, advanced fibrosis (stage ≥3) stands out as the most critical prognostic factor.[3,4] The development of advanced fibrosis significantly influences patients' prognosis and mortality, due to its association with severe liver-related complications, such as decompensated cirrhosis and hepatocellular carcinoma (HCC), as well as extrahepatic events, including cardiovascular diseases.[3,5,6] Thus, in patients with MASLD, early diagnosis and accurate staging of liver fibrosis are critical for improving clinical outcomes and guiding appropriate therapeutic interventions.

Unfortunately, in clinical practice, there remains an unmet need for reliable non-invasive tools for early diagnosis of advanced fibrosis. Traditional methods, like liver biopsy, are accurate but also invasive and carry risks, such as bleeding and infection, making them unsuitable for routine screening.[7] Thus, non-invasive tests (NITs) have been developed and increasingly utilized. The NAFLD fibrosis score (NFS) and Fibrosis-4 index (FIB-4) are the most extensively studied non-invasive marker panels, and are useful diagnostic tools for identifying advanced liver fibrosis.[8] Notably, European guidelines recommend the FIB-4 index as a first-line test in primary care.[9]

Among the current NITs used in this emerging clinical setting, the new Agile 3+ score seems highly promising.[10] The Agile 3+ score is based on the combination of aspartate aminotransferase/alanine aminotransferase ratio, platelet

* Corresponding authors. Addresses: Department of Internal Medicine, Gangnam Severance Hospital, Yonsei University College of Medicine, 211, Eonju-ro, Gangnam-gu, Seoul, Republic of Korea; Tel.: 82-2-2019-3310, fax: 82-2-3463-3882. (J.I. Lee), or Department of Internal Medicine, Yonsei University College of Medicine 50-1 Yonsei-ro, Seodaemun–gu, Seoul, 03722, Republic of Korea; Tel.: 82-2-2228-1988, fax: 82-2-393-6884. (J.Y. Park).
E-mail addresses: mdflorence@yuhs.ac (J.I. Lee), drpjy@yuhs.ac (J.Y. Park).
‡ Jung Il Lee and Jun Yong Park contributed equally to this work as co-corresponding authors.
https://doi.org/10.1016/j.jhepr.2025.101713

count, presence of type 2 diabetes mellitus (T2DM), sex, age, and liver stiffness measurement (LSM) values by vibration-controlled transient elastography (VCTE; FibroScan, Echosens, France).[11] Preliminary studies have reported that it shows good accuracy for the diagnosis of advanced fibrosis and the prediction of liver-related events in patients with MASLD,[12] and that it exhibits promising performance for predicting advanced fibrosis with high accuracy in comparison with LSM and FIB-4.[11]

Genetic factors play a significant role in MASLD pathogenesis and progression, suggesting that their integration could potentially improve the predictive accuracy of existing models.[13] Genome-wide association studies have yielded the identification of several variants that are significantly associated with MASLD. For example, one well-known genetic risk factor for MASLD is a coding variant in patatin-like phospholipase domain containing protein 3 (*PNPLA3*), an I-to-M substitution at position 148 (chr22:43928847, rs738409 C>G).[14] This rs738409 variant has been repeatedly found to be associated with MASLD or increased hepatic fat content.[15,16] MASLD susceptibility is also known to be associated with other genes,[17] including transmembrane 6 superfamily member 2 (*TM6SF2*), membrane-bound *O*-acyltransferase domain containing 7 (*MBOAT7*), glucokinase regulator (*GCKR*), and hydroxysteroid 17β-dehydrogenase (*HSD17B13*),[18] among others.[19,20]

Beyond genetic susceptibility, patients with MASLD and T2DM represent a particularly high-risk population requiring accurate fibrosis assessment. The prevalence of MASLD in patients with T2DM exceeds 70%, with approximately one in five having clinically significant fibrosis.[21] Studies have demonstrated that T2DM significantly increases the risk of severe hepatic and extrahepatic complications.[22] Specifically, longer T2DM duration is strongly associated with higher rates of liver-related events, including HCC and cirrhotic complications.[23] Given this elevated risk and the clinical urgency of identifying advanced fibrosis, refining diagnostic tools for patients with MASLD and T2DM is essential. Despite the established role of genetic variants in MASLD progression and the heightened risk profile of patients with T2DM, current NITs do not incorporate genetic information, limiting optimal risk stratification in this high-risk population.

In the present study, we aimed to assess whether incorporating these well-established genetic variants into existing NITs (including the Agile 3+, NFS, and FIB-4) would improve the diagnostic accuracy for predicting advanced fibrosis, particularly in the T2DM subgroup. The research objective was to refine existing NITs to enhance diagnostic assessment of patients with MASLD at the highest risk of disease progression.

Patients and methods

Study population

We analyzed patient cohorts from two university hospitals with tertiary hepatology clinics in Seoul, Korea: Severance Hospital (n = 414, December 2015–September 2023) and Gangnam Severance Hospital (n = 247, July 2022–June 2025) (Fig. S1). This study was conducted as a retrospective observational study, in which all patients who met the eligibility criteria during the study period were consecutively included. The included patients were Asian adults (age ≥19) who underwent liver biopsy for MASLD evaluation, with concomitant routine blood testing for biological markers and LSM by VCTE (FibroScan, Echosens, France). Twenty-four patients were excluded because they could not undergo LSM by VCTE. The exclusion criteria were hepatitis B or C virus infection, autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis, drug-induced liver injury, steatosis, Wilson's disease, hemochromatosis, excessive alcohol consumption (>30 g/day for males, >20 g/day for females), and malignancy diagnosed within the past year.

This study ultimately included 637 patients with histologically diagnosed MASLD and sufficient clinical data for NIT risk score calculation, including FibroScan. We chose liver biopsy as the reference standard because it is widely considered the gold standard for diagnosing advanced fibrosis, despite its known limitations. The pathologists assessing liver histology were blinded to the NIT and genotype results, while the operators performing LSM were not aware of the histological staging. The study cohort was analyzed to assess the accuracy of NITs for advanced fibrosis diagnosis. We designated the 399 patients from Severance Hospital as the derivation cohort and the 238 patients from Gangnam Severance Hospital as the validation cohort, thereby enabling an external validation structure.

Liver biopsy was conducted using liver specimens of ≥15 mm in length, or containing ≥10 complete portal tracts. Steatosis was assessed as the percentage of hepatocytes containing fat droplets (minimum 5%), and fibrosis was staged from 0–4 using the Kleiner classification.[24] A pre-specified subgroup analysis was performed for patients with T2DM to evaluate potential variability in diagnostic accuracy.

No formal power calculation was performed to determine sample size; rather, the total number of eligible patients across the two centers was included. This design reflects the real-world availability of data in these tertiary institutions. All participants were non-family members. Written informed consent was obtained from participants or their legal caregivers. The study protocol was approved by the Institutional Review Boards of Severance Hospital (IRB No. #4-2015-0184 and #4-2018-0537) and Gangnam Severance Hospital (IRB No. #3-2022-0166). This study was performed following the ethical standards of the latest amended Declaration of Helsinki.

NIT risk scores for advanced fibrosis detection

After overnight fasting, LSM was determined by VCTE with FibroScan (Echosens, Paris, France), using medium (M) and extra-large (XL) probes, as appropriate. LSM of <7.5 kPa indicated low risk, 7.5–9.5 kPa intermediate risk, and >9.5 kPa high risk of advanced fibrosis.[25]

VCTE results were considered valid if they met the following criteria: ≥10 valid measurements, a success rate of ≥60%, and an interquartile range/median liver stiffness of ≤30%. Risk scores were calculated using Agile 3+, FIB-4, and NFS.[8,11] Calculations were performed based on the existing literature (Table S1).[26] All clinical and laboratory covariates for NIT calculation were obtained on the same day as, or within 30 days of, the liver biopsy, to ensure temporal alignment of measurements.

Genotyping of study participants

All patients were genotyped using the Affymetrix Axiom PMRA chip (Affymetrix Inc., Santa Clara, CA, USA) following the manufacturer's protocol. Genotypes were called using Affymetrix Power Tools software. Quality control involved discarding samples with a call rate below 95%, and excluding single nucleotide polymorphisms (SNPs) with a call rate <98%, minor allele frequency <1%, and Hardy-Weinberg equilibrium p value $<10^{-6}$. After SNP acquisition, we determined the genotypes of five SNPs: rs738409 C>G in *PNPLA3*, rs58542926 C>T in *TM6SF2*, rs1260326 T>C in *GCKR*, rs641738 C>T in *MBOAT7-TMC4*, and the rs72613567 adenine insertion (A-INS) in *HSD17B13*.

Model development

Logistic regression was employed to model the risk of advanced liver fibrosis after determining that alternative methods, such as extreme gradient boosting, did not improve predictive performance. Univariable logistic regressions were performed using the Agile 3+, NFS, and FIB-4 scores as predictors to recalibrate the models to the derivation dataset (Severance Hospital), thus forming the "original" models.[27] Next, selected genotype variables were incorporated into each of the three original models, thereby creating the respective "genotype" models. To determine which of the five candidate genotype variables should be included in the genotype models, we utilized Akaike information criterion (AIC)-based selection (forward, backward, and stepwise) and the LASSO (least absolute shrinkage and selection operator).[28,29] AIC-based selection identified variables that minimized the AIC, while LASSO evaluated the importance of genotype variables by adjusting the lambda penalty (Tables S2 and S3).

Genotype information was combined with the raw score of an existing NIT in a multivariable logistic regression model. Specifically, the genotype information regarding *PNPLA3* rs738409 (C>G) and *TM6SF2* rs58542926 (C>T) was combined with the raw score of an existing NIT as follows:

Logit(probability of advanced fibrosis) = intercept + β_0 × raw NIT score + β_1 × genotype of *PNPLA3* rs738409 + β_2 × genotype of *TM6SF2* rs58542926,

where Logit is the log-odds function, β coefficients are logistic regression coefficients, the raw NIT score is calculated using the equations provided in Table S1, and the genotypes are additively included with their values ordered as *PNPLA3* rs738409 (CC = 0, CG = 1, and GG = 2), and *TM6SF2* rs58542926 (CC = 0, CT = 1, and TT = 2) (Table S4–S6).

This combination strategy is similar to how predictive variables have been combined into the Agile 3+ and NFS NITs.[8,11]

We then tested and validated the final models in the validation cohort (Gangnam Severance Hospital) to evaluate their diagnostic performance for advanced fibrosis.

Statistical analysis

Descriptive statistics were analyzed separately for the derivation cohort (Severance Hospital, n = 399) and validation cohort (Gangnam Severance Hospital, n = 238) cohorts. Continuous variables were presented as mean (SD), and compared by independent two-sample t test, assuming approximate normality. Categorical variables were reported as n (%), and associations were tested using chi-squared test or Fisher's exact test (for expected counts <5).

Logistic regression was employed to model the risk of advanced liver fibrosis. Model discrimination was assessed by the area under the receiver-operating characteristic curve (AUROC), comparing the original and genotype models for each prediction model. DeLong's test was used to assess the statistical significance of AUROC differences.[30] Model calibration was evaluated using the Hosmer-Lemeshow goodness-of-fit test, with p >0.05 indicating an adequate fit.[27] Due to the low sensitivity and statistical power of AUC comparisons,[31,32] we also conducted decision curve analysis (DCA) to evaluate clinical utility.[33,34] Net benefit was assessed using DCA to evaluate the models' clinical utility at varying threshold probabilities, and these net benefits were compared between models with and without genotype inclusion.[34] The net benefits for clinically plausible threshold probabilities were plotted as decision curves, enabling visual comparison between the original and genotype models. We also evaluated threshold probabilities of 10%, 15%, and 30%, which are decision points commonly used to guide advanced fibrosis diagnosis in patients with MASLD in clinical practice. The threshold probability represents the risk level at which a clinician would consider further diagnostic evaluation, such as liver biopsy, to confirm or exclude advanced fibrosis.

For risk stratification using dual cut-off values (low/rule-out and high/rule-in), different approaches were used for models with and without genetic information. For the models without genetic information (original NITs), we applied their respective previously established standard cut-off values.[8,11,26] For the newly developed genotype models, cut-off values were determined using a data-driven approach based on performance in the derivation cohort: the low (rule-out) cut-off was set to achieve a sensitivity of ≥85%, and the high (rule-in) cut-off was set to achieve a specificity of ≥85%. These data-driven cut-offs were based on the predicted probabilities of advanced fibrosis, as there are not yet established clinically interpreted risk scores for the new genotype models. Using these distinct cut-off strategies, negative predictive value (NPV) in the low-risk group, positive predictive value (PPV) in the high-risk group, and the percentage of biopsies avoided (*i.e.* the proportion of patients outside the indeterminate zone) were compared between the original and genotype models. The predicted probabilities of advanced fibrosis were used to indicate the dual cut-offs in the newly developed genotype models.

Model development was performed exclusively on the derivation dataset (Severance Hospital, Table S7–S10), with all validations conducted on the independent validation dataset (Gangnam Severance Hospital) to independently evaluate predictive performance. All statistical tests were two-sided, with a p value <0.05 considered to indicate statistical significance. Analyses were conducted using R version 4.3.2 (The R Foundation for Statistical Computing, Vienna, Austria).

Results

Baseline characteristics of the study population

The total dataset comprised 637 patients, including 399 in the derivation cohort (to construct the scores) and 238 in the validation cohort. These two cohorts had similar demographic, metabolic, and serological characteristics, in terms of the collected parameters (Table 1). The prevalence of advanced

fibrosis was 29.8% and 32.4% in the derivation and validation cohorts, respectively. In the full cohort of all participants, the minor allele frequencies were 0.477 for the G allele of rs738409, 0.031 for the T allele of rs58542926, 0.436 for the C allele of rs1260326, 0.189 for the T allele of rs641738, and 0.232 for the A allele of rs72613567. The minor allele frequencies of these five SNPs did not significantly differ between the derivation and validation cohorts (Table 1).

Overall diagnostic accuracy of NITs for advanced fibrosis

Table 2 presents the overall diagnostic accuracies of three NITs for predicting advanced fibrosis in the validation cohort. As expected, the Agile 3+ model demonstrated superior diagnostic accuracy compared to NFS and FIB-4. When applying the standard, previously established, cut-offs for the models without genetic information, the Agile 3+ showed high NPVs in the low (rule-out) group of 0.94 (95% CI 0.88–0.97), and higher PPVs in the high (rule-in) group of 0.73 (95% CI 0.63–0.81). The indeterminate zone, a critical measure of diagnostic uncertainty, was lowest for Agile 3+ (10.1%), compared to those for NFS (30.7%) and FIB-4 (25.0%), indicating that Agile 3+ yielded a clearer diagnostic outcome. Table S7 presents the overall diagnostic accuracies of three NITs for predicting advanced fibrosis in the derivation cohort. These results showed a trend similar to that observed in the test cohort.

Table 1. Characteristics of patients with metabolic dysfunction-associated steatotic liver disease (n = 637).

	Validation cohort (n = 238)	Derivation cohort (n = 399)	*p* value[†]
Age, years	49.3 ± 16.0	49.8 ± 15.5	0.69
Gender, n (%)			0.55
Male	117 (49.2)	185 (46.4)	
Female	121 (50.8)	214 (53.6)	
DM			0.82
No	117 (49.2)	191 (47.9)	
Yes	121 (50.8)	208 (52.1)	
BMI, kg/m^2	28.2 ± 5.0	28.3 ± 5.2	0.68
Platelet count, 10^3/μl	232.4 ± 77.0	228.9 ± 75.1	0.56
Total cholesterol, mg/dl	190.4 ± 57.5	199.9 ± 63.4	0.06
Protein, g/dl	7.3 ± 0.6	7.3 ± 0.5	0.85
Albumin, g/dl	4.4 ± 0.4	4.4 ± 0.4	0.77
AST, IU/L	61.9 ± 52.7	61.2 ± 46	0.86
ALT, IU/L	63.9 ± 63.7	72.5 ± 74.8	0.13
GGT, IU/L	93.7 ± 136.4	81.2 ± 118.4	0.24
MASH, n (%)	131 (55.0)	212 (53.1)	0.64
Advanced fibrosis, n (%)	77 (32.4)	119 (29.8)	0.56
Fibrosis stage			
Stage 0	10 (4.2)	22 (5.5)	0.41
Stage 1	129 (54.2)	212 (53.1)	
Stage 2	35 (14.7)	46 (11.5)	
Stage 3	24 (10.1)	57 (14.3)	
Stage 4	40 (16.8)	62 (15.5)	
Genotype information			
rs738409 chr22; *PNPLA3*			0.08
CC	52 (21.8)	72 (18)	
CG	90 (37.8)	187 (46.9)	
GG	96 (40.3)	140 (35.1)	
rs58542926 chr19: *TM6SF2*			0.14
CC	194 (81.5)	300 (75.2)	
CT	43 (18.1)	94 (23.6)	
TT	1 (0.4)	5 (1.3)	
rs641738 chr19: *MBOAT7-TMC4*			0.60
CC	150 (63.0)	253 (63.4)	
CT	73 (30.7)	128 (32.1)	
TT	15 (6.3)	18 (4.5)	
rs1260326 chr2: *GCKR*			0.81
TT	86 (36.1)	134 (33.6)	
TC	112 (47.1)	195 (48.9)	
CC	40 (16.8)	70 (17.5)	
rs72613567 chr4: *HSD17B13*			0.79
A/A	17 (7.1)	26 (6.5)	
–/A	89 (37.4)	141 (35.3)	
–/–	132 (55.5)	232 (58.1)	
Non-invasive tests			
FIB-4	2.3 ± 2.5	2.2 ± 2.0	0.52
NFS	–1.2 ± 2.3	–1.2 ± 2.3	0.85
LSM by FibroScan (kPa)	13.1 ± 9.7 (3.6–75.0)	11.6 ± 7.4 (3.3–67.1)	0.06

ALT, alanine aminotransferase; AST, aspartate aminotransferase; DM, diabetes mellitus; FIB-4, Fibrosis-4 index; GGT, gamma-glutamyltransferase; LSM, liver stiffness measurement; MASH, metabolic dysfunction-associated steatohepatitis; NAS, NAFLD activity score; NFS, NAFLD fibrosis score.
Values are presented as mean ± standard deviation or n (%).
[†]*p* values obtained by two-sample *t* test or Mann-Whitney *U* test under non-normality, and Chi-squared test or Fisher's exact test (when expected counts <5).

Table 2. Diagnostic performance of the Agile 3+, NFS, and FIB-4 models of advanced fibrosis prediction, with or without the *PNPLA3* and *TM6SF2* genotype variables, in the validation cohort (n = 238) of the MASLD study population.

	Agile 3+		NFS		FIB-4	
	With genotypes	Without genotypes	With genotypes	Without genotypes	With genotypes	Without genotypes
Rule-out cut-off pr. (≥85% Se.)	0.247	0.451	0.205	-1.455	0.198	1.30
Se. (95% CI)	0.90 (0.81–0.95)	0.90 (0.81–0.95)	0.83 (0.73–0.90)	0.85 (0.75–0.91)	0.88 (0.80–0.94)	0.86 (0.76–0.92)
Sp. (95% CI)	0.75 (0.68–0.81)	0.73 (0.66–0.79)	0.61 (0.53–0.68)	0.64 (0.57–0.71)	0.57 (0.50–0.65)	0.58 (0.51–0.66)
NPV (95% CI)	0.94 (0.88–0.97)	0.94 (0.88–0.97)	0.92 (0.85–0.96)	0.90 (0.83–0.94)	0.91 (0.84–0.95)	0.90 (0.83–0.94)
Indeterminate zone (%)	8.8%	10.1%	26.6%	30.7%	27.9%	25.0%
Rule-in cut-off pr. (≥85% Sp.)	0.388	0.679	0.448	0.676	0.356	2.67
Se. (95% CI)	0.84 (0.75–0.91)	0.83 (0.73–0.90)	0.63 (0.52–0.73)	0.46 (0.35–0.57)	0.59 (0.48–0.69)	0.56 (0.45–0.67)
Sp. (95% CI)	0.86 (0.80–0.90)	0.85 (0.79–0.90)	0.88 (0.82–0.92)	0.92 (0.86–0.95)	0.84 (0.78–0.89)	0.81 (0.75–0.86)
PPV (95% CI)	0.74 (0.64–0.82)	0.73 (0.63–0.81)	0.71 (0.59–0.80)	0.72 (0.58–0.82)	0.64 (0.52–0.74)	0.59 (0.47–0.69)

Sensitivity, specificity, NPV, and PPV are shown with 95% CIs. FIB-4, Fibrosis-4 index; MASLD, metabolic dysfunction-associated steatotic liver disease; NFS, NAFLD fibrosis score; NPV, negative predictive value; PPV, positive predictive value; pr., probability; Se., sensitivity; Sp., specificity.

Table 3 and Fig. S2 present the AUROCs of the NITs for advanced fibrosis diagnosis. All showed good predictive accuracy in terms of discrimination (Agile 3+, 0.919; NFS, 0.823; FIB-4, 0.789), while Agile 3+ performed better than NFS and FIB-4 (p <0.001 for both comparisons).

Changes in diagnostic accuracy of NITs with addition of genotype information

Next, we used DCA to evaluate the net benefit of adding genotype information to the Agile 3+, NFS, and FIB-4 models.

As shown in Fig. 1A, the decision curves indicated that incorporating genotype information into the Agile 3+ model provided an improved net benefit across a range of threshold probabilities from 0.1–0.5, compared to Agile 3+ alone or the strategy of treating all patients as having advanced fibrosis. Similarly, adding genetic information to the NFS model resulted in a higher net benefit within the threshold probability range of 0.1–0.5 (Fig. 1B). For the FIB-4 model, higher net benefit for including genotype information was seen within the threshold probability range of 0.1–0.5 (Fig. 1C). These findings suggest that the incorporation of genotype information generally enhanced the clinical decision-making process.

We also examined changes in diagnostic performance metrics including NPV in the low-risk (rule-out) group, PPV in the high-risk (rule-in) group, and the proportion of patients in the indeterminate zone (Table 2). For Agile 3+, adding genotypes maintained the low-risk group NPV at 0.94 (95% CI 0.88–0.97) and showed minimal change in the high-risk group PPV from 0.73 (95% CI 0.63–0.81) to 0.74 (95% CI 0.64–0.82). The indeterminate zone decreased from 10.1% to 8.8% (Fig. 2A). For the NFS model, incorporating genotypic variables resulted in an NPV of 0.92 (95% CI 0.85–0.96) compared to 0.90 (95% CI 0.83–0.94) without genotypes, while the high-risk group PPV remained comparable (0.72 *vs.* 0.71). The indeterminate zone decreased from 30.7% to 26.6% (Fig. 2B). For the FIB-4 model, genotype inclusion led to minimal change in the low-risk group NPV from 0.90 (95% CI 0.83–0.94) to 0.91 (95% CI 0.84–0.95) and an increase in the high-risk group PPV from 0.59 (95% CI 0.47–0.69) to 0.64 (95% CI 0.52–0.74). In contrast, the indeterminate zone increased from 25.0% to 27.9% (Fig. 2C).

To address potential cut-off differences between models, we performed a sensitivity analysis using identical data-driven thresholds (≥85% sensitivity for rule-out and ≥85% specificity for rule-in). When identical cut-offs were applied, genetic information reduced the indeterminate zone for both NFS (from 28.0% to 26.6%) and FIB-4 (from 28.7% to 27.8%) (Table S11).

Table 3 summarizes the AUROC values for Agile 3+, NFS, and FIB-4, with or without genotype information, in the full validation cohort (n = 238), and in the T2DM (n = 121) and non-T2DM (n = 117) subgroups. In the overall cohort, the addition of genotype information yielded modest AUROC improvements, with only NFS reaching statistical significance (Agile 3+: 0.919 to 0.925, p = 0.104; NFS: 0.823 to 0.845, p = 0.002; FIB-4: 0.789 to 0.807, p = 0.169). Importantly, in the T2DM subgroup, incorporating genetic information led to statistically significant AUROC improvements for NFS (0.806 to 0.859, p = 0.001) and FIB-4 (0.793 to 0.851, p = 0.010), while Agile 3+ showed a trend that did not reach significance (0.893 to 0.909, p = 0.058) (Fig. 3). In contrast, no improvements were observed in the non-T2DM subgroup.

Table 3. Discrimination (predictive accuracy) in terms of the AUROC of the Agile 3+, NFS, and FIB-4 models of advanced fibrosis prediction, with or without the *PNPLA3* and *TM6SF2* genotype variables, in the validation cohort and in subgroups stratified by T2DM status.

		Agile 3+		NFS		FIB-4	
Study cohorts and subgroups	AUROC comparison	With genotypes	Without genotypes	With genotypes	Without genotypes	With genotypes	Without genotypes
Validation cohort (n = 238)	AUROC (95% CI)	0.925 (0.891–0.960)	0.919 (0.884–0.954)	0.845 (0.792–0.898)	0.823 (0.767–0.879)	0.807 (0.748–0.865)	0.789 (0.730–0.848)
	AUROC difference (95% CI)		0.006 (-0.001–0.014)		0.023 (0.008–0.037)		0.018 (-0.007–0.042)
	DeLong test *p* value		0.104		0.002		0.169
T2DM subgroup (n = 121)	AUROC (95% CI)	0.909 (0.853–0.965)	0.893 (0.832–0.953)	0.859 (0.791-0.926)	0.806 (0.729–0.884)	0.851 (0.782-0.920)	0.793 (0.714–0.873)
	AUROC difference (95% CI)		0.016 (-0.001–0.033)		0.053 (0.023–0.082)		0.058 (0.014–0.102)
	DeLong test *p* value		0.058		0.001		0.010
Non-T2DM (n = 117)	AUROC (95% CI)	0.937 (0.891–0.984)	0.937 (0.891–0.984)	0.810 (0.715-0.906)	0.806 (0.710–0.902)	0.773 (0.676-0.869)	0.770 (0.675-0.866)
	AUROC difference (95% CI)		0.000 (-0.002–0.002)		0.004 (-0.004–0.013)		0.003 (-0.010–0.015)
	DeLong test *p* value		>0.999		0.330		0.709

p value for DeLong's test, which was used to compare the AUROCs between the models with and without genetic information. AUROC, area under the receiver-operating characteristic curve; FIB-4, Fibrosis-4 index; MASLD, metabolic dysfunction-associated steatotic liver disease; NFS, NAFLD fibrosis score.

Net benefit changes at specified threshold probabilities with addition of genotype information

The addition of *PNPLA3* and *TM6SF2* genotypic variables to the Agile 3+, NFS, and FIB-4 models resulted in modest improvements in net benefit for diagnosis of advanced fibrosis at clinically relevant threshold probabilities of 10%, 15%, and 30% (Table 4). The incremental benefit was limited for Agile 3+ but more evident for NFS and FIB-4, particularly at higher threshold probabilities. At a threshold probability of 10%, the net benefit for Agile 3+ showed a minimal increase from 28.4 to 28.5 per 100 patients either with or without genotypes, indicating limited additional gain at lower thresholds. As the threshold probability increased, inclusion of genotype information showed a clearer net benefit for Agile 3+: at 15%, the net benefit was 27.0 per 100 patients with genotypes *vs*. 25.6 without, and at 30%, 22.8 *vs.* 22.0 per 100 patients. In contrast, the NFS model demonstrated an increase at the 10% threshold (from 24.9 to 25.9 per 100 patients), and 30% (from 17.0 to 18.4) thresholds, though a minimal decrease was observed at the 15% threshold (from 22.5 to 22.4). Similarly, FIB-4 showed even more pronounced improvement with genotype inclusion at higher thresholds (*e.g*. from 21.7 to 22.1 per 100 patients at 15%, and from 13.0 to 16.9 at 30%). These findings suggest that incorporating *PNPLA3* and *TM6SF2* genotypes into existing NITs may enhance clinical decision-making by improving the net benefit, particularly for NFS and FIB-4, at higher threshold probabilities to justify more invasive procedures such as liver biopsy.

Changes in model calibration with addition of genotype information

In terms of the Hosmer-Lemeshow goodness-of-fit test *p* values, the inclusion of *PNPLA3* and *TM6SF2* genotype variables into the Agile 3+, NFS, and FIB-4 models had no significant impact on their respective calibration performance. For the Agile 3+ model, the Hosmer-Lemeshow test had a *p* value of 0.105 with genotypes and 0.186 without genotypes, indicating no significant difference in calibration performance with genotypic data inclusion (Table 5). Similarly, the NFS model showed high *p* values of 0.560 with genotypes and 0.797 without genotypes, suggesting adequate model fit regardless of genotype inclusion. For the FIB-4 model, the *p* value was 0.203 with genotypes compared to 0.092 without genotypes, and both *p* values suggested an acceptable fit without compromising reliability.

Discussion

Identifying patients with advanced fibrosis is important because they are at risk of unfavorable clinical outcomes. The present study is among the first to evaluate how currently available NITs are affected by inclusion of genetic information. Our findings demonstrate that incorporating *PNPLA3* and *TM6SF2* genotypes into existing NITs, particularly NFS and FIB-4, provides clinically meaningful improvements in risk stratification among patients with T2DM, primarily through enhanced discrimination and improved net clinical benefit, rather than through large improvements in global discrimination metrics across the entire cohort. While AUROC improvements in the overall cohort were modest, statistically

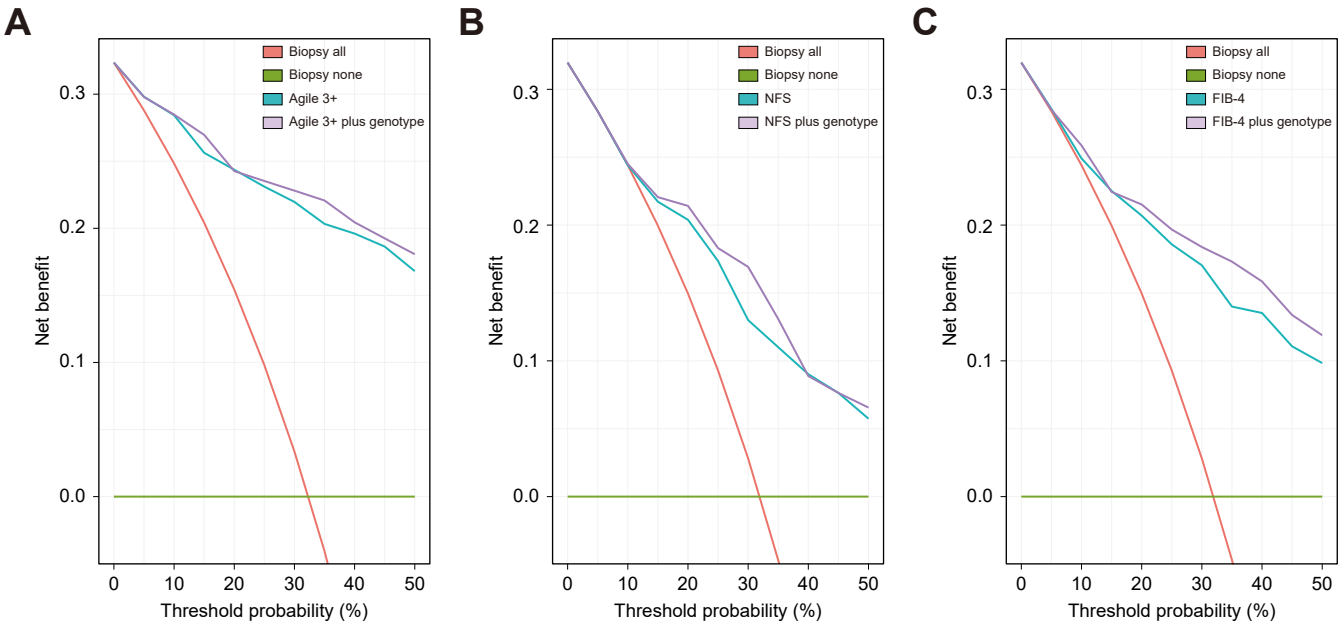


Fig. 1. Decision curves for Agile 3+, NFS, and FIB-4, with and without genetic information, for advanced fibrosis diagnosis. Net benefit was estimated using decision curve analysis; statistical details are provided in the Methods. FIB-4, Fibrosis-4 index; NFS, NAFLD fibrosis score.

significant gains were observed specifically in the T2DM subgroup for NFS and FIB-4. These findings highlight the potential usefulness of integrating genetic information into existing clinical models to optimize advanced fibrosis prediction in high-risk metabolic populations.

In particular, our results show that integrating genetic information significantly enhances diagnostic accuracy among patients with T2DM. This finding aligns with the well-established and highly reproducible principle that the detrimental effects of the high-risk *PNPLA3* genotype are significantly amplified by metabolic stressors, such as obesity, insulin resistance, and T2DM. This synergistic interaction was first demonstrated over a decade ago, in studies showing that morbid obesity unmasks the association between the *PNPLA3* risk variant and elevated liver enzymes,[35] and has since been consistently replicated in various settings, including among children with abdominal obesity[36] and across multiple genetic loci associated with fatty liver disease.[37]

Building on this foundational knowledge, recent large-scale prospective studies have confirmed that this gene–environment interaction directly translates into a higher risk of major adverse liver outcomes, including accelerated progression to cirrhosis, particularly among patients with T2DM.[38,39]

In line with these results, we noted that incorporating *PNPLA3* and *TM6SF2* genotypes led to statistically significant AUROC improvements in the T2DM subgroup for the widely-used NFS and FIB-4 scores. For the already high-performing Agile 3+ model, a similar trend of improvement was observed, although it did not reach statistical significance. To facilitate translation into practice, we generated a schematic workflow (Fig. S3) that outlines how genotype-augmented NITs may be incorporated into clinical decision-making.

Notably, the AUROC only reflects a model's predictive accuracy, without accounting for situations where false-negative results are more harmful than false-positive results. Thus, we also performed a DCA to identify threshold probabilities where the use of non-invasive criteria yielded the maximum net benefit for detecting advanced fibrosis.[34,40] These results supported that adding genotypes to each model increased the net benefit for advanced fibrosis diagnosis. Moreover, this advantage became more pronounced with an increasing threshold probability. This suggests that genetic data may be particularly valuable in clinical scenarios where a higher threshold (*e.g.* 30%) is required to justify a biopsy, compared to in cases managed more conservatively (*e.g.* biopsy recommended at a 10% probability). The increased net benefit observed on DCA – consistent with previous reports emphasizing threshold-based clinical metrics – indicates a meaningful impact on real-world decision-making beyond what global discrimination measures alone can capture.[34,40]

Importantly, DCA and the size of the indeterminate zone reflect different decision frameworks. While our dual cut-off policy increased the indeterminate zone for FIB-4, sensitivity analysis using identical data-driven thresholds demonstrated a reduction, indicating this was attributable to cut-off differences. Moreover, the genotype-augmented model nevertheless provided higher net benefit across clinically relevant thresholds, indicating improved decision-making utility even when more cases fall into the intermediate category under a dual cut-off strategy. Although the calibration *p* value for the FIB-4 model without genotypes was borderline, both models demonstrated overall acceptable fit.

Another study investigated the use of genetic information for predictions in MASLD.[41] However, it focused on predicting MASH rather than advanced fibrosis. Since the degree of fibrosis has the greatest clinical significance in MASLD, it is a strength of our present study that genetic information was specifically used for predicting advanced fibrosis.

The application of genetic information for risk stratification and prognostication in MASLD is a rapidly advancing field, with

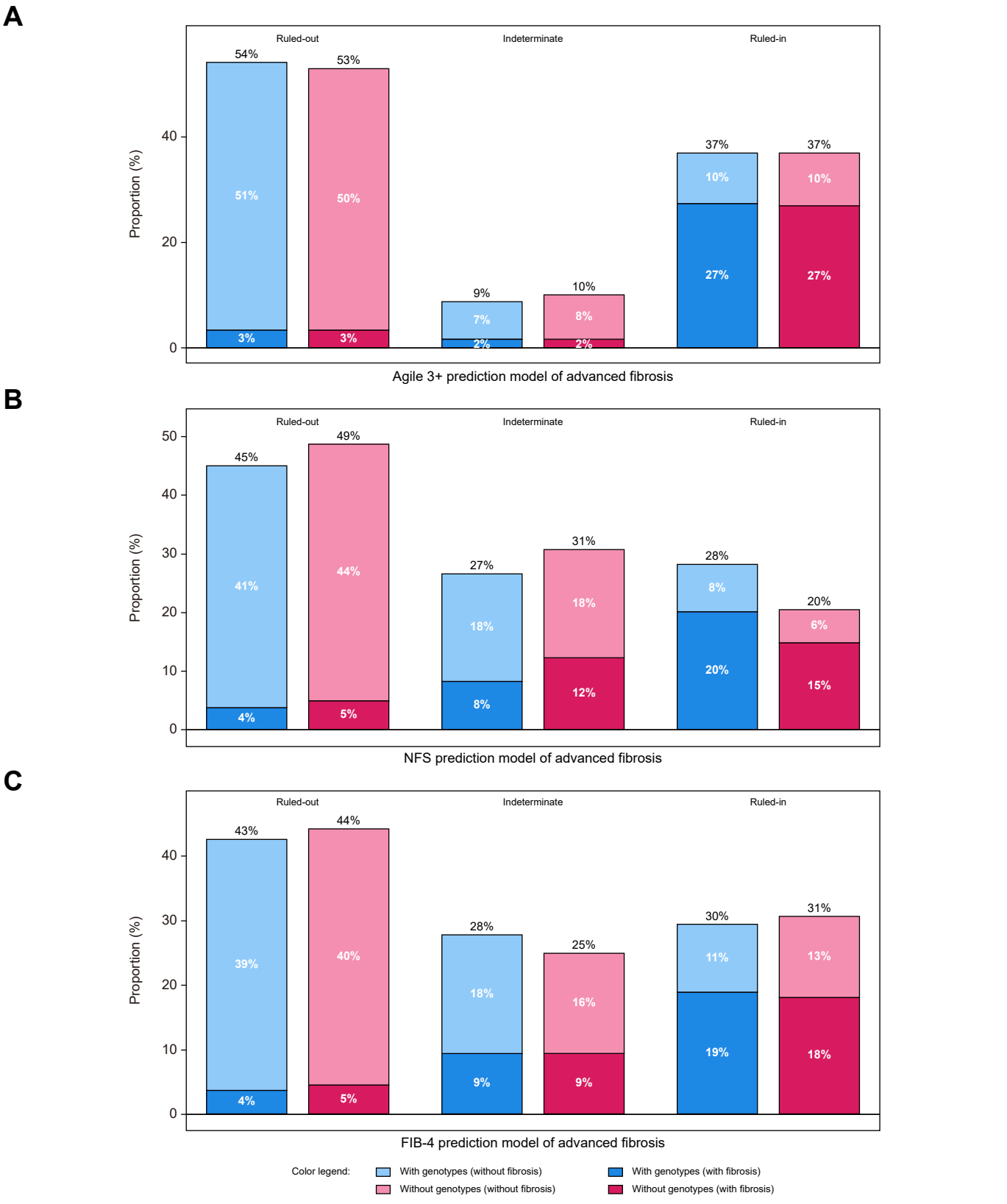


Fig. 2. Percentages of patients in the rule-out, indeterminate, and rule-in zones for advanced fibrosis diagnosis using Agile 3+, NFS, and FIB-4, with and without genetic information, in the validation cohort (n = 238). The cut-offs used were: (A) Agile 3+: with genotypes (rule-out <0.247, rule-in >0.388), without genotypes (rule-out <0.451, rule-in >0.679); (B) NFS: with genotypes (rule-out <0.205, rule-in >0.448), without genotypes (rule-out <-1.455, rule-in >0.676); (C) FIB-4:

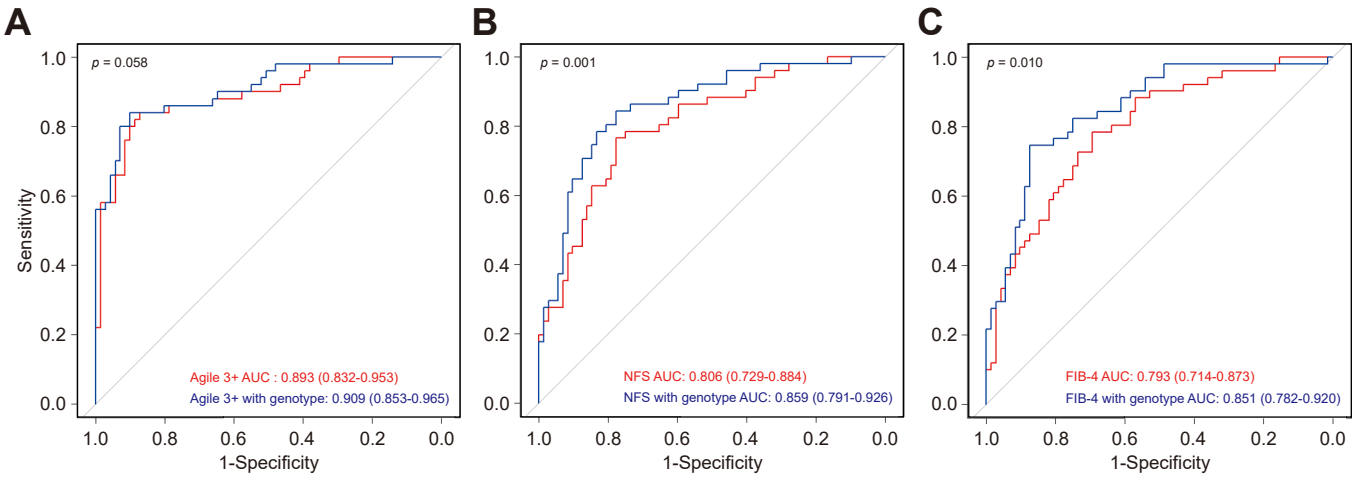


Fig. 3. AUROCs for advanced fibrosis diagnosis in the T2DM subgroup (n = 121) of the validation cohort, comparing each model with and without genotype information. (A) Agile 3+; (B) NFS; (C) FIB-4. The AUROCs were compared using DeLong's test. AUROC, area under the receiver-operating characteristic curve; FIB-4, Fibrosis-4 index; NFS, NAFLD fibrosis score; T2DM, type 2 diabetes mellitus.

Table 4. Comparison of net benefit by DCA for ruling-in advanced fibrosis, with or without the *PNPLA3* and *TM6SF2* genotype variables, for the Agile 3+, NFS, and FIB-4 prediction models in the validation cohort (n = 238) of the MASLD study population.

	Agile 3+		NFS		FIB-4	
Threshold probability of diagnosis for further biopsy	With genotypes	Without genotypes	With genotypes	Without genotypes	With genotypes	Without genotypes
10%	28.5	28.4	25.9	24.9	24.5	24.4
15%	27.0	25.6	22.4	22.5	22.1	21.7
30%	22.8	22.0	18.4	17.0	16.9	13.0

Net benefit by DCA for Agile 3+, NFS, and FIB-4 with and without *PNPLA3* and *TM6SF2* genotypes in the validation cohort. Net benefit is expressed per 100 patients at threshold probabilities of 10%, 15%, and 30%. Values represent the difference between the true-positive proportion and the false-positive proportion weighted by the odds of the threshold probability. DCA, decision curve analysis; FIB-4, Fibrosis-4 index; MASLD, metabolic dysfunction-associated steatotic liver disease; NFS, NAFLD fibrosis score.

a growing body of literature. Numerous studies have established that genetic data – often consolidated into polygenic risk scores (PRSs) – can predict the entire spectrum of MASLD-related outcomes. Notably, PRSs have been used to successfully predict the development of severe liver outcomes, such as cirrhosis and HCC, in both high-risk cohorts and the general population.[42,43] Genetic scores have also been demonstrated to predict the rate of fibrosis progression over a patient's lifetime,[44] and the transition from established cirrhosis to HCC.[19]

A key application of this research has been the integration of PRSs with established clinical fibrosis scores (*e.g.* FIB-4 and NFS) to refine their prognostic accuracy for severe liver disease.[45] Our present work builds upon this extensive foundation by specifically investigating the integration of genetic data with the high-performance Agile 3+ score for the prediction of advanced fibrosis.

The presently observed diagnostic accuracy of Agile 3+ is in line with previous findings. Sanyal *et al.*[11] reported the predictive performance of Agile 3+ in terms of discrimination, indicated by an AUROC of 0.90 (95% CI 0.88–0.91) in the derivation set, and 0.90 (95% CI 0.88–0.92) in the internal validation set. Similarly, we found an AUROC of 0.919 (95% CI 0.884–0.954). Therefore, our present analysis also externally validated the Agile 3+ score within an independent large Asian cohort with biopsy-proven MASLD. Since the Agile 3+ model performed exceptionally well, addition of genetic information had limited benefit in terms of AUROC improvement. However, genetic data incorporation provided value through modest reduction of the indeterminate zone, compared to the original model.

Because the Agile 3+ score requires VCTE, which may be limited by cost and availability, NFS and FIB-4 remain the most readily available, commonly used, and well-validated non-invasive tools for identifying patients with MASLD and advanced fibrosis.[9,46] The EASL and AASLD guidelines suggest FIB-4 as a first test, and performing LSM only in patients at intermediate-to-high risk of advanced fibrosis according to FIB-4.[9,47] However, FIB-4 and NFS scores are poorly correlated with liver stiffness, resulting in substantial proportions of false-positive and false-negative results.[48,49] Incorporating genetic information at the initial screening stage may enhance risk stratification, particularly in high-risk populations such as patients with T2DM. When elastography is unavailable, adding genomic information to serum-based fibrosis tests may provide additional value for risk assessment. Our present results

with genotypes (rule-out <0.198, rule-in >0.356), without genotypes (rule-out <1.30, rule-in >2.67). The rule-out and rule-in cut-offs for the models with genetic information were determined to achieve a sensitivity of ≥85% and a specificity of ≥85%, respectively, in the derivation cohort. For the models without genetic information, their standard, previously established cut-offs were applied. Statistical details are provided in the Methods. FIB-4, Fibrosis-4 index; NFS, NAFLD fibrosis score.

Table 5. Calibration of the Agile 3+, NFS, and FIB-4 models of advanced liver fibrosis prediction, with or without the *PNPLA3* and *TM6SF2* genotype variables, in the validation cohort (n = 238) of the MASLD study population.

	Agile 3+		NFS		FIB-4	
	With genotypes	Without genotypes	With genotypes	Without genotypes	With genotypes	Without genotypes
Hosmer-Lemeshow test *p* value	0.105	0.186	0.560	0.797	0.203	0.092

p value for the Hosmer-Lemeshow test, used to assess the goodness-of-fit (calibration) of the logistic regression models. A non-significant *p* value (*e.g.* >0.05) indicates a good fit. FIB-4, Fibrosis-4 index; MASLD, metabolic dysfunction-associated steatotic liver disease; NFS, NAFLD fibrosis score.

demonstrated that adding genotype information to NFS and FIB-4 improved the net benefit for diagnosing advanced fibrosis, at threshold probabilities between 10–50%.

A key strength of our study is the evaluation of genotype-enhanced NITs within a large biopsy-confirmed MASLD cohort. We demonstrated that incorporating the *PNPLA3* rs738409 and *TM6SF2* rs58542926 genotypes yielded statistically significant discrimination improvements for NFS and FIB-4 in patients with T2DM, along with consistent net benefit gains demonstrated by DCA. Furthermore, to our knowledge, no currently available NIT for advanced fibrosis incorporates genetic data, which can provide consistent and reliable insights, compensating for the variability of biomarkers that are influenced by external factors. Here, we demonstrated successful use of genetic information, with low cut-off points to exclude advanced fibrosis, and high cut-off points to identify advanced fibrosis. This novel approach enhanced the predictive accuracy and clinical utility of the tested NITs. Additionally, our large cohort of 637 patients with biopsy-confirmed MASLD enabled evaluation of the accuracy of our models by using liver biopsy as a reference. Finally, the specific genetic data utilized in this study is feasible to evaluate in clinical settings, requiring only two genetic markers: the rs738409 (C>G) variant in *PNPLA3* and rs58542926 (C>T) in *TM6SF2*.

The potential usefulness of genetic information extends beyond improving point-in-time diagnostic accuracy. In particular, it may help predict long-term disease progression, guiding earlier intervention and monitoring strategies for individuals at highest risk. Although the cost and availability of genotyping remain challenges, future cost-effectiveness analyses could clarify whether routine genotyping in MASLD improves overall outcomes and is economically feasible in diverse healthcare settings.

Genetic information is becoming integral for improving both disease prediction and patient stratification for drug development. Several pharmaceutical companies are developing therapies that target specific genetic markers. For example, Janssen's JNJ-75220795 small-interfering RNA targets the *PNPLA3* gene variant for MASLD treatment.[50] As more emerging therapies enter clinical trials, the combination of genotype-enhanced prediction and novel therapeutics could further personalize treatment pathways. This personalized medicine approach aligns therapeutic choices with individual genetic profiles, potentially maximizing efficacy while minimizing unnecessary interventions.

This study has several limitations. First, allele frequencies of genetic variants can significantly differ among different ethnic populations; thus, the predictive power and applicability of genetic markers identified in this study may not directly translate to other populations. Validation in diverse cohorts is necessary.

Second, different cut-off strategies were used for models with and without genetic information. Sensitivity analysis using identical thresholds showed consistent results, alleviating this concern. External validation with standardized cut-offs is nonetheless warranted.

Finally, our analyses were conducted in a tertiary care setting with a high prevalence of advanced fibrosis (29%). In populations with lower prevalence, the PPV would likely be lower and the NPV higher. Additionally, the tertiary care setting also led to enrichment of higher risk patients and a corresponding greater prevalence of the *PNPLA3* 'G' allele, which may limit generalizability to community-based populations. Nonetheless, we believe that the model offers clinical utility, as it was specifically designed to refine risk stratification among patients with suspected advanced fibrosis – a population that is more likely to carry the *PNPLA3* 'G' allele. In this setting, the ability to reduce unnecessary liver biopsies via more accurate non-invasive prediction is directly relevant to clinical practice.

In conclusion, incorporating *PNPLA3* and *TM6SF2* genetic information into existing NITs for MASLD provides limited but measurable improvements in advanced fibrosis risk stratification. DCA demonstrated consistent net benefit improvements across all three models, with additional statistically significant AUROC improvements for NFS and FIB-4 in patients with T2DM. While these gains are incremental, they may support clinical decision-making. Further validation in diverse populations and cost-effectiveness analyses are necessary before recommending widespread clinical implementation.

Affiliations

[1]*Department of Internal Medicine, Yonsei University College of Medicine, Yonsei Liver Center, Severance Hospital, 50-1 Yonsei-ro, Seodaemun–gu, Seoul, Republic of Korea;* [2]*Biostatistics Collaboration Unit, Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea;* [3]*Department of Internal Medicine, Yonsei University College of Medicine, Gangnam Severance Hospital, 211, Eonju-ro, Gangnam-gu, Seoul, Republic of Korea;* [4]*Department of Pharmacology, Yonsei University College of Medicine, Seoul, Republic of Korea*

Abbreviations

AIC, Akaike information criterion; AUROC, area under the receiver-operating characteristic curve; DCA, decision curve analysis; FIB-4, Fibrosis-4 index; GCKR, glucokinase regulator; HCC, hepatocellular carcinoma; HSD17B13, hydroxysteroid 17β-dehydrogenase type 13; LSM, liver stiffness measurement; MASLD, metabolic dysfunction-associated steatotic liver disease; MASH, metabolic dysfunction-associated steatohepatitis; MBOAT7, membrane-bound O-acyltransferase domain containing 7; NAFLD, non-alcoholic fatty liver disease; NFS, NAFLD fibrosis score; NITs, non-invasive tests; NPV, negative predictive value; PNPLA3, patatin-like phospholipase domain containing protein 3; PPV, positive predictive value; SNP, single nucleotide polymorphism; T2DM, type 2 diabetes mellitus; TM6SF2, transmembrane 6 superfamily member 2; VCTE, vibration-controlled transient elastography.

Financial support

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI23C1493 to D.Y.K. and 2020M3A9E403918422 to J.Y.P.).

Conflict of interest

The authors declare that they have no competing interests.
Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Concept and design: D.Yu.K., J.I.L., and J.Y.P.; data analyses: D.Yu.K. and H.S.Z.; data collection and management: J.S.L., H.W.L., M.N.K., B.K.K., S.U.K., D.Yo.K., S.H.A, H.Wo.L., and H.Y.G.; interpretation of data: D.Yu.K.; writing of the manuscript: D.Yu.K.; supervision: J.I.L., and J.Y.P. All authors approved the final version of the manuscript.

Data availability

Due to ethical restrictions related to patient confidentiality, the data are not publicly available but are available from the corresponding authors upon reasonable request.

Acknowledgements

MID (Medical Illustration & Design), as a member of the Medical Research Support Services of Yonsei University College of Medicine, providing excellent support with medical illustration.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2025.101713.

References

Author names in bold designate shared co-first authorship

[1] **Wong VW**, **Ekstedt M**, Wong GL, et al. Changing epidemiology, global trends and implications for outcomes of NAFLD. J Hepatol 2023;79(3):842–852.
[2] Anstee QM, Reeves HL, Kotsiliti E, et al. From NASH to HCC: current concepts and future challenges. Nat Rev Gastroenterol Hepatol 2019;16(7):411–428.
[3] Taylor RS, Taylor RJ, Bayliss S, et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. Gastroenterology 2020;158(6):1611–25. e12.
[4] Stern C, Castera L. Identification of high-risk subjects in nonalcoholic fatty liver disease. Clin Mol Hepatol 2023;29(Suppl):S196.
[5] Sanyal AJ, Van Natta ML, Clark J, et al. Prospective study of outcomes in adults with nonalcoholic fatty liver disease. N Engl J Med 2021;385(17):1559–1569.
[6] Mantovani A, Csermely A, Petracca G, et al. Non-alcoholic fatty liver disease and risk of fatal and non-fatal cardiovascular events: an updated systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2021;6(11):903–913.
[7] Rockey DC, Caldwell SH, Goodman ZD, et al. Liver biopsy. Hepatology 2009;49(3):1017–1044. https://doi.org/10.1002/hep.22742.
[8] Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 2007;45(4):846–854.
[9] Panel CPG, Berzigotti A, Tsochatzis E, et al. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis–2021 update. J Hepatol 2021;75(3):659–689.
[10] Nakatsuka T, Tateishi R, Sato M, et al. Agile scores are a good predictor of liver-related events in patients with NAFLD. J Hepatol 2023;79(3):e126–e127.
[11] Sanyal AJ, Foucquier J, Younossi ZM, et al. Enhanced diagnosis of advanced fibrosis and cirrhosis in individuals with NAFLD using FibroScan-based Agile scores. J Hepatol 2023;78(2):247–259.
[12] Calès P, Ravaioli F, Berger A, et al. Comparison of screening strategies with two new tests to score and diagnose varices needing treatment. Clin Res Hepatol Gastroenterol 2022;46(7):101925.
[13] **Chen Y**, **Du X**, Kuppa A, et al. Genome-wide association meta-analysis identifies 17 loci associated with nonalcoholic fatty liver disease. Nat Genet 2023;55(10):1640–1650.
[14] Kawaguchi T, Sumida Y, Umemura A, et al. Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese. PLoS One 2012;7(6):e38322.
[15] Romeo S, Kozlitina J, Xing C, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40(12):1461–1465.
[16] Xu R, Tao A, Zhang S, et al. Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. Sci Rep 2015;5:9284.
[17] **Kim DY**, **Park JY**. Genetic risk factors associated with NAFLD. Hepatoma Res 2020;6(85):10–20517.
[18] Lohmueller KE, Pearce CL, Pike M, et al. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 2003;33(2):177–182.
[19] Ghouse J, Sveinbjörnsson G, Vujkovic M, et al. Integrative common and rare variant analyses provide insights into the genetic architecture of liver cirrhosis. Nat Genet 2024;56(5):827–837. https://doi.org/10.1038/s41588-024-01720-y.
[20] **Sveinbjornsson G**, **Ulfarsson MO**, Thorolfsdottir RB, et al. Multiomics study of nonalcoholic fatty liver disease. Nat Genet 2022;54(11):1652–1663. https://doi.org/10.1038/s41588-022-01199-5.
[21] Cusi K, Abdelmalek MF, Apovian CM, 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 2025;48(7):1057–1082. https://doi.org/10.2337/dci24-0094.
[22] Lee CH, Lui DT, Lam KS. Non-alcoholic fatty liver disease and type 2 diabetes: an update. J Diabetes Investig 2022;13(6):930–940. https://doi.org/10.1111/jdi.13756.
[23] **Zhang X**, **Yip TC**, Tse YK, et al. Duration of type 2 diabetes and liver-related events in nonalcoholic fatty liver disease: a landmark analysis. Hepatology 2023;78(6):1816–1827. https://doi.org/10.1097/hep.0000000000000432.
[24] Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41(6):1313–1321.
[25] Wong VWS, Vergniol J, Wong GLH, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. Hepatology 2010;51(2):454–462.
[26] Wong VW-S, Adams LA, de Lédinghen V, et al. Noninvasive biomarkers in NAFLD and NASH—current progress and future promise. Nat Rev Gastroenterol Hepatol 2018;15(8):461–478.
[27] Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer International Publishing; 2019.
[28] Harrell FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer International Publishing; 2015.
[29] Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. Springer; 2009.
[30] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
[31] Pepe MS, Kerr KF, Longton G, et al. Testing for improvement in prediction model performance. Stat Med 2013;32(9):1467–1482. https://doi.org/10.1002/sim.5727.
[32] Vickers AJ, Cronin AM, Begg CB. One statistical test is sufficient for assessing new predictive markers. BMC Med Res Methodol 2011;11:13. https://doi.org/10.1186/1471-2288-11-13.
[33] Sadatsafavi M, Adibi A, Puhan M, et al. Moving beyond AUC: decision curve analysis for quantifying net benefit of risk prediction models. Eur Respir J 2021;58(5):2101186. https://doi.org/10.1183/13993003.01186-2021.
[34] Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn prognostic Res 2019;3:1–8.
[35] Romeo S, Sentinelli F, Dash S, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. Int J Obes (Lond) 2010;34(1):190–194. https://doi.org/10.1038/ijo.2009.216.
[36] Giudice EM, Grandone A, Cirillo G, et al. The association of PNPLA3 variants with liver enzymes in childhood obesity is driven by the interaction with abdominal fat. PLoS One 2011;6(11):e27933. https://doi.org/10.1371/journal.pone.0027933.
[37] Stender S, Kozlitina J, Nordestgaard BG, et al. Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci. Nat Genet 2017;49(6):842–847. https://doi.org/10.1038/ng.3855.

[38] Chen VL, Oliveri A, Miller MJ, et al. PNPLA3 genotype and diabetes identify patients with nonalcoholic fatty liver disease at high risk of incident cirrhosis. Gastroenterology 2023;164(6):966–77. e17.
[39] Chalasani N, Vilar-Gomez E, Loomba R, et al. PNPLA3 rs738409, age, diabetes, sex, and advanced fibrosis jointly contribute to the risk of major adverse liver outcomes in metabolic dysfunction-associated steatotic liver disease. Hepatology 2024;80(5):1212–1226. https://doi.org/10.1097/hep.0000000000000896.
[40] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26(6):565–574. https://doi.org/10.1177/0272989x06295361.
[41] Koo BK, Joo SK, Kim D, et al. Development and validation of a scoring system, based on genetic and clinical factors, to determine risk of steatohepatitis in Asian patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2020;18(11):2592–9. e10.
[42] **Bianco C**, **Jamialahmadi O**, **Pelusi S**, et al. Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. J Hepatol 2021;74(4):775–782. https://doi.org/10.1016/j.jhep.2020.11.024.
[43] Gellert-Kristensen H, Richardson TG, Davey Smith G, et al. Combined effect of PNPLA3, TM6SF2, and HSD17B13 variants on risk of cirrhosis and hepatocellular carcinoma in the general population. Hepatology 2020;72(3):845–856. https://doi.org/10.1002/hep.31238.
[44] Díaz LA, Alazawi W, Agrawal S, et al. High inherited risk predicts age-associated increases in fibrosis in patients with MASLD. J Hepatol 2025. https://doi.org/10.1016/j.jhep.2025.04.035.
[45] **De Vincentis A**, **Tavaglione F**, Jamialahmadi O, et al. A polygenic risk score to refine risk stratification and prediction for severe liver disease by clinical fibrosis scores. Clin Gastroenterol Hepatol 2022;20(3):658–673. https://doi.org/10.1016/j.cgh.2021.05.056.
[46] Karlsen TH, Sheron N, Zelber-Sagi S, et al. The EASL–Lancet Liver Commission: protecting the next generation of Europeans against liver disease complications and premature mortality. Lancet 2022;399(10319):61–116.
[47] Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology 2023;77(5):1797–1835.
[48] **Graupera I**, **Thiele M**, Serra-Burriel M, et al. Low accuracy of FIB-4 and NAFLD fibrosis scores for screening for liver fibrosis in the population. Clin Gastroenterol Hepatol 2022;20(11):2567–2576. e6.
[49] Hagström H, Talbäck M, Andreasson A, et al. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. J Hepatol 2020;73(5):1023–1029.
[50] Fabbrini E, Rady B, Koshkina A, et al. Phase 1 trials of PNPLA3 siRNA in I148M homozygous patients with MAFLD. N Engl J Med 2024;391(5):475–476.

Keywords: Metabolic Dysfunction-Associated Steatotic Liver Disease; Genotypes; Advanced Fibrosis; Non-Invasive Tests; Decision Curve Analysis; Type 2 diabetes mellitus.

Received 10 April 2025; received in revised form 20 November 2025; accepted 3 December 2025; Available online 17 December 2025