




A Genetic and Metabolic Staging System for Predicting the Outcome of Nonalcoholic Fatty Liver Disease

Grazia Pennisi ^{1*}, Rosaria Maria Pipitone,^{1*} Marco Enea,^{2,3} Antonio De Vincentis,⁴ Salvatore Battaglia,² Vito Di Marco,¹ Vincenzo Di Martino,¹ Federica Spatola,¹ Federica Tavaglione,^{4,5} Umberto Vespasiani-Gentilucci,² Rossella Zito,¹ Stefano Romeo ⁵, Calogero Cammà,^{1**} Antonio Craxi,^{1**} Stefania Grimaudo,^{1**} and Salvatore Petta ^{1**}

Nonalcoholic fatty liver disease (NAFLD) is an emerging cause of liver-related events (LREs). Here, we have assessed the ability of a composite score based on clinical features, metabolic comorbidities, and genetic variants to predict LREs. A total of 546 consecutive patients with NAFLD were recruited and stratified according to the fibrosis-4 (FIB-4) index. LREs were defined as occurrence of hepatocellular carcinoma or hepatic decompensation. Cox regression multivariate analysis was used to identify baseline variables associated with LREs. The UK Biobank was used as the validation cohort, and severe liver disease (incidence of cirrhosis, decompensated liver disease, hepatocellular carcinoma, and/or liver transplantation) was used as the outcome. LREs were experienced by 58 patients, only one of whom was in the cohort of patients with a FIB-4 score < 1.3. Multivariate Cox regression analysis of 229 patients with a FIB-4 score ≥ 1.3 highlighted clinical variables independently associated with the development of LREs, including older age, low platelet count, low albumin, low high-density lipoprotein cholesterol, certain genetic factors, and interactions between genetic factors and sex or diabetes. The area under the curve (AUC) for the model was 0.87 at 1, 3, and 5 years. Our novel Genetic and Metabolic Staging (GEMS) scoring system was derived from the Cox model linear predictor, ranked from 0 to 10, and categorized into five classes (0-5, 5-6, 6-7, 7-8, and 8-10). The risk of LREs increased from 4% in patients in the best class (GEMS score 0-5) to 91% in the worst (GEMS score 8-10). GEMS score was associated with incident severe liver disease in the study population (hazard ratio, 1.56; 95% confidence interval, 1.48-1.65; $P < 0.001$) as well as in the UK Biobank cohort where AUCs for prediction of severe liver disease at 1, 3, and 5 years were 0.70, 0.69, and 0.67, respectively. **Conclusion:** The novel GEMS scoring system has an adequate ability to predict the outcome of patients with NAFLD. (*Hepatology Communications* 2022;6:1032-1044).

Nonalcoholic fatty liver disease (NAFLD) is currently the most relevant liver disease worldwide, affecting about 25% of the general population⁽¹⁾ and leading to an increased risk of liver-related events (LREs), such as liver decompensation (LD) and hepatocellular carcinoma (HCC). Liver fibrosis is the main driver of these complications.⁽²⁻⁵⁾ The identification of risk factors and/or noninvasive

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, body mass index; CI, confidence interval; FIB-4, fibrosis-4; FLI, fatty liver index; GEMS, Genetic and Metabolic Staging; HCC, hepatocellular carcinoma; HDL, high-density lipoprotein; HR, hazard ratio; HSD17B13, hydroxysteroid 17-beta dehydrogenase 13; ICD-10, International Classification of Diseases, Tenth Revision; LD, liver decompensation; LRE, liver-related event; NAFLD, nonalcoholic fatty liver disease; PLT, platelet; PNPLA3, patatin-like phospholipase domain-containing protein 3; SLD, severe liver disease; TM6SF2, transmembrane 6 superfamily member 2; UKBB, UK Biobank.

Received May 31, 2021; accepted October 30, 2021.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1877/supinfo.

*These authors contributed equally to this work.

**These authors contributed equally to this work.

© 2021 The Authors. *Hepatology Communications* published by Wiley Periodicals LLC on behalf of American Association for the Study of Liver Diseases. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep4.1877

Potential conflict of Interest: Nothing to report.

scores predicting liver-related outcomes is helpful for stratifying the risk of LRE development.

Common noninvasive scores, like the NAFLD fibrosis score and the fibrosis-4 (FIB-4) index, are based on demographic, clinical, and biochemical parameters and were originally developed to stage liver fibrosis. These scores have also been validated as useful tools to stratify the risk of LRE occurrence in patients with NAFLD.⁽⁶⁻⁸⁾ However, hepatic complications occur only in some patients defined by these scores as at higher risk of developing LRE, and predictors able to accurately stratify the risk are lacking. Therefore it may be useful to consider individual components of metabolic syndrome (e.g., type 2 diabetes, low high-density lipoprotein [HDL] levels, high triglyceride levels, obesity, and arterial hypertension), many of which have been associated with poor clinical outcomes in patients with NAFLD.⁽⁹⁻¹¹⁾ Furthermore, genome-wide association studies and candidate gene studies have identified common gene variants, such as patatin-like phospholipase domain-containing protein 3 (*PNPLA3*) rs738409 C>G, transmembrane 6 superfamily member 2 (*TM6SF2*) rs58542926 C>T, and hydroxysteroid 17-beta dehydrogenase 13 (*HSD17B13*) rs72613567 T>TA, that have been associated with presence and severity of NAFLD, presence of HCC, and occurrence of LREs.⁽¹²⁻¹⁴⁾ A recent study proposed a polygenic score based on a combination of these gene variants to predict the risk of cirrhosis and HCC in the general population.⁽¹⁵⁾

In the context of this complex scenario, clinical, metabolic, and genetic features are involved in the development of NAFLD and in the increased risks of disease progression and liver complications. However, to date no studies have assessed the combined effects

of these factors on predicting outcomes of patients with NAFLD, and especially on stratifying the risk in patients with NAFLD with advanced liver fibrosis. The present study aimed to elaborate a composite score, including clinical features, metabolic comorbidities, and genetic background, to predict the occurrence of LREs in patients with NAFLD.

Patients and Methods

PALERMO COHORT

We analyzed data from 546 patients prospectively enrolled at the Gastrointestinal and Liver Unit of Palermo University Hospital. These patients had a histologic diagnosis of NAFLD or clinical diagnosis of advanced fibrosis/cirrhosis due to NAFLD. Specifically, in patients without histology, advanced fibrosis/cirrhosis was diagnosed by liver stiffness measurement (>11.5 kPa for M probe or >11 kPa for XL probe), and a diagnosis of NAFLD required the presence of ultrasonography-assessed steatosis plus at least one criterion of metabolic syndrome (obesity, diabetes, arterial hypertension, or dyslipidemia). Patients were included if they had blood samples available for genetic analysis and a follow-up of at least 6 months. Other causes of liver disease were ruled out, including alcohol intake (>20 g/day) as evaluated by a questionnaire, viral and autoimmune hepatitis, hereditary hemochromatosis, and alpha-1 antitrypsin deficiency. Patients with decompensated cirrhosis, HCC, and current use of steatosis-inducing drugs were excluded. The study was carried out in accordance with the principles of the Helsinki Declaration and with local and national

ARTICLE INFORMATION:

From the ¹Section of Gastroenterology and Hepatology, Dipartimento Di Promozione Della Salute, University of Palermo, Palermo, Italy; ²Dipartimento di Scienze Economiche, Aziendali e Statistiche, University of Palermo, Palermo, Italy; ³Dipartimento di Promozione della Salute, Materno Infantile, Medicina Interna e Specialistica di Eccellenza (PROMISE), University of Palermo, Palermo, Italy; ⁴Clinical Medicine and Hepatology Unit, Department of Internal Medicine and Geriatrics, Campus Bio-Medico University, Rome, Italy; ⁵Department of Molecular and Clinical Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Salvatore Petta, M.D., Ph.D.
Section of Gastroenterology and Hepatology
PROMISE, Policlinico Universitario Paolo Giaccone
Piazza delle Cliniche, 2

90127 Palermo, Italy
E-mail: salvatore.petta@unipa.it
Tel.: +39 091 23890673

law. Approval was obtained from the Comitato Etico Palermo 1 (ID 2014).

Clinical and metabolic data were collected at the time of enrollment. Body mass index (BMI) was calculated in kilograms for weight and in meters for height. Obesity was defined as BMI ≥ 30 kg/m². A diagnosis of type 2 diabetes was made according to standards set by the American Diabetes Association,⁽¹⁶⁾ using a value of fasting blood glucose ≥ 126 mg/dL and/or use of medications. High triglyceride levels were defined as serum triglycerides >150 mg/dL, and low HDL levels were defined as <40 mg/dL in male patients and <50 mg/dL in female patients.⁽¹⁷⁾ Statin use was also recorded. Arterial hypertension was defined as systolic blood pressure ≥ 140 mm Hg and/or diastolic blood pressure ≥ 90 mm Hg or use of blood pressure-lowering agents.⁽¹⁸⁾

At baseline, a 12-hour overnight fasting blood sample was drawn to determine serum levels of aspartate aminotransferase (AST), alanine aminotransferase (ALT), platelets (PLTs), albumin, total cholesterol, HDL cholesterol, triglycerides, and plasma glucose concentration. Serum tests were obtained at liver biopsy or at the day of clinical diagnosis of advanced fibrosis/cirrhosis. The FIB-4 scoring system combines age, AST, ALT, and PLTs⁽¹⁹⁾ and was used to stratify patients as having low, intermediate, or high risk for advanced fibrosis (<1.30 , 1.30 - 2.67 , and >2.67 , respectively).

The Kleiner scoring system⁽²⁰⁾ was used for histologic assessment of NAFLD and specifically to grade steatosis, lobular inflammation, and hepatocellular ballooning and to stage fibrosis from 0 to 4. Nonalcoholic steatohepatitis was defined by the presence of steatosis, ballooning, and lobular inflammation at once.

LREs were recorded during the entire follow-up period and were defined as development of LD (occurrence of ascites and/or bleeding varices and/or encephalopathy and/or jaundice) or HCC. Patients who developed LREs during follow-up were evaluated for available therapies and/or for liver transplantation, if appropriate.^(21,22) Ultrasound examination was employed every 6 months as stricter surveillance for HCC in patients with F3 fibrosis or cirrhosis, as per international guidelines.⁽²¹⁾ Patients with F0-F2 fibrosis underwent annual ultrasound follow-up. In the presence of cirrhosis, upper gastrointestinal endoscopy was performed at baseline and repeated

as recommended by clinical guidelines. Patients with progression to medium or large (F2 or F3) esophageal varices were treated with beta-blockers or elastic banding, whereas no prophylaxis was scheduled for patients with small (F1) varices.

Genotyping for *PNPLA3* rs738409 C>G, *TM6SF2* rs58542926 C>T, and *HSD17B13* rs72613567 T>TA were performed using the TaqMan single-nucleotide polymorphism genotyping allelic discrimination method (Applied Biosystems, Foster City, CA). Genotype calls were made by Sequence Detection System software v.2.3 (StepOne Plus; Applied Biosystems). Genotyping was conducted in a blinded fashion with regard to patient characteristics.

UK BIOBANK COHORT

Detailed study design and methods of the UK Biobank (UKBB) have been described.⁽²³⁾ Briefly, the UKBB is a large, prospective, cohort study including approximately 500,000 participants (aged 40-69 years) recruited from 2006 to 2010 in 22 assessment centers throughout the United Kingdom. The UKBB study was approved by the North West Multicenter Research Ethics Committee (reference number 11/NW/0382). Participants were identified from National Health Service patient registers and provided informed consent to the study. Health-related information and laboratory and genetic data were collected using highly standardized procedures.

For use in the present study, we excluded individuals with 1) self-reported history or hospital diagnosis of viral hepatitis, severe liver disease (SLD, see definition later in this article), or other causes of liver disease (International Classification of Diseases, Tenth Revision [ICD-10] codes B18, B19, C22.0, E83.0, E83.1, I85.0, I85.9, K70.3, K70.4, K70.9, K71, K72.1, K72.9, K74.1, K74.2, K74.3, K74.4, K74.5, K74.6, K75.2, K75.3, K75.4, K75.8, K75.9, K76.6, K76.7, K76.8, K76.9, R18, Z94.4; Supporting Table S1); 2) self-reported history or diagnosis from the cancer register of liver cancer (ICD-10 C22); 3) excessive alcohol consumption (self-reported; ≥ 30 g/day and ≥ 20 g/day for men and women, respectively). We also excluded individuals who reported non-European ancestry, withdrew consent for any reason, or were missing data for any Genetic and Metabolic Staging (GEMS) variable (described below). Our final analyses included 303,075 individuals, of whom 134,177 were at intermediate or

high risk for advanced liver fibrosis as assessed by FIB-4. A diagnosis of steatosis was made for individuals scoring ≥ 60 on the fatty liver index (FLI), a validated system for the noninvasive diagnosis of fatty liver.⁽²⁴⁾

Genotyping and arrays used in the UKBB study have been described in detail.⁽²⁵⁾ Genotype data were available for approximately 490,000 participants. *PNPLA3* rs738409 C>G (p.I148M), *TM6SF2* rs58542926 C>T (p.E167K), and *HSD17B13* rs72613567:TA were assayed using two similar genotyping arrays (Affymetrix UK BiLEVE and UKBB Axiom arrays) and coded as 0, 1, or 2 for noncarriers, heterozygous carriers, and homozygous carriers of the minor allele, respectively.

Follow-up data on health-related events and mortality were obtained through linkage of the National Health Service records, including in-hospital admissions as well as death and cancer registers (UKBB data fields 41270, 40001, 40002, and 40006). We defined SLD as a composite diagnosis of cirrhosis, decompensated liver disease (esophageal varices with or without bleeding, portal hypertension, hepatorenal syndrome, liver failure), HCC, and/or liver transplantation (ICD-10 C22.0, I85.0, I85.9, K70.3, K70.4, K72.1, K72.9, K74.1, K74.2, K74.6, K76.6, K76.7, Z94.4; Supporting Table S1) in any of the above-mentioned records.

Individuals were excluded if they had received a hospital diagnosis of viral hepatitis or other causes of liver disease (ICD-10 B18, B19, E83.0, E83.1, K71, K74.3, K74.4, K74.5, K75.2, K75.3, K75.4, K75.8, K75.9; Supporting Table S1) before the diagnosis of SLD.

The length of follow-up for each participant was determined from the date of baseline assessment until either the date of SLD diagnosis, date of death, or date of end of follow-up for the assessment center attended (January 31, 2018), whichever occurred first.

STATISTICAL ANALYSES

Data are reported as mean \pm SD for continuous variables and as percentages for categorical variables. The statistical tests applied were *t* test, analysis of variance, and chi-squared test, as appropriate. All analyses were performed using R statistical software version 3.6.1. A Cox proportional hazards model was used to evaluate the effect of metabolic, clinical, and genetic factors on the risk of LRE occurrence.⁽²⁶⁾

We considered $P < 0.05$ as statistically significant. Confidence intervals (CIs) of hazard ratios (HRs) were provided at a 95% significance level. The proportionality of risk factors was evaluated by Schoenfeld residuals.⁽²⁷⁾

The best model was chosen after a backward selection on the basis of both the likelihood ratio test and the Akaike information criterion,⁽²⁸⁾ including factors such as sex, diabetes status, age; levels of HDL, albumin, and PLTs; and presence of *PNPLA3*, *HSD17B13*, and *TM6SF2*. The most complete models we considered included an interaction term between each of the individual genetic variables and all the clinical and metabolic variables, but those interaction terms involving three-level factors (i.e., the categorized version of age and PLTs) that presented convergence problems. On the basis of the best model, we created a genetic–metabolic–clinical risk score of LRE occurrence. The first step in score construction was to consider the linear predictor model, a weighted sum of the risk factors with weights given by the estimated coefficients. The theoretical range of this score varies from zero (lowest risk; each variable is set to zero) to 10 (highest risk; each variable is set to one). The second and final step was to discretize the model linear predictor to create five different risk profiles as follows: 0, 0-5; 1, 5-6; 2, 6-7; 3, 7-8; 4, 8-10. The accuracy of the scoring predictions was evaluated in terms of homogeneity, discrimination, monotonicity of the gradient, and calibration. The diagnostic accuracy of GEMS was validated on the UKBB cohort, and all analyses were conducted in the overall cohort as well as in high-risk subgroups for SLD, namely those with FIB-4 ≥ 1.3 , FLI ≥ 60 ,⁽²⁴⁾ and diabetes. The estimated probability of LREs was derived from the Kaplan-Meier estimator.

Results

PATIENT FEATURES

Baseline characteristics of the 546 patients with NAFLD are shown in Supporting Table S2. The diagnosis of NAFLD was supported by histology in 456 cases (83.5%), while advanced fibrosis/cirrhosis was clinically diagnosed in 90 cases (16.5%). Specifically, 78 patients with advanced fibrosis/cirrhosis had a liver stiffness measurement >15 kPa⁽²⁹⁾ and

12 were between 11 kPa and 15 kPa (10 of these cases also had esophageal varices and the other 2 had signs of portal hypertension [splenomegaly and portal vein ectasia]). A total of 112 patients (20.5%) took statin and 159 (29.1%) took metformin.

We found that 317 patients had FIB-4 < 1.3, indicating low risk of advanced liver fibrosis, while 229 patients had FIB-4 ≥ 1.3, indicating intermediate to high risk. Baseline features of patients stratified for FIB-4 are summarized in Table 1.

OCCURRENCE OF LRES

During a median follow-up of 73.8 months, 58 patients of the entire cohort experienced LREs, with a likelihood of 1.9%, 6.2%, and 9.7% at 1, 3, and 5 years, respectively. We recorded 53 events of LD and 15 of HCC. When patients were stratified according to FIB-4 score, 57 of the patients experiencing LREs were in the cohort with FIB-4 ≥ 1.3 and only 1 was in the cohort with FIB-4 < 1.3 (Fig. 1). Consistent with the low risk of LREs, patients with FIB-4 < 1.3 were excluded from further analyses.

RISK FACTORS ASSOCIATED WITH LRES IN PATIENTS WITH FIB-4 ≥ 1.3

During a median follow-up of 66.1 months, 57 patients experienced LREs, with a likelihood of 4.5%, 14.8%, and 23.6% at 1, 3, and 5 years, respectively. We recorded 52 events of LD and 15 of HCC.

In the multivariate Cox regression analysis, the following clinical variables were independently associated with LRE occurrence: age between 55 and 65 years (HR, 13.96; 95% CI, 2.9-67.23; $P = 0.001$); age older than 65 years (HR, 17.96; 95% CI, 3.66-88.12; $P < 0.001$); PLTs between 110,000 and 150,000/mm³ (HR, 6.89; 95% CI, 2.74-17.35; $P < 0.001$); PLTs < 110,000/mm³ (HR, 13.54; 95% CI, 5.53-33.13; $P < 0.001$); albumin < 4 g/L (HR, 1.5; 95% CI, 1.00-3.78; $P = 0.04$); metabolic variables, such as low HDL cholesterol (HR, 1.88; 95% CI, 1.02-3.44; $P = 0.04$); and genetic variables, such as *TM6SF2* rs58542926 CT/TT (HR, 1.94; 95% CI, 1.00-3.77; $P = 0.04$) and *HSD17B13* rs72613567 T/TA (HR, 1.83; 95% CI, 1.07-3.43; $P = 0.04$) (Table 2). Interactions between *PNPLA3* rs738409 and male sex (HR, 0.32; 95% CI, 0.10-0.98; $P = 0.04$) and between *PNPLA3* rs738409

and diabetes (HR, 5.16; 95% CI, 1.30-20.41; $P = 0.01$) were also independently associated with the development of LREs (Table 2). The generated Cox regression model showed good diagnostic accuracy for predicting LREs at 1, 3, and 5 years (area under the curve [AUC], 0.87; Fig. 2). When genetic variables and their interaction products were excluded from the model, the diagnostic accuracy was significantly lower (0.83 at 1 year and 0.84 at 3 and 5 years; chi-squared, 17.27 [degrees of freedom, 8]; $P = 0.02$). The AUCs of the model were significantly higher than those for the FIB-4 score (AUC, 0.68, 0.68, and 0.70 at 1, 3, and 5 years, respectively; $P < 0.001$ for all). Moreover, we compared the calibration of both GEMS and FIB-4 using the Hosmer-Lemeshow test. A good calibration resulted for GEMS ($P = 0.5$), while we found FIB-4 was not well calibrated ($P < 0.001$).

Considering the potential impact of both metformin and statin use on LRE risk, we did additional analyses, including both variables, but neither statin (HR, 1.11; $P = 0.73$ in univariate analysis and HR, 1.29; $P = 0.45$ in multivariate analysis) nor metformin (HR, 0.58; $P = 0.09$ in univariate analysis and HR, 0.73; $P = 0.39$ in multivariate analysis) showed significant effects.

When considering LD and HCC separately, the Cox regression multivariate model maintained high-predictive accuracy (LD AUC of 0.84, 0.85, and 0.86 at 1, 3, and 5 years, respectively; HCC AUC of 0.83 at 1 year and 0.84 at 3 and 5 years) (Supporting Figs. S1 and S2).

THE GEMS SYSTEM

Consistent with the clinical, metabolic, and genetic variables significantly associated with LREs in the Cox multivariate regression model, we created GEMS, a simple score for predicting LREs in patients with NAFLD and at risk for liver fibrosis by FIB-4. GEMS was calculated from the weighted sum of the risk factors according to the following formula: $1.163 - 0.438(\text{PNPLA3 CG/GG}) + 0.421(\text{male sex}) - 0.413(\text{diabetes}) + 2.635(55 \leq \text{age} < 65) + 2.888(\text{age} > 65) + 0.632(\text{low HDL}) + 0.668(\text{albumin} < 4 \text{ g/dL}) + 1.935(110,000/\text{mm}^3 < \text{PLTs} < 150,000/\text{mm}^3) + 2.605(\text{PLTs} < 110,000/\text{mm}^3) + 0.602(\text{HSD17B13 TTA/TATA}) + 0.661(\text{TM6SF2 CT/TT}) - 1.146(\text{interaction PNPLA3 CG/GG and male sex}) + 1.641(\text{interaction PNPLA3 CG/GG and diabetes})$.

TABLE 1. BASELINE DEMOGRAPHIC, METABOLIC, LABORATORY, HISTOLOGIC, AND GENETIC FEATURES OF ENTIRE COHORT STRATIFIED FOR LOW RISK OF FIBROSIS (FIB-4 < 1.3) AND INTERMEDIATE/HIGH RISK OF FIBROSIS (FIB-4 ≥ 1.3)

	FIB-4 < 1.3 (n = 3 17)	FIB-4 ≥ 1.3 (n = 229)	PValue
Mean age, years	43 ± 12.3	61.6 ± 9.1	<0.001
Age ≥ 50 years	34.1%	90.8%	<0.001
Male sex	70%	56.8%	0.001
Mean BMI	29.74 ± 5	31.8 ± 5.7	0.34
Obesity (BMI ≥ 30 kg/m ²)	40.1%	61.6%	<0.001
Blood glucose, mg/dL	98.9 ± 33.3	113.8 ± 35.4	0.01
Total cholesterol, mg/dL	200 ± 43.4	180.2 ± 48	0.32
HDL, mg/dL	49.2 ± 13.9	48.8 ± 16.9	0.03
HDL <40 mg/dL in males	33.1%	38.4%	0.20
<50 mg/dL in females			
Triglycerides, mg/dL	146 ± 97.5	131 ± 63.1	0.004
Triglycerides ≥ 150 mg/dL	36.9%	28.4%	0.03
AST, IU/L	39.5 ± 20.2	49.1 ± 31.6	0.001
ALT, IU/L	78.1 ± 53.1	57.9 ± 41	0.02
Albumin, g/L	4.6 ± 0.3	4.2 ± 0.4	<0.001
Albumin ≥ 4 g/L	3.1%	25.3%	<0.001
PLTs, 10 ³ /mm ³	262.6 ± 77.7	160.3 ± 67.2	<0.001
PLTs ≥ 110 × 10 ³ /mm ³	0	21.4%	<0.001
Type 2 diabetes	22.7%	57.6%	<0.001
Arterial hypertension	29.3%	58.5%	<0.001
Statin users	17.9%	24.0%	0.11
Metformin users	30.0%	27.9%	0.68
<i>PNPLA3</i> rs738409 CC/CG/GG	35/45.7/19.2%	25.8/46.3/27.9%	0.01
<i>TM6SF2</i> rs58542926 CC/CT/TT	72.5/24.3/2.2%	72.9/24.4/2.6%	0.95
<i>HSD17B13</i> rs72613567 T/TA/TATA	65.6/30.6/3.8%	74.2/22.7/3.1%	0.09
Time of follow-up, months	79.4 ± 42.9	66.1 ± 37.3	0.03

Data are given as mean ± SD or as percentage (%) of cases.

In this formula, the value 1.163 is a normalization constant allowing the theoretical range of the score to vary from 0 to 10. Parentheses indicate an identity function as follows: if a variable within the brackets is true, then the corresponding term will score 1, otherwise it will be 0. Note that *PNPLA3* CG/GG, male sex, diabetes, low HDL, *HSD17B13* TTA/TATA, and *TM6SF2* CT/TT are all binary variables coded as 1 or 0, indicating the presence or absence of that characteristic (e.g., male sex = 0 means a female patient), while the numerical variables age, albumin, and PLTs are made binary by the comparison with the corresponding cutoffs. The interaction terms will be 1 when both variables involved are 1, otherwise the term will score 0. For example, for a patient profile having *PNPLA3* CG/GG, female sex, age 60, diabetes, low HDL, albumin 3.5 g/dL, PLTs 100,000 mm³, no *HSD17B13* TTA/TATA, and no *TM6SF2* CT/

TT, the score will be calculated as follows: 1.163 – 0.438 × 1 + 0.421 × 0 – 0.413 × 1 + 2.635 × 1 + 2.888 × 0 + 0.632 × 1 + 0.668 × 1 + 1.935 × 0 + 2.605 × 1 + 0.602 × 0 + 0.661 × 0 – 1.146 × 0 + 1.641 × 1 = 8.493.

The theoretical range of this score varied from 0 to 10, where 0 represented total absence of risk of LREs while 10 represented the highest risk of LREs. The GEMS score was categorized into the following five classes of risk: 0-5, 5-6, 6-7, 7-8, 8-10, as shown in Supporting Table S3.

GEMS scores categorized in five classes of risk provided a good prediction of LREs. The crude rate of LREs at the end of follow-up among GEMS risk classes is shown in Fig. 3. Notably, the risk increased from 4% in patients in the best class (GEMS 0-5) to 91% in patients in the worst class (GEMS 8-10). Consistent with this result, Kaplan-Meier analysis (Fig. 4) also showed that the overall cumulative



FIG. 1. LREs recorded during follow-up in entire cohort and in patients stratified according to FIB-4.

TABLE 2. MULTIVARIATE COX REGRESSION ANALYSIS OF CLINICAL, METABOLIC, AND GENETIC VARIABLES ASSOCIATED WITH LRES IN PATIENTS WITH NAFLD WITH FIB-4 ≥ 1.3

	HR (95% CI)	PValue
Male sex	1.52 (0.76-3.06)	0.23
Age 55-65 years	13.96 (2.90-67.23)	0.001
Age > 65 years	17.96 (3.66-88.12)	<0.001
PLTs 110,000-150,000/mm ³	6.89 (2.74-17.35)	<0.001
PLTs < 110,000/mm ³	13.54 (5.53-33.13)	<0.001
Albumin <4 g/L	1.95 (1.00-3.78)	0.04
Low HDL	1.88 (1.02-3.44)	0.04
Diabetes	0.66 (0.32-1.38)	0.27
PNPLA3 rs738409	0.64 (0.18-2.28)	0.49
HSD17B13 rs72613567	1.83 (1.07-3.43)	0.04
TM6SF2 rs58542926	1.94 (1.00-3.77)	0.04
PNPLA3 × sex	0.32 (0.10-0.98)	0.04
PNPLA3 × diabetes	5.16 (1.30-20.41)	0.01

rate of LREs progressively increased according to GEMS risk classes. In particular, the rate of LREs was lowest in patients with a GEMS score from 0 to 5 (1.1% at 1 and 3 years and 3.5% at 5 years) and highest in patients with a score from 8 to 10 (38.6%, 79.5%, and 89.8% at 1, 3, and 5 years, respectively) (Fig. 4).

INTERNAL VALIDATION

Internal validation, calibration, and discrimination tests were performed to validate the diagnostic accuracy of GEMS scoring. An appropriate calibration of the model (Brier score, 13; Hosmer-Lemeshow test, *P* = 0.5) is displayed in Supporting Fig. S3 and reflects good predictive accuracy (AUC, 85.1).

EXTERNAL VALIDATION IN THE UKBB COHORT

Baseline characteristics of the UKBB individuals stratified by incident SLD are reported in Supporting Table S4.

The GEMS score was significantly associated with incident SLD in the entire population (HR, 1.56; 95% CI, 1.48-1.65; *P* < 0.001) and in patients with FIB-4 > 1.3 (HR, 1.55; 95% CI, 1.43-1.67; *P* < 0.001) (Supporting Table S5). Similar results were observed in other subgroups of patients at risk for SLD, including those with FLI ≥ 60, with FLI ≥ 60 and FIB-4 > 1.3, with diabetes, or with diabetes and FIB-4 > 1.3 (Supporting Table S5). AUCs for GEMS prediction of SLD at 1, 3, and 5 years in the entire population were 0.70, 0.69, and 0.67, respectively (Supporting Table S6); AUCs for subgroups of patients at risk for liver disease are reported in Supporting Table S6.

Across all groups, the rate of SLD was lowest in patients in the best class (GEMS 0-5) and highest in those in the worst class (GEMS 8-10) (Fig. 5). Similarly, Kaplan-Meier analysis showed that the overall cumulative rate of SLD progressively and significantly increased according to GEMS classes of risk across all groups (Fig. 6).

Discussion

In a cohort of patients with a histologic diagnosis of NAFLD or with clinical diagnosis of advanced fibrosis/cirrhosis related to NAFLD, we found that a composite

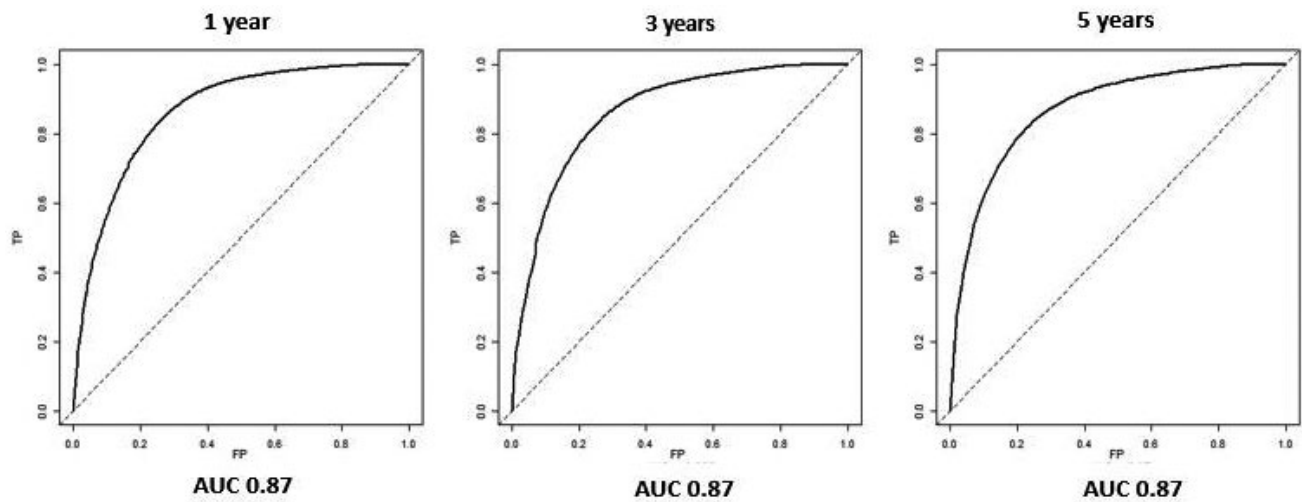


FIG. 2. ROC curves of the model, including clinical, metabolic, and genetic features, for prediction of LREs at 1, 3, and 5 years in patients with NAFLD and FIB-4 ≥ 1.3 .

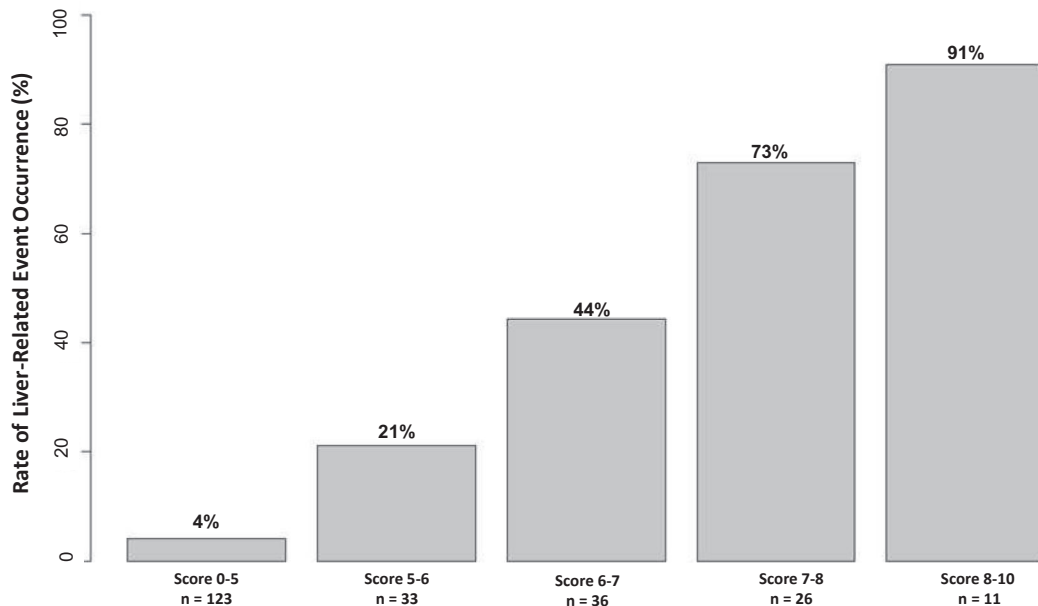


FIG. 3. Crude rate of LREs at the end of follow-up among GEMS risk classes.

score that included clinical, metabolic, and genetic variables had good accuracy for predicting LRE occurrence. This score also accurately stratified the risk of incident SLD in individuals from the UKBB population.

NAFLD is a growing cause of liver diseases leading to hepatic complications, such as LD and HCC. In our cohort, the overall rate of LRE occurrence was 1.9%, 6.2%, and 9.7% at 1, 3, and 5 years, respectively.

By stratifying patients by FIB-4 score, 57 patients with FIB-4 ≥ 1.3 experienced LREs whereas only 1 patient with FIB-4 < 1.3 developed LREs. The higher risk of LREs in patients with FIB-4 ≥ 1.3 confirms FIB-4 as an accurate tool not only to noninvasively stratify the risk of liver fibrosis but also to help predict LREs.⁽⁶⁻⁸⁾ In accordance with the available literature,⁽²⁻⁵⁾ our data emphasize the role of hepatic fibrosis

Score	Events (%)		
	1 year	3 year	5 year
0 - 5	1.1	1.1	3.5
5 - 6	6	10	19.5
6 - 7	8.4	34	40
7 - 8	12	45.5	74.8
8 - 10	38.6	79.5	89.8

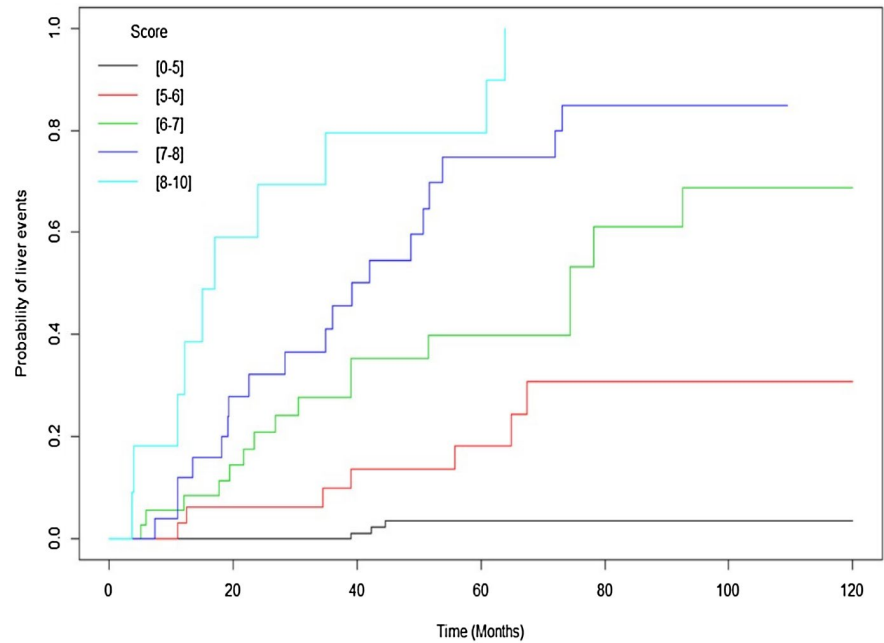


FIG. 4. Kaplan-Meier curves of overall cumulative rate of LREs according to GEMS risk classes.

as a main driver of the progression of liver disease and the development of liver-related complications.

In the present study, we found that a score based on a combination of clinical, metabolic, and genetic risk factors has good diagnostic accuracy for predicting the risk of LREs in patients with NAFLD and at intermediate to high risk of hepatic fibrosis as indicated by FIB-4 scoring. In a large US cohort of patients with NAFLD with low risk of advanced fibrosis (according to FIB-4), Kanwal et al.⁽³⁰⁾ showed that metabolic traits (obesity, hypertension, diabetes, and dyslipidemia) were significantly associated with development of cirrhosis and HCC and that the level of risk was proportional to the number of metabolic dysfunctions. Gellert-Kristensen et al.⁽¹⁵⁾ recently developed a polygenic risk score that included common variants of *PNPLA3*, *TM6SF2*, and *HSD17B13* genes. They reported that in a general population setting the risk of developing cirrhosis and HCC increased with the number of at-risk alleles. Another polygenic risk score, based on the combination of *PNPLA3*, *TM6SF2*, glucokinase regulator (*GCKR*), and membrane bound O-acyltransferase domain containing 7 (*MBOAT7*) risk variants and adjusted for *HSD17B13* status, predicted the presence of HCC in patients with NAFLD as well as in the UKBB general population, although with suboptimal accuracy.⁽³¹⁾

Together these studies reported that in different settings metabolic or genetic scores can help stratify the risk for liver cirrhosis and HCC. To the best of our knowledge, the present study is the first to demonstrate that the addition of genetic to other clinical and metabolic features could improve the prediction of liver outcomes in a cohort of patients with NAFLD and at moderate to high risk of fibrosis. The clinical–metabolic model without genetic variables had a significantly lower accuracy than the model including clinical, metabolic, and genetic variables (AUC of 0.87 at 1, 3, and 5 years when genetic variables were included as compared to AUC of 0.83 at 1 year and 0.84 at 3 and 5 years without genetic variables). Consistent with our results, a recent study on patients with a histologic diagnosis of alcohol-related liver disease showed that adding genetic data (*TM6SF2* rs58542926 and *PNPLA3* rs738409 variants) to clinical and metabolic traits improved accuracy for the prediction of liver fibrosis.⁽³²⁾

Calzadilla-Bertot et al.⁽³³⁾ recently developed a prognostic score that includes clinical (AST/ALT ratio, bilirubin, international normalized ratio) and metabolic (type 2 diabetes) parameters as well as esophageal varices (ABIDE). Although ABIDE showed good accuracy, it was designed to evaluate

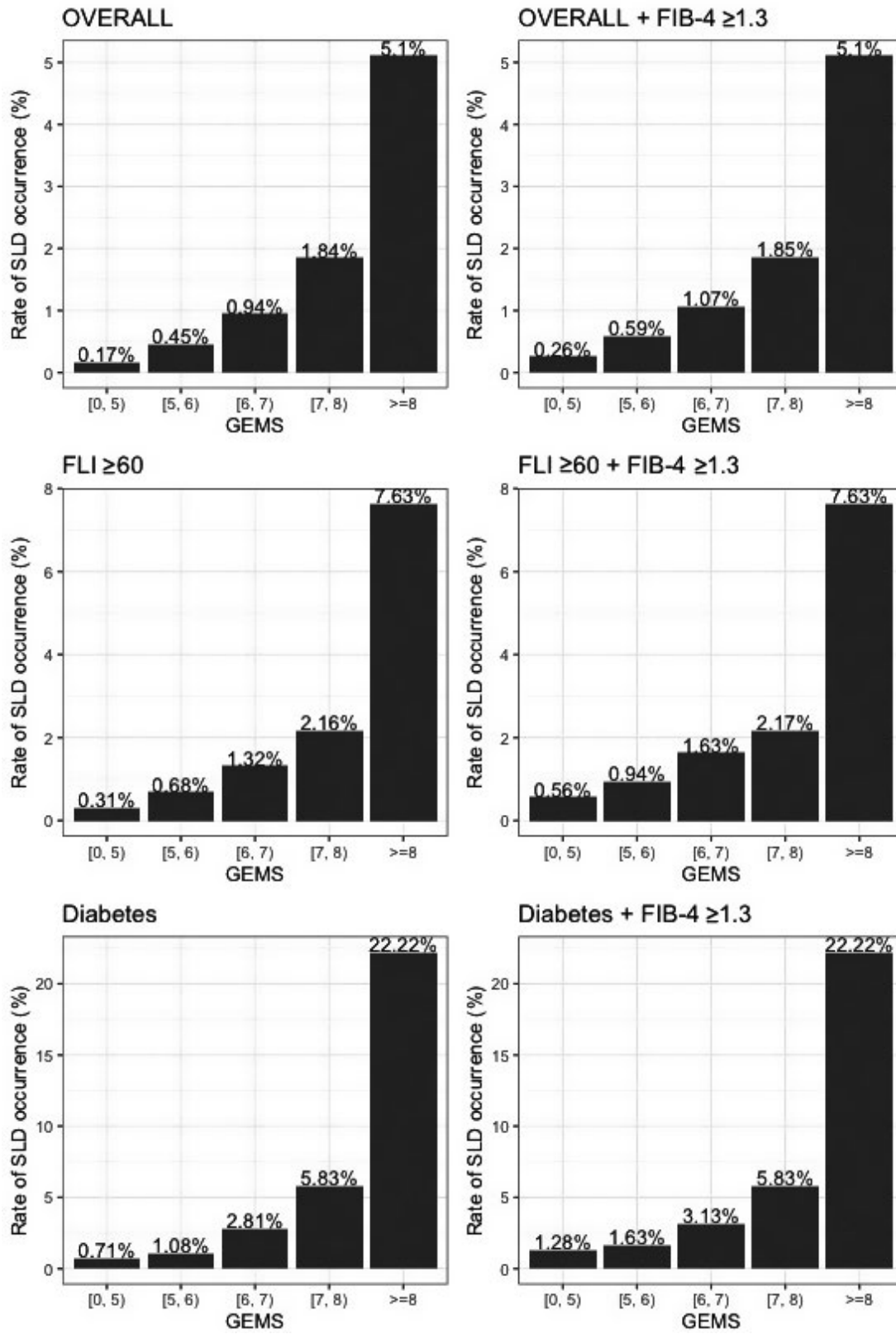


FIG. 5. Risk of incident SLD across GEMS classes in the overall UKBB cohort and in at-risk groups.

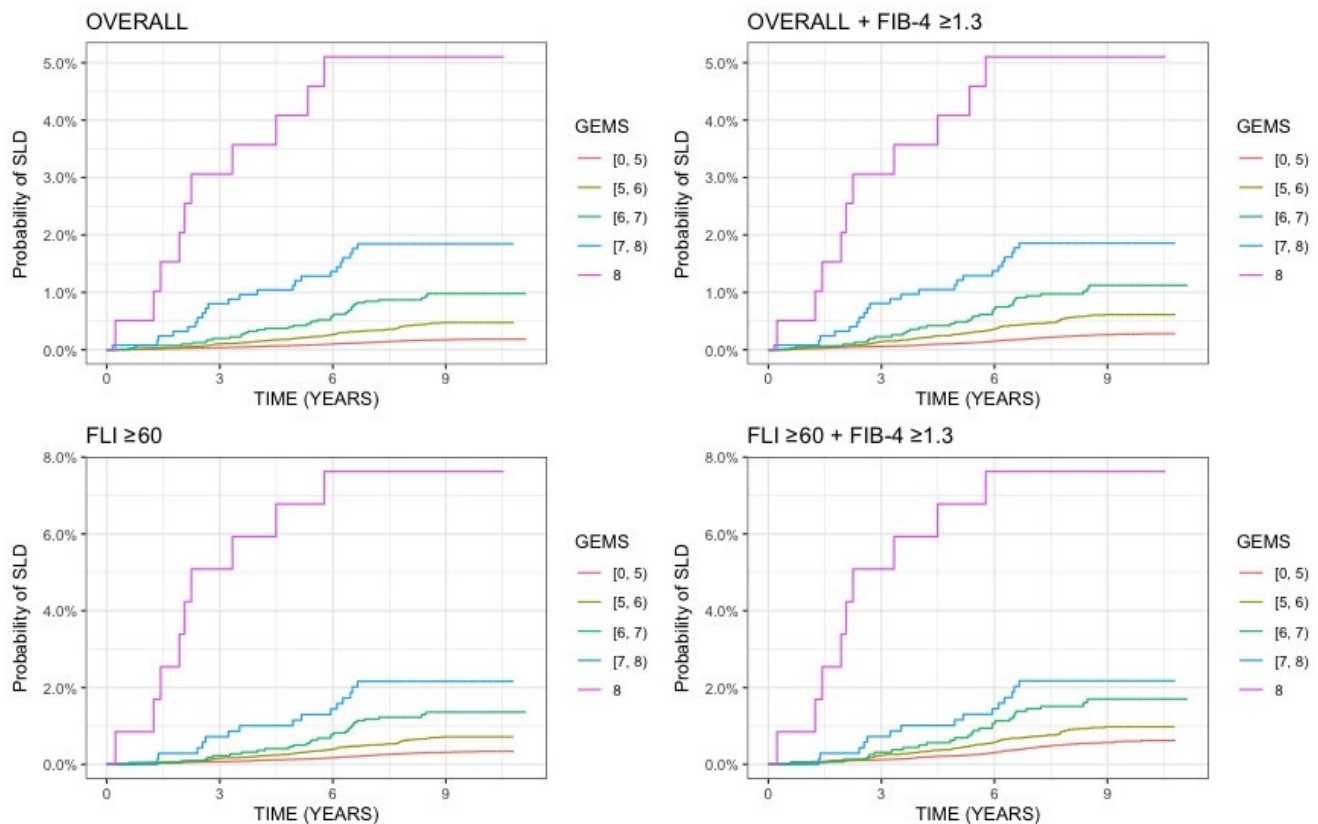


FIG. 6. Cumulative incidence of SLD during follow-up across GEMS classes in the overall UKBB cohort and in at-risk groups. Log-rank P value for trend $< 2 \times 10^{-16}$ for all plots.

only a single outcome (LD) in patients with compensated cirrhosis due to NAFLD. Notably, the GEMS score was able to predict LREs in a wider spectrum of patients with NAFLD without the need for invasive and expensive tests while also being able to predict separately the occurrence of both LD and HCC.

The present study also showed that interactions between the *PNPLA3* rs738409 C>G variant and both diabetes and sex significantly affect the risk of LRE occurrence. There was a positive interaction between the *PNPLA3* variant and diabetes in increasing the risk of LRE development. This finding is in line with previous evidence that the effect of the *PNPLA3* variant on liver damage is more pronounced in patients at higher metabolic risk⁽³⁴⁾ and with evidence suggesting that patients with diabetes carrying the *PNPLA3* G-allele had higher circulating free fatty acid concentrations and greater adipose tissue insulin resistance compared with noncarriers.⁽³⁵⁾ The present study also found a negative interaction between the *PNPLA3*

variant and male sex on the risk of developing LREs. This finding agrees with evidence from a meta-analysis of patients with biopsy-proven NAFLD that showed a negative correlation between male sex and the effect of the *PNPLA3* rs738409 variant on liver fat content.⁽³⁶⁾ The *PNPLA3* gene is under the control of several factors, including components of the sterol regulatory element binding protein (SREBP) pathway.⁽³⁷⁾ It should be noted that sex hormones regulate the expression of lipogenic genes, including SREBP-1c, and share the control of fat homeostasis.⁽³⁸⁾

From a clinical point of view, GEMS is a simple score with good accuracy for stratifying the risk of LREs in patients with NAFLD at intermediate to high risk of advanced liver fibrosis as measured by FIB-4. Patients with the lowest GEMS scores had the lowest likelihood of developing LREs, whereas those with the highest GEMS scores had the highest likelihood. Because we only observed a single patient with FIB-4 < 1.3 who experienced LREs, that particular

population would not appear to benefit much from scoring predictions. To minimize the risk of missing patients that will develop liver-related complications later, we suggest that GEMS be repeated in these patient populations every 1-3 years. This strategy is in line with European Association for the Study of the Liver guidelines⁽³⁹⁾ that recommend repeating fibrosis markers during follow-up in patients with NAFLD at low risk of fibrosis. It is a feasible strategy that could help physicians during clinical practice to stratify at baseline and during follow-up those patients who are at high risk of complications and to customize their follow-up accordingly.

Patients evaluated in the Palermo cohort were recruited from a single tertiary center, and their clinical features, metabolic comorbidities, and genetic backgrounds may differ from the general population. Therefore, we validated GEMS in the general population as well as in high-risk subgroups (FIB-4 > 1.3, FLI ≥ 60, diabetes) from the UKBB, and we confirmed its association and its ability to stratify the risk of incident SLD. The somewhat suboptimal AUC derived from the UKBB general population is in line with evidence for lower accuracy of noninvasive tests⁽⁴⁰⁾ in predicting cirrhosis and its complications in the general population as a consequence of the relatively low incidence of LREs. However a relevant concern of our study is related to the fact that the outcome evaluated in the validation cohort (SLD) is different from LREs used to develop the score in the training cohort. This is a key point that could at least partially explain the drop in the accuracy of the score in the validation cohort and one that strongly demands external validation by using a more comparable outcome before suggesting the wide use of the score. Furthermore, future studies will also validate GEMS in cohorts of patients with NAFLD at risk for liver disease from other referral centers and of different ancestry.

There are a number of limitations to the present study. First, some metrics could be optimized as follows: the metabolic impact of diabetes on NAFLD could have been better reflected by hemoglobin A1c levels than by serum glucose and the use of liver stiffness for diagnosing severe fibrosis/cirrhosis may have overestimated the severity of baseline liver disease. Second, there may have been a slight overparameterization of the underlying regression model due to the moderate number (4.5) of events per variable. This may have resulted in an overfitted model. However,

we argue that the usual rule of thumb of 10 events per variable might be too conservative, especially for the Cox model. According to the simulation study reported by Vittinghoff and McCulloch,⁽⁴¹⁾ bias in a Cox model is relatively uncommon even with as few as five events per variable.

The lower AUC in the validation cohort seems to indicate a lower reliability of GEMS when applied to new data. Unfortunately, we were not able to assess how much of the drop in AUC was due to lower reliability and how much was due, as discussed above, to a different outcome considered in the validation cohort.

In conclusion, GEMS is the first composite score combining clinical features, metabolic comorbidities, and genetic common variants to 1) accurately predict LREs in patients with NAFLD who are at intermediate to high risk of advanced fibrosis as scored by FIB-4 and 2) stratify the risk of incident SLD in the general population.

Acknowledgment: We thank the staff and the participants of the UKBB study. This research has been conducted using the UKBB resource (application 37142).

REFERENCES

- 1) Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M, et al. Global epidemiology of non-alcoholic fatty liver disease—meta-analytic assessment of prevalence, incidence and outcomes. *Hepatology* 2016;64:73-84.
- 2) Adams LA, Lymp JF, St. Sauver J, Sanderson SO, Lindor KD, Feldstein A, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology* 2005;129:113-121.
- 3) Ekstedt M, Franzén LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* 2006;44:865-873.
- 4) Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 2015;149:389-397.e10.
- 5) Bhala N, Angulo P, van der Poorten D, Lee E, Hui JM, Saracco G, et al. The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study. *Hepatology* 2011;54:1208-1216.
- 6) Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol* 2013;66:719-725.
- 7) Crossan C, Tsochatzidis EA, Longworth L, Gurusamy K, Davidson B, Rodríguez-Perálvarez M, et al. Cost-effectiveness of non-invasive methods for assessment and monitoring of liver fibrosis and cirrhosis in patients with chronic liver disease: systematic review and economic evaluation. *Health Technol Assess* 2015;19:1-409, v-vi.

- 8) Crossan C, Majumdar A, Srivastava A, Thorburn D, Rosenberg W, Pinzani M, et al. Referral pathways for patients with NAFLD based on non-invasive fibrosis tests: diagnostic accuracy and cost analysis. *Liver Int* 2019;39:2052-2060.
- 9) Angulo P, Keach JC, Batts KP, Lindor KD. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. *Hepatology* 1999;30:1356-1362.
- 10) Neuschwander-Tetri BA, Clark JM, Bass NM, Van Natta ML, Unalp-Arida A, Tonascia J, et al.; NASH Clinical Research Network. Clinical, laboratory and histological associations in adults with nonalcoholic fatty liver disease. *Hepatology* 2010;52:913-924.
- 11) Angulo P, Bugianesi E, Bjornsson ES, Charatcharoenwithaya P, Mills PR, Barrera F, et al. Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 2013;145:782-789.e4. <http://dx.doi.org/10.1053/j.gastro.2013.06.057>
- 12) Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2008;40:1461-1465.
- 13) Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* 2014;46:352-356.
- 14) Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, et al. A protein-truncating HSD17B13 variant and protection from chronic liver disease. *N Engl J Med* 2018;378:1096-1106.
- 15) Gellert-Kristensen H, Richardson TG, Davey Smith G, Nordestgaard BG, Tybjaerg-Hansen A, Stender S. Combined effect of PNPLA3, TM6SF2, and HSD17B13 variants on risk of cirrhosis and hepatocellular carcinoma in the general population. *Hepatology* 2020;72:845-856.
- 16) Introduction: Standards of Medical Care in Diabetes-2020. *Diabetes Care* 2020;43(Suppl. 1):S1-S2.
- 17) Huang PL. A comprehensive definition for metabolic syndrome. *Dis Model Mech* 2009;2:231-237.
- 18) Williams B, Mancia G, Spiering W, Rosei EA, Azizi M, Burnier M, et al; ESC Scientific Document Group. 2018 ESC/ESH guidelines for the management of arterial hypertension. *Eur Heart J* 2018;39:3021-3104. Erratum in: *Eur Heart J*;2019;40:475.
- 19) Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al.; APRICOT Clinical Investigators. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 2006;43:1317-1325.
- 20) Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al.; Nonalcoholic Steatohepatitis Clinical Research Network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313-321.
- 21) European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;69:182-236. Erratum in: *J Hepatol* 2019;70:817.
- 22) Cillo U, Burra P, Mazzaferro V, Belli L, Pinna AD, Spada M, et al. A multistep, consensus-based approach to organ allocation in liver transplantation: toward a "blended principle model". *Am J Transplant* 2015;15:2552-2561. <http://dx.doi.org/10.1111/ajt.13408>
- 23) Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12:e1001779.
- 24) Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The fatty liver index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC Gastroenterol* 2006;6:33.
- 25) Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203-209.
- 26) Cox DR. Regression models and life-tables. *J Roy Stat Soc: Ser B (Methodol)* 1972;34:187-202.
- 27) Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81:515-526. <http://dx.doi.org/10.1093/biomet/81.3.515>
- 28) Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. *Proceedings of the 2nd International Symposium on Information Theory*. Budapest, Hungary: Akademiai Kiado; 1973:267-281.
- 29) Wong VW, Irls M, Wong GL, Shili S, Chan AW, Merrouche W, et al. Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease. *Gut* 2019;68:2057-2064.
- 30) Kanwal F, Kramer JR, Li L, Dai J, Natarajan Y, Yu X, et al. Effect of metabolic traits on the risk of cirrhosis and hepatocellular cancer in nonalcoholic fatty liver disease. *Hepatology* 2020;71:808-819.
- 31) Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanon I, et al. Non-invasive stratification of hepatocellular carcinoma risk in non-alcoholic fatty liver using polygenic risk scores. *J Hepatol* 2021;74:775-782.
- 32) Israelsen M, Juel HB, Detlefsen S, Madsen BS, Rasmussen DN, Larsen TR, et al.; GALAXY and MicroLiver consortia. Metabolic and genetic risk factors are the strongest predictors of severity of alcohol-related liver fibrosis. *Clin Gastroenterol Hepatol* 2020; <https://doi.org/10.1016/j.cgh.2020.11.038>.
- 33) Calzadilla-Bertot L, Vilar-Gomez E, Wong VW, Romero-Gomez M, Aller-de la Fuente R, Wong GL, et al. ABIDE: an accurate predictive model of liver decompensation in patients with non-alcoholic fatty liver-related cirrhosis. *Hepatology* 2021;73:2238-2250.
- 34) Stender S, Kozlitina J, Nordestgaard BG, Tybjaerg-Hansen A, Hobbs HH, Cohen JC. Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci. *Nat Genet* 2017;49:842-847.
- 35) Zaharia OP, Strassburger K, Knebel B, Kupriyanova Y, Karusheva Y, Wolkersdorfer M, et al.; GDS Group. Role of patatin-like phospholipase domain-containing 3 gene for hepatic lipid content and insulin resistance in diabetes. *Diabetes Care* 2020;43:2161-2168.
- 36) Sookoian S, Pirola CJ. Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of non-alcoholic fatty liver disease. *Hepatology* 2011 Jun;53:1883-1894.
- 37) Huang Y, He S, Li JZ, Seo Y-K, Osborne TF, Cohen JC, et al. A feed-forward loop amplifies nutritional regulation of PNPLA3. *Proc Natl Acad Sci* 2010;107:7892-7897. <http://dx.doi.org/10.1073/pnas.1003585107>
- 38) D'Em TM, Souza SC, Aronovitz M, Obin MS, Fried SK, Greenberg AS. Estrogen regulation of adiposity and fuel partitioning. Evidence of genomic and non-genomic regulation of lipogenic and oxidative pathways. *J Biol Chem* 2005;280:35983-35991.
- 39) European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO clinical practice guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol* 2016;64:1388-1402.
- 40) Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Ability of noninvasive scoring systems to identify individuals in the population at risk for severe liver disease. *Gastroenterology* 2020;158:200-214.
- 41) Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. *Am J Epidemiol* 2007;165:710-718.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1877/suppinfo.