LiverPRO for the prediction of significant liver fibrosis in primary care: Development, validation, and prognostic evaluation of a novel score

Katrine P Lindvig PhD[1,2], Katrine H Thorhauge MD[1,2], Johanne K Hansen MD[1,2], Maria Kjærgaard PhD[1,2], Camilla D Hansen PhD[1,2], Stine Johansen MD[1,2], Ellen Lyngbeck[1,2], Mads Israelsen PhD[11,2], Peter Andersen[1], Katrine Bech MD[1,2], Nikolaj Torp MD[1,2], Helle L Schnefeld[1,2], Sönke Detlefsen PhD[2,3], Sören Möller[4], Isabel Graupera PhD[5,6,7,8], Morten B Trelle PhD[9], Steen Antonsen MD. PhD[9], Rebecca Harris MD PhD[10], Line L Kårhus[11] Kirsten Bjørnsbo[11], Charlotte Brøns[12], Torben Hansen MD PhD [13], Andreas Geier[14], Heiner Wedemeyer PhD[15,16], Stefan Zeuzem[17,18], Jörn M. Schattenberg PhD[19], Pere Ginès PhD[5,6,7,8], Indra Neil Guha PhD[10], Aleksander Krag PhD [1,2], Maja Thiele PhD [1,2]

Author affiliations

[1]Department of Gastroenterology and Hepatology, Odense University Hospital, Kløvervænget 10, 5000 Odense C, Denmark
[2]Institute of Clinical Research, University of Southern Denmark, J.B. Winsløws Vej 19, 5000 Odense C, Denmark
[3]Department of Pathology, Odense University Hospital, J.B. Winsløws Vej 15, 5000 Odense C, Denmark
[4]Open Patient data Explorative Network, Odense University
[5]Liver Unit, Hospital Clinic of Barcelona, Barcelona, Spain
[6]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain
[7]Centro de Investigación en Red de Enfermedades Hepáticas y Digestivas (Ciberehd), Barcelona, Spain.
[8]Faculty of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain
[9]Department of Clinical Biochemistry, Odense University Hospital Svendborg, Baagøes Álle 15, 5700 Svendborg, Denmark
[10]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, United Kingdom
[11] Center for Clinical Research and Prevention, Copenhagen University Hospital - Bispebjerg and Frederiksberg, Copenhagen, Denmark
[12] Steno Diabetes Center Copenhagen, Herlev, Denmark
[13] Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen Faculty of Health and Medical Sciences, Copenhagen, Denmark
[14]Department of Hepatology, University of Wuerzburg, Oberduerrbacher Str. 6, 97080 Wuerzburg, Germany
[15]Department of Gastroenterology, Hepatology, Infectious Diseases and Endocrinology, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany
[16]Leberstiftungs-GmbH Deutschland, Carl-Neuberg-Str. 1, 30625, Hannover, Germany
[17]Department of Medicine I, Gastroenterology, Hepatology and Endocrinology, University Hospital Frankfurt, Frankfurt am Main, Germany
[18]Goethe University, Frankfurt, Germany
[19]Department of Internal Medicine II, Saarland University Medical Center, Kirrberger Str. 100, 66421 Homburg, Germany

Corresponding author

Aleksander Krag, Department of Gastroenterology and Hepatology, Odense University Hospital, Kløvervænget 10, entrance 112, 11th floor, DK-5000 Odense C, Denmark. Email: Aleksander.krag@rsyd.dk. Telephone: +45 2140 9915

Keywords: Liver fibrosis, algorithm, biomarker, diagnosis, prognosis, prediction

Research in context

Evidence before this study

Early detection of fibrosis is important to facilitate lifestyle interventions, disease monitoring, and pharmacological treatment and thus reduce patient morbidity and mortality. Non-invasive tests have been developed to detect advanced fibrosis; however, recent evidence suggests that patients with significant fibrosis may also be at high risk of liver-related events. We searched PubMed up to November 10, 2023, with the terms "liver fibrosis" AND ("biomarker" OR "test" OR "serological marker" OR "marker" OR "tool" OR "score") AND "primary care", with no date or language restrictions, to find information about availability of non-invasive tests to detect liver fibrosis in primary care. The search retrieved 208 articles. Existing simple diagnostic scores, such as the Fibrosis 4 index (FIB-4), have better diagnostic accuracy than individual routine liver blood test components; however, they result in a high number of false positives in low prevalence settings. Other scores use a fixed set of variables, which are not always available in heterogeneous healthcare systems, limiting their flexibility. The commercially available biomarkers are expensive and not widely implemented. The lack of accurate diagnostic tools in primary care results in both futile investigations in individuals who do not have significant liver fibrosis, and in under-referral of individuals with advanced fibrosis. Primary care physicians urgently need new tools to detect significant and advanced fibrosis with high diagnostic accuracy.

Added value of this study

We developed LiverPRO, an inexpensive, CE marked and practical prediction model that includes patient's age and nine routine blood tests for use across the spectrum of steatotic liver disease. As the algorithm uses different combinations of standard variables in multivariable fractional polynomials, LiverPRO is a flexible and dynamic model and can adapt to user requirements and availability of individual routine blood tests. LiverPRO exhibited good diagnostic and prognostic accuracy for significant fibrosis, advanced fibrosis, and liver-related events in independent, high- and low-prevalence cohorts. LiverPRO performed with comparable accuracy as the Enhanced Liver Fibrosis test and Liver Risk Score, and was superior to FIB-4 and the NAFLD Fibrosis Score.

Implications of all the available evidence

LiverPRO may aid patients and clinicians to make informed referral decisions by providing accurate estimations of the risk for significant liver fibrosis and future liver-related events. A decision support tool that is based solely on a dynamic set of common and inexpensive biochemical variables that can be integrated with existing automated laboratory systems and adapted to local health systems will ensure optimized implementation in primary care. Use of LiverPRO could potentially reduce the costs, resources, and patient anxiety associated with the many false positive results, and futile referrals to secondary care, produced by an approach such as FIB-4.

Abstract

Background: Significant liver fibrosis is associated with future adverse events in patients with steatotic liver disease (SLD). We designed a software tool for detection of significant liver fibrosis in primary care.

Methods: We developed and validated LiverPRO using six independent cohorts representing SLD related to alcohol and/or metabolic dysfunction. We used significant fibrosis (histology stage ≥F2) and advanced fibrosis (≥F3) as outcomes for variable selection in the development cohort and built the model with fractional polynomial regression. We independently validated the tool for prediction of elevated liver stiffness by transient elastography (TE ≥8 kPa and ≥12 kPa) and liver-related events (LRE).

Findings: In the development cohort (n=462), we derived 466 multivariable models consisting of age in combination with 3–9 variables from a list of nine blood tests. LiverPRO diagnosed significant fibrosis with good accuracy (TE≥8 kPa AUC 0·86, 95% CI 0·83–0·90). In the DECIDE validation cohort (n=6,468), LiverPRO detected TE ≥8 kPa with good accuracy (AUC 0·80, 0·78–0·82), comparable to ELF (AUC 0·78, 0·75–0·80) and the LiverRisk score (AUC 0·81, 95% CI 0·79–0·84), but superior to FIB-4 (AUC 0·69, 0·66–0·72) and NAFLD Fibrosis Score (AUC 0·74, 0·72–0·77). Findings were consistent in three other validation cohorts (n=2,554) from Denmark, Germany, and England, albeit accuracy was slightly lower. With a rule-out cut-off of <25% (no further examinations required) LiverPRO correctly classified 82% of participants with significant fibrosis, and with a rule-in cut-off of >65% (referral to hepatologist required) LiverPRO correctly classified 95% of participants. LiverPRO strongly predicted LREs (C-statistic >0·8) in 470,795 participants from the UK Biobank. On the basis of these results, LiverPRO was certified according to IVDR class b, obtaining European CE approval in 2024.

Interpretation: LiverPRO reliably identifies significant liver fibrosis and elevated liver stiffness, and predicts the 10-year risk of LREs in primary care. It serves as a versatile decision support tool, with the added advantage of adaptability to liver blood test availability.

Funding:

EU Horizon 2020 research and innovation program and Novo Nordisk Foundation

Introduction

Steatotic liver disease (SLD) includes metabolic dysfunction-associated steatotic liver disease (MASLD; includes steatosis associated with high cholesterol, diabetes, and obesity), MetALD (i.e., those with MASLD who consume greater amounts of alcohol per week), and alcohol-related liver disease (ALD), and is present in over one third of the world's adult population.[1–3] Progression to fibrosis and ultimately cirrhosis and decompensated disease occurs over years. Cirrhosis represents the end stage of progressive liver fibrosis and is the 11th most common cause of death globally.[4] Risk factors for steatotic liver include metabolic syndrome and excessive alcohol consumption; however, only 9% and 5% of individuals in a population with excessive alcohol consumption and diabetes or obesity, respectively, will progress to advanced fibrosis (histological fibrosis stage ≥F3).[5,6] Histological fibrosis stage is the best predictor of liver-related outcomes; diagnostic studies have focused on identifying advanced fibrosis due to progression to decompensation within 3–5 years.[7] However, recent data show that patients with significant fibrosis (F2) are also at high risk of liver-related events (LREs), especially when alcohol is the main etiology.[8,9] Thus, early detection of significant fibrosis (F2) is of utmost importance to facilitate pharmaceutical treatment, stop fibrosis progression, implement lifestyle interventions, and initiate disease monitoring, thereby reducing morbidity and mortality.[10,11]

Liver fibrosis may be detected using inexpensive non-invasive tests; however, each comes with its own limitations. The Fibrosis 4 index (FIB-4), which is recommended by the European Association for the Study of the Liver (EASL) and American Association for the Study of Liver Diseases as the first-tier screening test to rule out liver fibrosis in primary care, has shown moderate to good diagnostic accuracy for advanced fibrosis in high-prevalence populations; however, it has a false positive rate of 28% in low-prevalence populations and a substantial false negative rate, especially in at-risk populations.[5,7,12,13] Several simple, blood based scores have been developed to improve diagnostic accuracy beyond FIB-4 and to predict risk of advanced liver disease and risk of decompensation and LREs.[14–18] However, many are expensive; none has received regulatory approval as a diagnostic or monitoring biomarker, been made available on a commercial platform, or even made the transition from academia to clinical practice. Many scores include clinical variables, like body mass index (BMI) and presence of diabetes for the NAFLD fibrosis score (NFS), and are therefore not suitable for automated testing. Few have been validated in low-prevalence cohorts, and all require a fixed set of variables to enable calculation of the scores. Consequently, a flexible score that can be automated based on available variables is desirable.

The current lack of such automated diagnostic tools in primary care results in both futile referrals and investigations in individuals who do not have significant liver fibrosis and in under-referral of individuals with advanced fibrosis. We used standard liver blood test results to develop and validate a simple diagnostic and prognostic score for significant liver fibrosis in high- and low-prevalence populations. Our aim was to develop a flexible score both in terms of liver blood test availability and applicability to different patient populations across the spectrum of SLD.

Methods

This work adheres to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines (**Table S1** in supplementary material) and the World Medical Association's Declaration of Helsinki.[19] All participants provided informed written and oral consent prior to inclusion, and research ethical approvals were obtained for all study cohorts.[20]

Participants

We developed our model in a prospective, biopsy-controlled cohort acquired from Odense University Hospital, Denmark, which included both MetALD and ALD patients. This cohort has been described previously.[7,8] The model was validated in four independent cohorts from Denmark (DECIDE[21] and Inter99 study[22] cohorts), Germany (German SLD Registry[23]), and the United Kingdom (UK; Scarred Liver Project [SLP] cohort[24]) (**Figure 1**). The DECIDE cohort included patients with MASLD, MetALD, and ALD. The German SLD Registry only included patients with MASLD, whereas the Inter99 and SLP cohorts included participants from the general population. The UK Biobank, a large-scale biomedical database and research resource containing genetic, lifestyle, and health information from UK participants, served as the prognostic evaluation cohort.[25] See supplementary material for further detail on patient cohorts.

Data on physical examination, comorbidities, laboratory variables, liver diagnostics, and lifestyle factors were obtained from participants in all cohorts, where possible. Prior to 2016, all participants in the development cohort were biopsied; however, after this date only participants with transient elastography (TE) ≥6 kilopascal (kPa) were biopsied, as TE <6 kPa was associated with a negative predictive value of 100% for advanced fibrosis.[26] Quality criteria were applied to biopsy specimens, and a central pathologist was used. Reliability criteria were applied for vibration-controlled TE in all cohorts. All participants (where possible) in validation cohorts underwent TE as per standard procedure[27], except for the UK Biobank.

Development of the LiverPRO score

We developed the LiverPRO score as a set of multivariable models. LiverPRO selects the algorithm with the best performance (highest area under the receiver operating characteristic curve [AUC]) based on the availability of biochemical blood tests, allowing the software to calculate the best model that is available to users (**Figure S1** in supplementary material). Originating from the University of Southern Denmark, LiverPRO is a medical device software being implemented by Evido, who is facilitating regulatory approval, including CE marking under *in vitro* diagnostic medical devices regulation in Europe and Food and Drug Administration 510(k) clearance in the United States (US). To transition from development and validation into clinical practice, it is crucial to establish a digital infrastructure and a quality management system compliant with ISO 13485 standards to support regulatory processes and enable product market entry.

<u>Selection of candidate variables</u>

We selected candidate variables based on existing evidence, from a search in MEDLINE using the MeSH terms “liver cirrhosis,” “liver diseases,” “liver diseases, alcoholic,” “non-alcoholic fatty liver disease,” “diagnosis,” and “biomarkers”. Between 2018 and 2022, 250 human studies were identified and from these, we selected 25 widely available predictors of fibrosis (including patient age, gender, alcohol consumption, smoking, basic clinical investigations, and laboratory variables) for univariable regression analyses (**Table S2** in supplementary material).

<u>Model development</u>

We first conducted univariable logistic regression analyses for each of the 25 candidate predictors using biopsy-verified significant and advanced fibrosis as the outcome. We selected predictors for further assessment if they were associated ($p<0{\cdot}1$) with both significant and advanced fibrosis in the

development cohort. An alpha level of 0.1 was chosen to strike a balance between mitigating complexity and capturing pertinent information; a higher alpha value of 0·2 would likely lead to an influx of input variables due to inherent correlation among biochemical variables whereas a lower alpha value of 0·05 could have resulted in overlooking significant variables. The selected input variables included age (years), aspartate aminotransferase (AST; U/L), alkaline phosphatase (ALP; U/L), gamma-glutamyl transferase (U/L), INR, albumin (g/L), sodium (mmol/L), bilirubin(mg/dL), platelet count (10^9/L) and cholesterol (mmol/L). Multivariable models were developed by combining three to nine of the biochemical variables in all possible permutations. Age was consistently included in the models due to its known association with significant and advanced fibrosis.[28] For each combination of all available variables, a logistic regression model for significant fibrosis was constructed using multivariable fractional polynomials to allow for non-linear effects on a log-odds scale. LiverPRO is calculated using logistic regression where the coefficients are translated into predicted probabilities which provide the percentage risk of having significant or advanced fibrosis and thus the end result of the LiverPRO calculation. This serves as the foundation for all subsequent regression models detailed in this manuscript. Multivariable fractional polynomials model the effect of a predictor using a specific class of polynomials, with the possibility of reducing the number of predictors or simplifying the polynomial function using an approximate closed-test principle, thereby reducing the possibility of type I error. To reduce the risk of over- or underestimation in case of extreme values, most variables were capped. We applied fractional polynomials of the second order with powers chosen from the following set: –0·5, –1, –2, –3, 0, 0·5, 1, 2, and 3.

LiverPRO is a medical device software that can be configured during the installation process to offer flexibility in accommodating different numbers of input variables from the available options. The simplest models included three biochemical variables and age, while the largest models included nine biochemical variables and age. This means that LiverPRO comprises 466 sub-models, each designed to accept a unique combination of input parameters. When provided with a specific set of input values, the LiverPRO software automatically utilizes individual sub-model AUCs to determine the most suitable sub-model available that accepts the provided inputs or a subset thereof. These individual sub-model AUCs, utilized for model selection, were calculated using data from the DECIDE cohort, where all input variables are present. See **Figure S2** in the supplementary material for distribution of AUCs across different input parameter combinations and sub-models.

Other indices

We compared LiverPRO to FIB-4 index and NAFLD Fibrosis Score (NFS) in all validation cohorts.[29–31] In the DECIDE cohort, we also compared LiverPRO to the LiverRisk score and Enhanced Liver Fibrosis (ELF) test.[17,32] Due to differences in data collection procedures across cohorts, the LiverPRO model included different numbers and types of variables for different cohorts. For example, for the SLP cohort LiverPRO was calculated based on a 7-variable model comprising AST, albumin, ALP, bilirubin, cholesterol, sodium, and platelets, and for the German SLD Registry LiverPRO was calculated based on eight variables as only sodium was not available. For the Inter99 cohort LiverPRO was calculated based on a 5-variable model comprising AST, albumin, cholesterol, sodium, and platelets.

Outcomes

The main outcomes of interest in the development cohort were biopsy-verified significant liver fibrosis (≥F2) or advanced liver fibrosis (≥F3), while the main outcome measures in the validation

cohorts were elevated liver stiffness measure (LSM), defined as TE ≥8 kPa as a surrogate marker of significant fibrosis and TE ≥12 kPa for advanced fibrosis. For evaluation of the prognostic performance, LREs and liver-related mortality were the main outcomes of interest. LREs were defined as the occurrence of ascites, spontaneous bacterial peritonitis, varices requiring treatment (secondary prophylaxis), variceal bleeding, liver failure-induced jaundice, hepatorenal syndrome, or hepatocellular carcinoma. Non-spontaneous bacterial peritonitis infections were not included in the analysis. These clinical outcomes were derived by two physicians/researchers through systematic review of participants' electronic medical records from the development cohort and based on International Classification of Disease version 10 (ICD-10) codes (**Table S3** in supplementary material) for the UK Biobank.[8]

Statistical analyses

Statistical analyses were conducted using Stata version 17 (Stata Corporation, USA), and Python 3·11 software. Descriptive statistics were used to summarize patient demographic and clinical characteristics. Continuous variables were described using means, standard deviation or 95% confidence intervals (95% CIs) where applicable, or medians with interquartile range (IQR), and categorical variables were described using frequencies and proportions.
We assessed calibration, representing the model's fit to the data, through Akaike information criterion (AIC), where lower scores indicate superior fit. Supplementary discrimination, indicating the test's ability to predict fibrosis, was evaluated using AUC. AUC values range from 0 to 1, with 1 indicating perfect predictive accuracy.

<u>Diagnostic performance</u>

Diagnostic performance of LiverPRO and other indices are the primary analysis, and assessed by calibration and discrimination in both development and validation cohorts. AUC results were considered excellent for values 0·9–1, good for values 0·8–0·9, fair for values 0·7–0·8, poor for values 0·6–0·7, and failed for values 0·5–0·6. AUC comparison was conducted using the DeLong test. Subgroup analyses were performed to assess the diagnostic performance of LiverPRO and other indices in individuals with ALD (in development and DECIDE cohorts) and MASLD (in German SLD and DECIDE cohorts). Differentiation by etiology was not possible in the other validation cohorts (i.e., SLP, Inter99).

Cut-off values for LiverPRO were established in the DECIDE cohort and used to assess the clinical performance of the model to predict significant (TE ≥8 kPa) and advanced (TE ≥12 kPa) fibrosis. First, subjects were stratified into LiverPRO rule-out (i.e., low risk), grey zone, or rule-in (i.e., high risk) groups. Cut-off values were calculated for each potential value of LiverPRO between 0% and 100% (0% and 100% score indicating 0% and 100% chance of significant fibrosis (TE ≥8 kPa), respectively). The rule-out cut-off was selected to achieve an 80% sensitivity, while the rule-in cut-off was chosen for a specificity of 90%, enabling detection of true positives while minimizing false positives. Participants with missing blood samples where a LiverPRO score could not be calculated were excluded. Participants in the rule-in group have significant/advanced fibrosis and should be referred to a hepatologist, whereas no further examinations for SLD are required in primary care for those in the rule-out group. Repeat testing would be recommended in participants categorized in the grey zone. Cut-off values for other indices were based on previous research as follows: FIB-4 (<1·3 rule-out, >2·67 rule-in), ELF test (<9·8 rule-out, ≥10·5 rule-in); NFS (<–1·46 rule-out, >0·68 rule-in); and

as the LiverRisk score is a score that aims to mirror the FibroScan, we used <8 rule-out and >12 rule-in.[77]

Prognostic performance

We assessed the prognostic performance of LiverPRO and other indices to predict LREs and mortality in the development cohort and the UK Biobank. NFS and ELF test could not be calculated in the UK Biobank due to missing input parameters. We used Cox regression with time-to-event analyses and investigated the relationship between survival time and LiverPRO and other indices. Furthermore we used Fine-Gray regression to investigate LREs with death as a competing event. Prevalent cases as defined from ICD-10 codes used to define liver-related events, were excluded prior to the analysis and calculation of the different indices. LiverPRO was not re-developed for prognostic purposes; rather the LiverPRO score was used as a single variable in the regression analyses. Proportional hazards and linearity assumptions of Cox regression were evaluated through diagnostic plots (**Figure S3** in supplementary material). Kaplan–Meier plots were created to show the survival curves of the different risk groups. Overall prognostic accuracy was presented using Harrell's C-statistic and hazard ratios (HR), which measure the capacity of the prognostic model to distinguish between participants survival. The starting point was defined as the time of test calculation, while the endpoint was determined by either a LRE or the conclusion of the follow-up period. Participants were tracked from enrollment until death, loss to follow-up, or October 1, 2020. For participants in the UK Biobank, the censoring date was November 30, 2022.

Results

Participants

A total of 479,843 participants were included: 462 in the development cohort; 6,486 in DECIDE; 1,367 in SLP; 390 in Inter99; 711 in German SLD Registry; and 470,795 in UK Biobank. Participant demographics and clinical characteristics are displayed in **Table 1**. The mean age ranged from 51 to 67 years and 48–75% of participants were male. Excessive alcohol consumption was reported by all participants in the development cohort and 7–40% of participants in the other cohorts. Diabetes was present in >40% of participants in the SLP and German SLD Registry compared to <15% in all other cohorts. Between 69–100% of participants underwent TE, except for participants in the UK Biobank. The prevalence of TE 8–12 kPa ranged from 4–19% and the prevalence of TE ≥12 kPa ranged from 1–35% (**Table 1**). In the development cohort 60% of participants had ALD, whereas in the DECIDE cohort 31% and 59% had MASLD and no SLD, respectively (**Table 1**)

LiverPRO development

In univariable regression analyses, 13 of the 25 identified candidate predictors of fibrosis correlated ($p<0{\cdot}1$) with both significant and advanced fibrosis in the development cohort. Three variables (creatinine, mean corpuscular volume, triglycerides) were subsequently removed as they exhibited high covariance with other, stronger predictors. The final variable list included nine biochemical variables (AST, gamma-glutamyl transferase [GGT], ALP, total cholesterol, sodium, international normalized ratio [INR], bilirubin, albumin, platelets) and age (**Table S4** in supplementary material). A total of 466 unique, multivariable models were constructed; the models with the highest AUC and lowest AIC scores were identified as the models with the best performance and goodness of fit, respectively (**Table S4** and **Figure S4** in supplementary material).

LiverPRO performed well in terms of discrimination but showed moderate calibration, with an overestimation of liver stiffness prevalence above 8 kPa (Figure S4) However the calibration towards the outcome in the validation cohorts were not perfectly calibrated, most likely due to missing input parameters, not allowing the best performing sub-models to be used, and because of the difference between outcome used for development (biopsies) and that used when testing calibration (TE>8 kPa) (**Figure S4**). The diagnostic accuracy (AUC) of LiverPRO in the development cohort was 0·86 (95% CI 0·83–0·90) for significant fibrosis (≥F2) and 0·89 (95% CI 0·85–0·92) for advanced fibrosis (≥F3; **Table 2**). These findings are comparable with LSM to predict significant and advanced fibrosis (**Table S5** in supplementary material). The ELF test predicted a comparable level of significant and advanced fibrosis in the development cohort, whereas the FIB-4 index was less accurate than LiverPRO to predict significant fibrosis, and NFS was less accurate for advanced fibrosis (**Figure 2A–D**; **Table 2**).

LiverPRO cut-offs

LiverPRO was calculated for 97% (n=5,869) of participants in the DECIDE cohort, the cohort used to evaluate cut-off values. Participants were categorized as rule-out if their score was <25% (n=3,581, 61%), grey zone if their score was 25–65% (n=1,905, 32%), and rule-in if their score was >65% (n=383, 7%). With a rule-out cut-off of <25%, LiverPRO correctly classified 82% (sensitivity) of participants with TE ≥8 kPa, higher than FIB-4 (>1.3) and ELF (>9.8) test which correctly classified 56% and 46%, respectively. Findings were comparable for advanced fibrosis although test sensitivities were higher (**Table 3**). With a rule-in cut-off of >65%, LiverPRO correctly classified 95% of participants with TE ≥12 kPa, comparable to the other indices (**Table S6** in supplementary material). To ensure equitable evaluation of the various non-invasive tests, we have added a supplementary analysis with adjusted the cutoffs to achieve a 80% sensitivity and 90% specificity, for all the non-invasive tests in the DECIDE cohort and included the performance results in **Table S7** in the supplementary material.

Diagnostic performance

In the DECIDE cohort, LiverPRO detected TE ≥8 kPa with good accuracy (AUC 0·80, 95% CI 0·78–0·82), comparable to the LiverRisk score (0·81 95% CI 0·79–0·84) and the ELF test (AUC 0·78, 95% CI 0·75–0·80), but superior to FIB-4 (AUC 0·69, 95% CI 0·66–0·72), and NFS (AUC 0·74, 95% CI 0·72–0·77), (**Table 2**). LiverPRO detected TE ≥8 kPa with moderate accuracy in the SLP (AUC 0·69, 0·65–0·72), Inter99 (AUC 0·70, 0·57–0·84), and German SLD (AUC 0·72, 0·68–0·76) cohorts. FIB-4 and NFS were calculated in all validation cohorts, whereas the ELF test and LiverRisk score was calculated in the development cohort, and in the DECIDE cohort. The diagnostic accuracy of LiverPRO for predicting TE ≥12 kPa (AUC 0·86, 95% CI 0·82–0·89) in the DECIDE cohort was comparable to the LiverRisk score, ELF test, FIB-4, and NFS. Findings were consistent in the other three validation cohorts, albeit accuracy was lower (**Table 2**). A head-to-head analysis were performance benchmarks are conducted only on participants with all indices present can be found in the supplementary material (**Table S8**).

In participants with ALD, LiverPRO and LiverRisk score were superior for diagnosing significant fibrosis in the DECIDE cohort (TE ≥8 kPa; LiverPRO AUC 0·89, 95% CI 0·83–0·94, LiverRisk score AUC 0·89, 95% CI 0·83-0·94) compared to the ELF test, FIB-4, and NFS (**Table S9** in supplementary material). In the subgroup of participants with MASLD, the diagnostic accuracy of LiverPRO for

significant fibrosis (TE ≥8 kPa; AUC 0·72, 95% CI 0·68–0·77) was comparable to the LiverRisk score, ELF test, and NFS but superior to FIB-4.

Prognostic performance

Liver-related events in the Developmentcohort

In the development cohort, LiverPRO had a C-statistic of 0·78 (95% CI 0·73–0·84) for predicting LREs, comparable to FIB-4, NFS, ELF test, and LiverRisk score (**Table 4**).

Liver-related events in the UK Biobank

In the UK Biobank, 2865 LREs were observed during the study period. We excluded 403 participants from the UK biobank due to known liver diseases before the calculation of the different indices. LiverPRO predicted LREs with a C-statistic of 0·74 (95%CI 0·73–0·75) compared to 0·67 (95% CI 0·66–0·68) for FIB-4 and 0.72 (95% CI 0·71–0·73) for LiverRisk score (**Table 4**). Furthermore, a HR of 17·1 (95%CI 15·6–18·9; p<0·001) was observed for LiverPRO in the high-risk (>65%) versus low-risk category (<25%). Similarly, a HR of 15·2 (95% CI 13·7–16·9; p<0·001) was estimated for FIB-4 >2·67 compared to FIB-4 <1·3, and LiverRisk score >12 had a HR of 43·3 (95% CI 37·7–49·6; p<0·001) compared to LiverRisk score <8. When including death as a competing event for LREs, sub-hazard ratios were consistent with the hazard ratios (**Table S10** in the supplementary material).

Liver-related mortality in the Developmentcohort

The follow-up period was median 4.4 years (IQR 2·8–6·2), hereof, 75 of 462 (16%) participants in the development cohort died; the median survival time was 3·6 years (95% CI 2·9–4·3 years); (**Figure 3A–D**). Given that 16% of participants died, 84% of participants were censored, 98% at the end of the study and 2% were lost to follow-up. The participants with high-risk LiverPRO scores had a cumulative all-cause mortality rate of 30·9%, whereas participants with grey zone and low-risk LiverPRO scores had a cumulative all-cause mortality rate of 10·8% and 6·2%, respectively. The LiverPRO score had a C-statistic of 0·87 (95% CI 0·78–0·97) for predicting liver-related mortality at two years in the development cohort, comparable to FIB-4 0·80 (95% CI 0·66–0·94), ELF test 0·90 (95% CI 0·86–0·94), and NFS 0·77 (95% CI 0·59–0·94) (**Table 4**).

Liver-related mortality in the UK Biobank

During the follow-up period, 40,133 of 470,795 (8·5%) participants in the UK Biobank died. The LiverPRO score had a C-statistic 0·75 (95% CI 0·73–0·76) for predicting liver-related mortality in the UK Biobank, slightly better than FIB-4 0·71 (95% CI 0·69–0·72), and comparable to LiverRisk score 0·73 (95% CI 0·71–0·74) (**Figure 4A–F**) (**Table 4**).

Discussion

We describe the development and validation of LiverPRO, a CE-marked risk prediction model for significant liver fibrosis, liver stiffness, and LREs in patients with SLD from metabolic dysfunction and/or excess alcohol intake. LiverPRO can calculate a score from 466 combinations of age and three to nine routine blood tests, making it cheap, flexible, and easily adaptable to various healthcare systems.

LiverPRO showed good accuracy for significant fibrosis in a high-prevalence cohort and for liver stiffness in the low-prevalence DECIDE cohort. Slightly lower accuracy in the SLP, German SLD, and Inter99 cohorts was likely due to missing input variables, leading to the use of lower-priority models.

LiverPRO's accuracy was comparable to the LiverRisk score and the ELF test. However, the ELF test requires specific, costly equipment (Siemens Atellica or Centaur platforms), while LiverRisk, though also using routine blood tests, lacks LiverPRO's flexibility as it relies on a fixed set of six tests and age.

Two studies reported poor correlation of FIB-4 and NFS scores with liver stiffness, resulting in many false positives and false negatives, making them suboptimal for case-finding or screening in low-prevalence populations[5,12]. These tests were developed for high-prevalence populations in secondary or tertiary care[33,34]. EASL proposed a three-tier approach (FIB-4, TE, and serum-based markers) to diagnose liver fibrosis and reduce liver biopsies[29]. Validation in a low-prevalence cohort found that 40% would have a positive FIB-4 result and be referred for TE; further testing reduced FIB-4 false positives by 42% with FibroTest and 29% with FibroMeter[35]. However, the EASL algorithm did not reduce referrals for secondary care[5]. Using a more specific test like LiverPRO in primary care could reduce false positives and unnecessary patient anxiety[5,36].

Recent publication of the LiverRisk score highlights the need for advanced models using multivariable input to improve risk stratification and referral pathways in low prevalence cohorts[17]. Advanced fibrosis models, like LiverPRO, offer better diagnostic and prognostic potential than older tests such as FIB-4 by utilizing more complex data[18]. Key features for future models include ease of automation, integration into existing electronic laboratory systems, and market access authorization, such as CE marking in Europe and FDA approval in the US. LiverPRO is the first CE-certified medical device software product (IVDR class b) using routine laboratory tests, integrating seamlessly into laboratory information systems for immediate clinical use. Further research is needed to determine LiverPRO's optimal role in liver fibrosis referral pathways, evaluate its cost-benefit ratio, and investigate diagnostic and prognostic concordance and discordance. LiverPRO includes a feature to avoid overestimations by refusing to calculate algorithms containing INR if the INR is above 2.0 for patients on warfarin or direct oral anticoagulants.

This study showed that LiverPRO strongly predicted LREs (C-statistic 2-years 0.8), comparable to ELF and LiverRisk score, and superior to FIB-4. LiverPRO identified more high-risk events (891, 31%) compared to FIB-4 (550, 19%). While LiverPRO's diagnostic superiority over FIB-4 and NFS is evident, its prognostic capabilities need further research to fully leverage advanced statistical modeling. Implementing LiverPRO in primary care could enable early detection of compensated liver disease.

Our study has several strengths. We designed a diagnostic algorithm that integrated accessible biochemical parameters from a cross-sectional biopsy-proven study of asymptomatic participants at risk of ALD due to ongoing or prior excessive alcohol consumption. The sample size of both the development and validation cohorts was substantial. As the data collection procedures differed by cohort, the LiverPRO model had an opportunity to simulate real-world practice in which blood sample collection procedures may differ across clinics, hospitals, or regions. Some limitations also need consideration. First, the choice of TE as a surrogate marker for liver fibrosis in the validation cohorts can be debated, as well as the chosen cut offs. We chose TE as a surrogate marker because of its global availability and high number of publications, although it is not the perfect marker and the accuracy to diagnose fibrosis is moderate. Second, the missing input variables for the SLP, German SLD, and Inter99 cohorts resulted in calculation of models of lower priority and some loss of accuracy. Third, our development cohort consisted only of participants with alcohol-related SLD

rather than that associated both with alcohol and metabolic dysfunction. The development cohort had a higher prevalence of liver fibrosis compared to the background population, resulting in only moderate calibration for estimating elevated liver stiffness in the low prevalence cohort. Nonetheless, LiverPRO was designed to diagnose fibrosis in at-risk individuals, thus we believe this cohort was adequate as a development cohort. In addition, most validation cohorts were skewed towards a high proportion of alcohol consumers; however, the German SLD registry included only MASLD participants. Fourth, LiverPRO was validated in northern European, predominantly Caucasian populations so applicability in other parts of the world and in other ethnic groups remains unknown. Non-invasive tests for liver fibrosis have been shown to preform differently in various ethnic groups.[37] It will be important to further validate LiverPRO in other populations, as this score might be a good solution in low- and middle-income countries where availability of more advanced diagnostic tests, such as TE or direct fibrosis markers like ProC3 or ELF test is variable, but MASLD is increasing at an alarming rate.[36] However, it should be noted that the accuracy and discriminative power of the model applied depends on the combination of variables that are available in a particular clinical setting. Fifth, dietary habits, which may be a confounding variable for liver disease, were not analysed in this study. Sixth, retrospective study design is associated with issues around selection bias, confounding, data quality, and generalizability. Finally, Cox regressions models and HRs can present challenges with regard to interpretation and there may be inherent bias associated with their use.[38] The large HR estimates and CIs observed in our study may suggest sparse-data bias; however, as our study comprises a substantial number of individuals in the high-risk category with observed events, we anticipate minimal sparse-data bias and rather the large effect sizes observed are related to the differences between a LiverPRO of a 100% risk of liver fibrosis and a LiverPRO of 0% risk of liver fibrosis. Nevertheless, despite these complexities, HRs remain commonly employed and a preferred metric by many researchers.

In conclusion, LiverPRO is an accurate, user-friendly, CE-certified diagnostic and prognostic tool that can improve the assessment of patients with risk factors for significant liver fibrosis in primary care. LiverPRO is based solely on age and biochemical variables, and can therefore be integrated with existing automated laboratory data.

Acknowledgments

This research was conducted using data from UK Biobank, a major biomedical database (project ID number: 94859). We wish to acknowledge the valuable contribution made by the nurses and secretaries in the outpatient clinic at the Department of Gastroenterology and Hepatology, Odense University Hospital, as well as the contributions from Odense Patient Explorative Network (made by Anna Mejldal) and the GALAXY and LiverScreen consortia. Writing and editorial assistance was provided by Dr Julia Granerod of Dr JGW Consulting Ltd.

This study received funding from the Danish Advanced Technology Foundation and the European Union's Horizon 2020 Research and Innovation Program (grant agreement number 668031, GALAXY consortium). The Novo Nordisk Foundation's Challenge program (grant agreement number 16OC0016692 MicrobLiver consortium) funded data collection for the development cohort. The Novo Nordisk Foundation (grant agreement number 20OC0059393, DECIDE project) and the European Union's Horizon 2020 Research and Innovation Program (grant agreement number H2020 847989, LiverScreen consortium) funded data collection for the DECIDE cohort. The Innovation Fund Denmark (grant agreement number 9122-00119A, InnoExplorer grant), The Eureka and the

Innovation Fund Denmark funded through the EuroStars grant (grant agreement number E83 – LiverPRO 2043-00004B), and the University of Southern Denmark's PhD stipends, and Region of Southern Denmark's PhD stipends funded KL's (first author) PhD program. The Toyota Foundation and AP Moeller Foundation funded purchase of the FibroScan and Aixplorer systems for ultrasound elastography. Siemens Healthcare A/S (Ballerup, Denmark) provided the assays used for the Enhanced Liver Fibrosis test without restrictions. The German SLD-Registry is financially supported by Advanz Pharma Specialty Medicine Deutschland GmbH/Intercept Pharma Europe Ltd. and Gilead Sciences GmbH (grants to Deutsche Leberstiftung) as well as Novo Nordisk Pharma GmbH (directly via Leberstiftungs-GmbH). These agencies had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Disclosures

KL, MT, and **AK** are founders of the spin-out company Evido, which holds a commercial license with the University of Southern Denmark to use the invention LiverPRO commercially. **KL** has received speaker fees from Siemens Healthcare and Novo Nordisk. **MT** has received speaker fees from GSK, Echosens, Norgine, Siemens Healthcare, Takeda, Madrigal, Tillotts Pharma, and advisory fees from GE Healthcare, GSK, AstraZaneca, Boehringer Ingelheim, Novo Nordisk. **AK** has served as speaker for Novo Nordisk, Norgine, Siemens and Nordic Bioscience and participated in advisory boards for Norgine, Siemens, Resalis Therapeutics, Boehringer Ingelheim and Novo Nordisk, all outside the submitted work. Research support; Norgine, Siemens, Nordic Bioscience, Astra, Echosense. Board member and co-founder Evido. **AG**: served as a steering committee member or advisor for AbbVie, Alexion, Bayer, BMS, Eisai, Falk, Gilead, Heel, Intercept, Ipsen, Merz, MSD, Novartis, Pfizer, Roche, Sanofi-Aventis and as speaker for AbbVie, Advanz, Alexion, BMS, Burgerstein, CSL Behring, Falk, Gilead, Intercept, Merz, MSD, Novartis, NovoNordisc, Roche. He received research support from Intercept and Falk (both NAFLD CSG), Novartis**. HW**: Honoria for consulting and speaking from AstraZeneca; Honoria from Abbott, Falk, Gilead Sciences, GSK, Intercept/Advanz, Roche. **SZ**: XXX. **IG** has received ISCIII and Pfizer grants. **PG** has received Research funding from Gilead & Grifols. Consulted or attended advisory boards for Gilead, RallyBio, SeaBeLife, Merck, Sharp and Dohme (MSD), Ocelot Bio, Behring, Roche Diagnostics International and Boehringer Ingelheim, and speaker fees from Pfizer. **JKH** has received speaker fee from Norgine. **KB, SM** declares no conflicts of interest.

Author contributions

Conceptualization: AK, MT, and KL. Methodology, software, and writing original draft: KL, MK, PA, and MT. Data collection: MK, CDH, MI, JKH, HLS, SJ, KB, JM, NT, SD, SA, AG, HW, SZ and MBT. Data curation: PA, MK, and KL. Validation: RH, NG, LK, KSB, CB, KB, PG, IG, and TH. Writing, reviewing, and editing: MT, AK, MK, KHT, JM, SA, KB, MBT, SD, TH. Supervision: AK and MT.

References

1 Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. *Hepatol Baltim Md* 2023; **77**: 1335–47.

2 Luo N, Zhang X, Huang J, Chen H, Tang H. Prevalence of steatotic liver disease and associated fibrosis in the United States: Results from NHANES 2017-March 2020. *J Hepatol* 2024; **80**: e70–1.

3 Kalligeros M, Vassilopoulos A, Vassilopoulos S, Victor DW, Mylonakis E, Noureddin M. Prevalence of Steatotic Liver Disease (MASLD, MetALD, and ALD) in the United States: NHANES 2017-2020. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc* 2023; : S1542-3565(23)00914-X.

4 Asrani SK, Devarbhavi H, Eaton J, Kamath PS. Burden of liver diseases in the world. *J Hepatol* 2019; **70**: 151–71.

5 Kjaergaard M, Lindvig KP, Thorhauge KH, *et al.* Using the ELF test, FIB-4 and NAFLD fibrosis score to screen the population for liver disease. *J Hepatol* 2023; **79**: 277–86.

6 Abeysekera KWM, Valenti L, Younossi Z, *et al.* Implementation of a liver health check in people with type 2 diabetes. *Lancet Gastroenterol Hepatol* 2024; **9**: 83–91.

7 Thiele M, Madsen BS, Hansen JF, Detlefsen S, Antonsen S, Krag A. Accuracy of the Enhanced Liver Fibrosis Test vs FibroTest, Elastography, and Indirect Markers in Detection of Advanced Fibrosis in Patients With Alcoholic Liver Disease. *Gastroenterology* 2018; **154**: 1369–79.

8 Rasmussen DN, Thiele M, Johansen S, *et al.* Prognostic performance of 7 biomarkers compared to liver biopsy in early alcohol-related liver disease. *J Hepatol* 2021; **75**: 1017–25.

9 Loomba R, Ratziu V, Harrison SA. Expert Panel Review to Compare FDA and EMA Guidance on Drug Development and Endpoints in Nonalcoholic Steatohepatitis. *Gastroenterology* 2022; **162**: 680–8.

10 Mathurin P, Bataller R. Trends in the management and burden of alcoholic liver disease. *J Hepatol* 2015; **62**: S38-46.

11 Ginès P, Castera L, Lammert F, *et al.* Population screening for liver fibrosis: Toward early diagnosis and intervention for chronic liver diseases. *Hepatol Baltim Md* 2022; **75**: 219–28.

12 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 Off Clin Pract J Am Gastroenterol Assoc* 2022; **20**: 2567-2576.e6.

13 Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, *et al.* AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatol Baltim Md* 2023; **77**: 1797–835.

14 Åberg F, Luukkonen PK, But A, *et al.* Development and validation of a model to predict incident chronic liver disease in the general population: The CLivD score. *J Hepatol* 2022; **77**: 302–11.

15 Hydes T, Moore M, Stuart B, *et al.* Can routine blood tests be modelled to detect advanced liver disease in the community: model derivation and validation using UK primary and secondary care data. *BMJ Open* 2021; **11**: e044952.

16 Connoley D, Patel PJ, Hogan B, *et al.* The Enhanced Liver Fibrosis test maintains its diagnostic and prognostic performance in alcohol-related liver disease: a cohort study. *BMC Gastroenterol* 2021; **21**: 268.

17 Serra-Burriel M, Juanola A, Serra-Burriel F, *et al.* Development, validation, and prognostic evaluation of a risk score for long-term liver-related outcomes in the general population: a multicohort study. *Lancet Lond Engl* 2023; **402**: 988–96.

18 Blanes-Vidal V, Lindvig KP, Thiele M, Nadimi ES, Krag A. Artificial intelligence outperforms standard blood-based scores in identifying liver fibrosis patients in primary care. *Sci Rep* 2022; **12**: 2914.

19 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. *Br J Surg* 2015; **102**: 148–58.

20 World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013; **310**: 2191–4.

21 ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT03308916 (accessed Jan 9, 2024).

22 Bjørnsbo KS, Brøns C, Aadahl M, *et al.* Protocol for the combined cardiometabolic deep phenotyping and registry-based 20-year follow-up study of the Inter99 cohort. *BMJ Open* 2024; **14**: e078501.

23 Geier A, Rau M, Pathil-Warth A, *et al.* Clinical characteristics of patients with non-alcoholic fatty liver disease (NAFLD) in Germany - First data from the German NAFLD-Registry. *Z Gastroenterol* 2023; **61**: 60–70.

24 Harman DJ, Ryder SD, James MW, *et al.* Obesity and type 2 diabetes are important risk factors underlying previously undiagnosed cirrhosis in general practice: a cross-sectional study using transient elastography. *Aliment Pharmacol Ther* 2018; **47**: 504–15.

25 UK Biobank. https://www.ukbiobank.ac.uk/ (accessed Jan 9, 2024).

26 Thiele M, Detlefsen S, Sevelsted Møller L, *et al.* Transient and 2-Dimensional Shear-Wave Elastography Provide Comparable Assessment of Alcoholic Liver Fibrosis and Cirrhosis. *Gastroenterology* 2016; **150**: 123–33.

27 Dietrich CF, Bamber J, Berzigotti A, *et al.* EFSUMB Guidelines and Recommendations on the Clinical Use of Liver Ultrasound Elastography, Update 2017 (Long Version). *Ultraschall Med Stuttg Ger 1980* 2017; **38**: e16–47.

28 Wu X-N, Xue F, Zhang N, *et al.* Global burden of liver cirrhosis and other chronic liver diseases caused by specific etiologies from 1990 to 2019. *BMC Public Health* 2024; **24**: 363.

29 European Association for Study of the Liver. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. *J Hepatol* 2021; **75**: 659–89.

30 Newsome PN, Sasso M, Deeks JJ, *et al.* FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol* 2020; **5**: 362–73.

31 Pennisi G, Enea M, Pandolfo A, *et al.* AGILE 3+ Score for the Diagnosis of Advanced Fibrosis and for Predicting Liver-related Events in NAFLD. *Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc* 2023; **21**: 1293-1302.e5.

32 Guha IN, Parkes J, Roderick P, *et al.* Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. *Hepatol Baltim Md* 2008; **47**: 455–60.

33 Sterling RK, Lissen E, Clumeck N, *et al.* Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatol Baltim Md* 2006; **43**: 1317–25.

34 Angulo P, Hui JM, Marchesini G, *et al.* The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatol Baltim Md* 2007; **45**: 846–54.

35 Canivet CM, Costentin C, Irvine KM, *et al.* Validation of the new 2021 EASL algorithm for the noninvasive diagnosis of advanced fibrosis in NAFLD. *Hepatol Baltim Md* 2023; **77**: 920–30.

36 Srivastava A, Gailer R, Tanwar S, *et al.* Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease. *J Hepatol* 2019; **71**: 371–8.

37 Bril F, Gray M. Noninvasive tests to identify liver fibrosis in metabolic dysfunction-associated steatotic liver disease are affected by race. *Obes Silver Spring Md* 2024; **32**: 612–22.

38 Hernán MA. The hazards of hazard ratios. *Epidemiol Camb Mass* 2010; **21**: 13–5.

Table 1. Participant demographic and clinical characteristics

	Development cohort	**Validation cohorts**				**Prognostic validation cohort**
	(n=462)	**DECIDE** (n=6,468)	**Scarred Liver Project** (n=1,367)	**Inter99** (n=390)	**German SLD Registry** (n=711)	**UK Biobank** (n=470,795)
Country	Denmark	Denmark	UK	Denmark	Germany	UK
Age, mean years (SD)	56 (10·4)	56 (8·5)	58 (13·9)	67 (6·3)	51 (13·2)	57 (8.1)
Male sex, n (%)	349 (75%)	3114 (48%)	839 (61%)	153 (54%)	346 (49%)	215,568 (56%)
Ethnicity, % Caucasian	NA	99%	76%*	NA	57%*	NA
Metabolic comorbidities						
BMI (kg/m2), mean (SD)	27 (5·3)	29 (6·0)	29 (5·7)	27 (4·7)	33 (7·1)	NA
Normal weight, (BMI<25) n, (%)	148 (32%)	1444 (22%)	297 (22%)	109 (38%)	51 (7%)	NA
Overweight, (BMI 25-30) n, (%)	173 (37%)	2257 (35%)	444 (32%)	110 (39%)	209 (29%)	NA
Obesity, (BMI>30) n, (%)	141 (31%)	2767 (43%)	626 (46%)	66 (23%)	451 (63%)	NA
Type 2 diabetes, n (%)	64 (14%)	592 (9%)	671 (49%)	29 (7%)	289 (41%)	27,024 (6%)
Dyslipidaemia, n (%)	47 (10%)	2801 (44%)	735 (54%)	181 (46%)	526 (74%)	220,778 (47%)
Hypertension, n (%)	332 (72%)	1878 (29%)	597 (44%)	171 (44%)		
Excessive alcohol consumption, n (%)	462 (100%)	1,851 (29%)	543 (40%)	63 (16%)	NA	10,103 (7%)
Liver stiffness measurements, n (%)	462 (100%)	6,451 (99·7%)	1,357 (93%)	270 (69%)	711 (100%)	NA
Liver stiffness (kPa), median (IQR)	6·5 (4·8–11·7)	4·5 (3·7–5·5)	5·4 (4·3–6·8)	4·6 (3·6–5·6)	6·8 (4·9-11)	NA
TE 8-12 kPa, n(%)	83 (19%)	262 (4%)	162 (12%)	17 (6%)	99 (14%)	NA
TE ≥ 12 kPa, n(%)	123 (27%)	140 (2%)	89 (7%)	5 (1%)	246 (35%)	NA
SLD categories						
MASLD	24 (5%)	2014 (31%)	NA	NA	NA	NA
MetALD	19 (4%)	461 (7·1%)	NA	NA	NA	NA
ALD	278 (60%)	198 (3·1%)	NA	NA	NA	NA
No SLD	140 (30%)	3795 (59%)	NA	NA	NA	NA
Liver biopsy, n (%)**	356 (77%)	239 (3.7%)	NA	NA	127 (18%)	NA
Mild fibrosis (F0–1), n (%)	163 (46%)	77 (1.2%)	NA	NA	57 (45%)	NA
Moderate fibrosis (F2), n (%)	106 (30%)	82 (1.3%)	NA	NA	21 (17%)	NA
Advanced fibrosis (≥F3), n (%)	86 (24%)	80 (1.2%)	NA	NA	48 (39%)	NA

NAS score, mean (SD)	3 (2·1)	3 (1·9)	NA	NA	NA	NA
Degree of steatosis, n (%)						
None (<5%)	158 (45%)	40 (17%)	NA	111 (41%)	25 (4%)	NA
Low (5-33%)	85 (24%)	92 (38%)	NA	69 (26%)	45 (6%)	NA
Moderate (>33-66%)	73 (21%)	69 (29%)	NA	52 (19%)	105 (15%)	NA
Severe (>66%)	39 (11%)	38 (16%)	NA	38 (14%)	323 (45%)	NA
Biochemistry, median (IQR)						
ALT (U/L)	31 (22–48)	25 (19–34)	24 (18–34)	23 (18–31)	48 (32–76)	20 (15–27)
AST (U/L)	34 (25–51)	24 (21–30)	24 (20–30)	31 (28–36)	35 (26–51)	24 (21–29)
ALP (U/L)	80 (66–111)	71 (59–86)	77 (63–94)	NA	82 (66–100)	80 (67–96)
GGT (U/L)	72 (34–190)	26 (17–42)	NA	NA	58 (31–117)	26 (19–41)
INR	1·0 (0·9–1·1)	0·9 (0·9–1·0)	NA	NA	1.0 (0.9–1.0)	NA
Albumin (g/L)	42 (40–45)	45 (44–47)	40 (37–44)	45 (43–46)	45 (41.5–47)	45 (43–47)
Bilirubin (mg/dL)	10 (7–14)	8 (6–11)	10 (7–13)	NA	10 (7–14)	8 (6–10)
Platelets (10^9/L)	232 (186–286)	245 (211–283)	248 (209–291)	246 (215–283)	238 (196–286)	248 (213–287)
Sodium (mmol/L)	140 (138–141)	140 (139–141)	140 (138–141)	141 (140–142)	NA	NA
Cholesterol (mmol/L)	5 (4·3–5·9)	5·1 (4·4–5·8)	5 (3·8–5·5)	5·4 (4·6–6·1)	4·8 (3·9–5·6)	5·7 (4·9–7·2)
HbA1C, (mmol/mol)	36 (33–39)	36 (34–39)	43 (36–54)	38 (36–40)	40 (36–48)	35 (33-38)
Fasting glucose level (mmol/L)	6·2 (5·7–6·9)	5·6 (5.3–6.1)	NA	NA	NA	4·9 (4·6–5·3)
Indirect indices, median (IQR)						
FIB-4 index	1·52 (1·0–2·5)	1·13 (0·9–1·5)	1·17 (0·9–1·6)	1·78 (1·45–2·3)	1·11 (0·72–1·72)	1·25 (0·98–1·58)
NAFLD Fibrosis Score	–0·8 (–1·9–0·3)	–1·43 (–2·2–0·6)	–1·29 (–2·3–0·3)	0·82 (0·1 –1·3)	-1·53 (-2·67– -0·24)	NA
ELF test	9·25 (8·6–10·3)	8·9 (8·4–9·4)	NA	NA	NA	NA
LiverRisk score	7.0 (5·8–9·9)	5·4 (4·8–6·1)	NA	NA	NA	4·8 (4·3–5·5)

Diabetes defined as hemoglobin A1c >48 mmol/L in at least two consecutive measurements or antidiabetic medication prescribed. Obesity defined as BMI ≥30 (kg/m2). Dyslipidemia defined as either low HDL cholesterol (≤1·03 mmol/L for men and ≤1·29 mmol/L for women), high triglycerides (≥1·7 mmol/L), or medical treatment. Excessive alcohol consumption defined as either current, previous, or both with ≥21/14 units per week for men/women over a period of more than 5 years. In the Development cohort, the Danish definition of a unit alcohol of 12g. was used. In the UK Biobank the definition of excessive alcohol consumption is calculated on a subgroup of 148,317 (31%) and based on the definition from the new steatotic liver disease nomenclature, counting >20g/day for female and >30g/day for male. *In Scarred Liver Project information on Ethnicity is obtained in a related project. In German SLD information on ethnicity missing in 40%.

**In the Development cohort, liver biopsy was performed in 363 (79%) of participants. In the DECIDE cohort, liver biopsy was performed in a subgroup of patients with a TE ≥8 kPa. In both cohorts, liver biopsy was not performed when clear signs of cirrhosis were evident based on ultrasound examinations. In the German SLD cohort, liver biopsy was performed on clinical judgement. Degree of steatosis was assessed by liver biopsy as low, moderate, or severe steatosis or from controlled attenuation parameter by TE. ALP: Alkaline phosphatase; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; BMI: Body mass index; dL: Decilitre; ELF: Enhanced Liver Fibrosis; FAST: FibroScan-AST; FIB-4: Fibrosis-4; g: Gram; GGT: Gamma-glutamyl transferase; INR: International normalized ratio; IQR: Interquartile range; kPa: Kilopascal; mg: Milligram; mmol: Millimole; NA: Not available; NAFLD: Non-alcoholic fatty liver disease; SLD: Steatotic liver disease; UK: United Kingdom; U/L: Units per litre.

Table 2. Diagnostic performance (AUC) of LiverPRO and other indices in the development and validation cohorts

	Development cohort	Validation cohorts			
		DECIDE	**Scarred Liver Project***	**Inter99****	**German SLD Registry**
LiverPRO+		n=6,468	n=1,385	n=251	n= 711
≥F2	0·86 (0·83–0·90)	NA	NA	NA	NA
≥F3	0·89 (0·85–0·92)	NA	NA	NA	NA
TE≥8 kPa	0·87 (0·84–0·91)	0·80 (0·78–0·82)	0·69 (0·65–0·72)	0·70 (0·57–0·84)	0·72 (0·68–0·76)
TE≥12 kPa	0·91 (0·89–0·95)	0·86 (0·82–0·89)	0·77 (0·72–0·83)	0·79 (0·48–0·99)	0·74 (0·70–0·77)
FIB-4 index++					
≥F2	0·77 (0·73–0·82)	NA	NA	NA	NA
≥F3	0·84 (0·79–0·89)	NA	NA	NA	NA
TE≥8 kPa	0·79 (0·75–0·84)	0·69 (0·66–0·72)	0·58 (0·53–0·62)	0·57 (0·44–0·69)	0·67 (0·63–0·71)
TE≥12 kPa	0·86 (0·81–0·90)	0·80 (0·76–0·84)	0·70 (0·64–0·76)	0·58 (0·33–0·83)	0·69 (0·64–0·73)
NAFLD Fibrosis Score+++					
≥F2	0·81 (0·76–0·87)	NA	NA	NA	NA
≥F3	0·73 (0·68–0·78)	NA	NA	NA	NA
TE≥8 kPa	0·77 (0·73–0·82)	0·74 (0·72–0·77)	0·66 (0·62–0·69)	0·54 (0·42–0·67)	0·68 (0·63–0·72)
TE≥12 kPa	0·81 (0·76–0·86)	0·82 (0·79–0·85)	0·75 (0·70–0·80)	0·51 (0·10–0·92)	0·65 (0·61–0·70)
ELF test++++					
≥F2	0·84 (0·80–0·88)	NA	NA	NA	NA
≥F3	0·92 (0·89–0·95)	NA	NA	NA	NA
TE≥8 kPa	0·85 (0·81–0·88)	0·78 (0·75–0·80)	NA	NA	NA
TE≥12 kPa	0·94 (0·91–0·96)	0·89 (0·86–0·92)	NA	NA	NA
LiverRisk score+++++					
≥F2	0·81 (0·76–0·85)	NA	NA	NA	NA
≥F3	0·78 (0·73–0·83)	NA	NA	NA	NA
TE≥8 kPa	0.85 (0.82–0.89)	0·81 (0·79–0·84)	NA	NA	NA
TE≥12 kPa	0.85 (0.81–0.89)	0·87 (0·83–0·90)	NA	NA	NA

Numbers are presented as diagnostic performance by AUC and numbers in percentages are 95% confidence intervals (CI).

[+]**LiverPRO** had 2 mising in the Development cohort, and none missing across the other cohorts. [*]LiverPRO is calculated in the Scarred Liver Project based on a 7-variable model comprising AST, ALB, ALP, bilirubin, cholesterol, sodium, and platelets.[**]LiverPRO is calculated in the Inter99 cohort based on a 5-variable model comprising AST, ALB, cholesterol, sodium, and platelets. [++]**FIB-4** had 18 mising in the Development cohort, 249 in the DECIDE cohort, 170 in the Scarred Liver Project, 132 in the Inter99 cohort, and 10 in the German SLD Registry. [+++]**NAFLD Fibrosis score** had 25 mising in the Development cohort, 287 in the DECIDE cohort, 187 in the Scarred Liver Project, 155 in the Inter99 cohort, and 104 in the German SLD Registry. [++++]**ELF test** had 5 mising in the Development cohort, and 3,095 in the DECIDE cohort. ELF test was not analysed in the Scarred Liver Project, the Inter99 cohort, and in the German SLD Registry. [+++++]**LiverRisk score** had 15 mising in the Development cohort, and 344 in the DECIDE cohort. LiverRisk score was not analysed in the Scarred Liver Project, the Inter99 cohort, and the German SLD Registry.

AUC: Area under the curve; CI: Confidence interval; ELF: Enhanced Liver Fibrosis; FIB-4: Fibrosis-4; NA: Not available; NAFLD: Non-alcoholic fatty liver disease; TE: Transient elastography.

Table 3. Clinical performance of LiverPRO and other indices to predict significant liver fibrosis in included cohorts

	Rule-out*			Grey zone	Rule-in*			
	n	Sensitivity	Specificity	NPV	n	n	Sensitivity	Specificity	PPV
Development cohort									
LiverPRO	113 (25%)	93·2% (88·7–96·3)	39·1% (33·0–45·3)	88·5% (81·1–93·7)	197 (43%)	152 (33%)	59·4% (52·1–66·4)	88·6% (84·1–92·3)	79·7% (72·2–86·0)
FIB-4 index	178 (40%)	80·1% (73·6–85·6)	56·4% (50·1–62·9)	78·7% (71·9–84·6)	164 (37%)	102 (23%)	44·1% (36·8–51·5)	95·9% (92·6–98·0)	89·1% (80·9–94·7)
NAFLD Fibrosis Score	149 (34%)	83·2% (77·1–88·3)	49·2% (42·6–55·7)	79·1% (71·6–85·3)	207 (48%)	80 (18%)	36·2% (29·3–43·6)	98·7% (96·4–99·7)	95·7% (88·0–99·1)
ELF test	300 (66%)	63·4% (56·1–70·2)	90·0% (85·7–93·5)	76.4% (71·2–81·1)	51 (11%)	106 (23%)	48·7% (41·4–56·0)	97·6% (94·9–99·1)	93·9% (87·3–97·7)
LiverRisk score	270 (61%)	67·6% (60·4 –74·2)	84·0% (78·8–88·4)	77·1% (71·5–82.0)	104 (23%)	72 (16%)	30·9% (24·3–38·0)	96·7% (93·6–98·6)	87·9% (77·5–94·6)
DECIDE cohort									
LiverPRO	3,899 (60%)	80·6% (76·4–84·3)	63·1% (61·8–64·3)	98% (97·5–98·4)	2,157 (33%)	412 (6%)	33·6% (29·0–38·4)	95·5% (94·9–96·0)	33·0% (28·5–37·8)
FIB-4 index	4,087 (66%)	53·8% (48·5–58·9)	66·9% (65·7–68·2)	95·8% (95·1–96·4)	2,014 (32%)	118 (2%)	11·3% (8·3–15·0)	98·7% (98·4–99·0)	35·6% (27·0–44·9)
NAFLD Fibrosis Score	3,035 (49%)	80·8% (76·4–84·7)	51·1% (49·8–52·3)	97·7% (97·1–98·2)	2,936 (48%)	210 (3%)	16·5% (12·9–20·7)	97·5% (97·0–97·9)	29·5% (23·4–36·2)
ELF test	2,889 (86%)	45·9% (39·5–52·4)	88·1% (86·9–89·2)	95·5% (94·6–96·2)	358 (11%)	126 (4%)	22·7% (17·6–28·5)	97·7% (97·1–98·2)	43·7% (34·8–52·8)
LiverRisk score	5,783 (94%)	35·3% (30·4–40·5)	96·3% (95·8–96·8)	95·9% (95·4–96·4)	289 (5%)	52 (1%)	10·1% (7·2–13·7)	99·7% (99·6–99·9)	71·2% (56·9–82·9)
SLP cohort									
LiverPRO	304 (22%)	88·5% (84·0–92·1)	24·8% (22·3–27·4)	90·1% (86·2–93·2)	668 (49%)	395 (29%)	49·4% (43·2–55·7)	75·9% (73·3–78·4)	32·7% (28·1–37·5)

FIB-4 index	709 (52%)	50·4% (44·0–56·8)	61·8% (58·6–64·9)	82·5% (79·5–85·2)	427 (31%)	231 (17%)	10·8% (7·2–15·3)	96·4% (95·0–97·5)	44·3% (31·5–57·6)
NAFLD Fibrosis Score	438 (32%)	77·5% (71·8–82·5)	41·0% (37·9–44·3)	87·2% (83·7–90·2)	565 (41%)	177 (13%)	23·3% (18·2–29·0)	87·2% (84·9–89·3)	32·8% (25·9–40·2)
Inter99									
LiverPRO	134 (34%)	90·9% (70·8–98·9)	9·7% (6·3–14·1)	92·3% (74·9–99·1)	201 (52%)	55 (14%)	36·4% (17·2–59·3)	83·1% (77·8–87·5)	16·0% (7·2–29·1)
FIB-4 index	43 (11%)	95·2% (76·2–99·9)	18·2% (13·4–23·9)	97·6% (87·4–99·9%)	184 (47%)	163 (42%)	19·0% (5·5–41·9)	89·8% (85·1–93·4)	14·8% (4·2–33·7)
NAFLD Fibrosis Score	5 (1·3%)	100% (82·4–100)	2·4% (0·8–5·6)	100% (47·8–100)	104 (27%)	126 (32%)	68·4% (43·4–87·4)	48·1% (41·1–55·1)	10·8% (5·9–17·8)
German SLD									
LiverPRO	241 (34%)	79·1% (74·5–83·3)	46·2% (41·0–51·4)	70·1% (63·9–75·8)	328 (46%)	142 (20%)	33·0% (28·1–38·3)	92·3% (89·1–94·9)	80·3% (72·8–86·5)
FIB-4 index	435 (61%)	52·8% (47·3–58·3)	75·7% (71·0–80·0)	63·7% (59·0–68·2)	178 (25%)	88 (12%)	23·6% (19·1–28·5)	97·5 (95·4–98·9)	89·9 (81·5–95·2)
NAFLD Fibrosis Score	314 (44%)	59·9% (54·0–65·5)	62·6% (57·0–68·0)	62·4% (56·8–67·8)	219 (31%)	74 (10%)	21·8% (17·2–26·9)	96·8% (94·2–98·5)	86·5% (76·5–93·3)

*Cut-offs are as follows: LiverPRO: <25% rule-out, >65% rule-in; FIB-4: 1·3 rule-out, >2·67 rule-in; NAFLD Fibrosis Score: –1·46 rule-out, >0·68 rule-in; ELF test: 9·8 rule-out, 10·5 rule-in; LiverRisk score: <8 rule out, >12 rule in. Numbers in percentages are 95% confidence intervals (CI). ELF test and LiverRisk score was not calculated in the following three cohorts: SLP cohort, Inter99 and in German SLD.

ELF: Enhanced Liver Fibrosis; FIB-4: Fibrosis-4; kPa: Kilopascal; NAFLD: Non-alcoholic fatty liver disease; NPV: Negative predictive value; PPV: Positive predictive value; TE: Transient elastography.

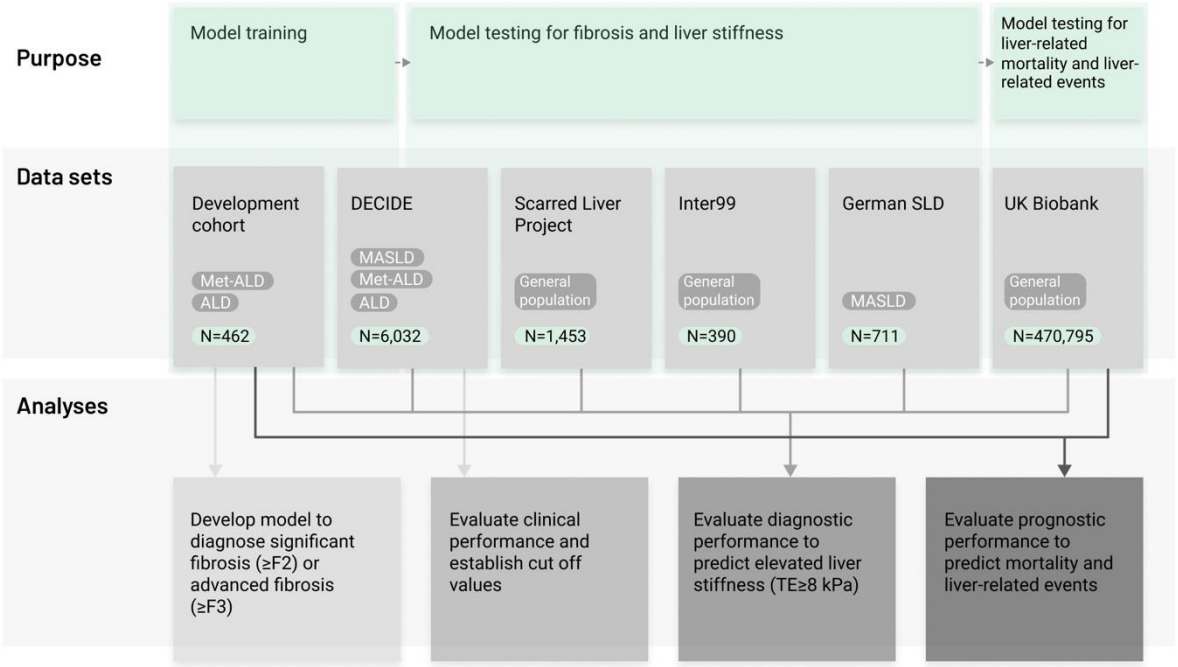
Table 4. Prognostic performance* of LiverPRO and other indices in the development cohort and UK Biobank

	Development cohort (n=462)				UK Biobank (n=470,795)			
	Liver-related mortality		Liver-related		Liver-related mortality		Liver-related events	
LiverPRO								
Harrel's C	0·79 (0·71–0·86)		0·78 (0·73–0·84)		0·75 (0·73–0·76)		0·74 (0·73–0·75)	
Harrel's C (2 years)	0·87 (0·78–0·97)		0·86 (0·80–0·92)		0·78 (0·72–0·86)		0·80 (0·77–0·84)	
Harrel's C (5 year)	0·76 (0·67–0·86)		0·78 (0·72–0·84)		0·72 (0·69–0·76)		0·77 (0·75–0·79)	
HR Low risk	1 (Ref)		1 (Ref)		1 (Ref)		1 (Ref)	
HR Moderate risk	2·2 (0·6–7·8)		2·3 (0·9–5·6)		2·3 (2·1–2·6)		2·5 (2·2–2·7)	
HR high risk	7·8 (2·4–25·7)		8·6 (3·7–19·9)		20·1 (17·7–22·9)		17·1 (15.6–18·9)	
Events	Events	Total n	Events	Total n	Events	Total n	Events	Total n
- Low risk	3 (7%)	113	6 (7%)	113	460 (29%)	265441	742 (26%)	265661
- Moderate	11 (25%)	194	23 (27%)	194	744 (47%)	182443	1226 (43%)	182657
- High risk	30 (68%)	152	55 (66%)	152	365 (23%)	20840	891 (31%)	20641
Total	44	459	84	459	1569	455775	2859	468959
FIB-4								
Harrel's C	0.75 (0.67–0.84)		0·82 (0·77–0·87)		0·71 (0·69–0·72)		0·67 (0·66–0·68)	
Harrel's C (2 years)	0·80 (0·66–0·94)		0·87 (0·81–0·93)		0·72 (0·63–0·79)		0·75 (0·71–0·79)	
Harrel's C (5 year)	0·73 (0·62–0·85)		0·83 (0·76–0·90)		0·72 (0·68–0·76)		0·73 (0·71–0·75)	
HR Low risk	1 (Ref)		1 (Ref)		1 (Ref)		1 (Ref)	
HR Moderate risk	1·3 (0·5–3·5)		1·8 (0·8–3·8)		2·2 (1·9–2·4)		1·8 (1·7–2·0)	
HR high risk	7.1 (3·2–15·5)		14·1 (7·3–26·9)		21·6 (18·8–24·8)		15·2 (13·7–16·9)	
Events	Events	Total n	Events	Total n	Events	Total n	Events	Total n
- Low risk	8 (18%)	177	11 (13%)	177	460 (29%)	249455	968 (34%)	249639
- Moderate	8 (18%)	164	16 (19%)	164	744 (47%)	195844	1333 (47%)	196082
- High risk	28 (64%)	102	56 (68%)	102	365 (23%)	10476	550 (19%)	10570
Total	44	443	83	443	1569	455775	2851	455775
NAFLD Fibrosis Score								
Harrel's C	0·71 (0·61–0·81)		0·77 (0·71–0·83)		NA		NA	
Harrel's C (2 years)	0·77 (0·59–0·94)		0·84 (0·77–0·91)		NA		NA	
Harrel's C (5 year)	0·69 (0·56–0·81)		0·76 (0·67–0·84)		NA		NA	
HR Low risk	1 (Ref)		1 (Ref)		NA		NA	
HR Moderate risk	0·7 (0·3–1·6)		1·2 (0·6–2·2)		NA		NA	
HR high risk	5·2 (2·5–10·8)		9·9 (5·4–18·1)		NA		NA	
Events	Events	Total n	Events	Total n	NA	NA	NA	NA
- Low risk	10 (23%)	149	14 (17%)	149				
- Moderate	9 (21%)	207	22 (27%)	207				
- High risk	24 (56%)	80	45 (56%)	80				
Total	43	436	81	436	NA	NA	NA	NA
ELF test								
Harrel's C	0·85 (0·81–0·90)		0·86 (0·82–0·90)		NA		NA	
Harrel's C (2 years)	0·90 (0·86–0·94)		0·90 (0·86–0·95)		NA		NA	
Harrel's C (5 year)	0·83 (0·76–0·91)		0·87 (0·82–0·92)		NA		NA	
HR Low risk	1 (Ref)		1 (Ref)		NA		NA	
HR Moderate risk	3·8 (1·2–12·0)		4·4 (2·1–9·5)		NA		NA	
HR high risk	16·0 (7·1–36·3)		17·9 (10·1–31·8)		NA		NA	

Events - Low risk - Moderate - High risk	Events 7 (16%) 5 (11%) 32 (73%)	Total n 299 55 102	Events 15 (18%) 12 (14%) 56 (68%)	Total n 299 55 102	NA	NA	NA	NA
Total	44	456	83	456	NA	NA	NA	NA
LiverRisk score								
Harrel's C	0·75 (0·69–0·82)		0.79 (0.75–0.83)		0·73 (0·71–0·74)		0·72 (0·71–0·73)	
Harrel's C (2 years)	0·80 (0·70–0·90)		0.84 (0.78–0.89)		0·79 (0·71–0·87)		0·81 (0·77–0·85)	
Harrel's C (5 year)	0·74 (0·66–0·82)		0.80 (0.74–0.86)		0·71 (0·67–0·75)		0·76 (0·74–0·78)	
HR Low risk	1 (Ref)		1 (Ref)		1 (Ref)		1 (Ref)	
HR Moderate risk	2·9 (1·3–6·2)		3·7 (2·1–6·4)		12·3 (10·7–14·2)		11·2 (10·1–12·4)	
HR high risk	6·1 (3·0–12·7)		7·8 (4·6–13·4)		58·0 (49·2–68·4)		43·3 (37·7–49·6)	
Events - Low risk - Moderate - High risk	Event 12 (27%) 14 (32%) 18 (41%)	Total 270 104 72	Event 22 28 33	Total 270 104 72	Events 995 (71%) 247 (18%) 166 (12%)	Total n 406560 8958 1446	Events 1919 (74%) 433 (17%) 231 (9%)	Total n 406560 8958 1446
Total	44	446	83	446	1408	416964	1684	416964

*Presented as Harrell's C statistic and hazard ratios (HR) with 95% confidence intervals. Participants with events, shows the total number of participants with a liver-related event, in-between the three risk groups for each indices. Hazard ratios compared high risk vs. low risk. Cut-offs used were LiverPRO >65% vs. <25%. For FIB-4 >2·67 vs. <1·3. For ELF test >10·5 vs. <9·8. For NFS >0·68 vs. <–1·46 LiverRisk score >12 vs. <8.
ELF: Enhanced Liver Fibrosis; FAST: FibroScan-AST; FIB-4: Fibrosis-4; LRE: Liver-related event; NA: Not available; NAFLD: Non-alcoholic fatty liver disease.

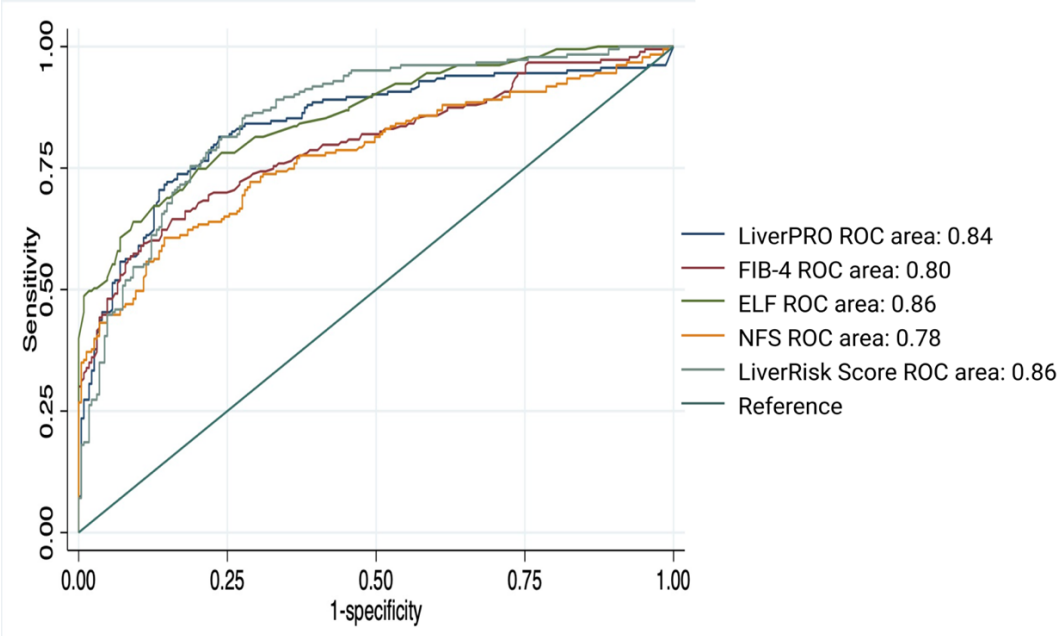
Figure 1. Study design. We developed our model in a prospective, biopsy-controlled cohort which included both MetALD and ALD patients (development cohort). The model was validated in four independent cohorts from Denmark (DECIDE and Inter99 study cohorts), Germany (German SLD Registry), and the United Kingdom (Scarred Liver Project cohort). The DECIDE cohort included patients with MASLD, MetALD, and ALD. The and German SLD Registry only included patients with MASLD, whereas the Inter99 and Scarred Liver Project cohorts included participants from the general population. The UK Biobank, a large-scale biomedical database and research resource containing genetic, lifestyle, and health information from UK participants, served as the prognostic evaluation cohort.



ALD: Alcohol-related liver disease; MASLD: Metabolic dysfunction-associated steatotic liver disease; MetALD: Metabolic and alcohol related/associated liver disease; SLD: Steatotic liver disease; TE: Transient elastography; UK: United Kingdom.

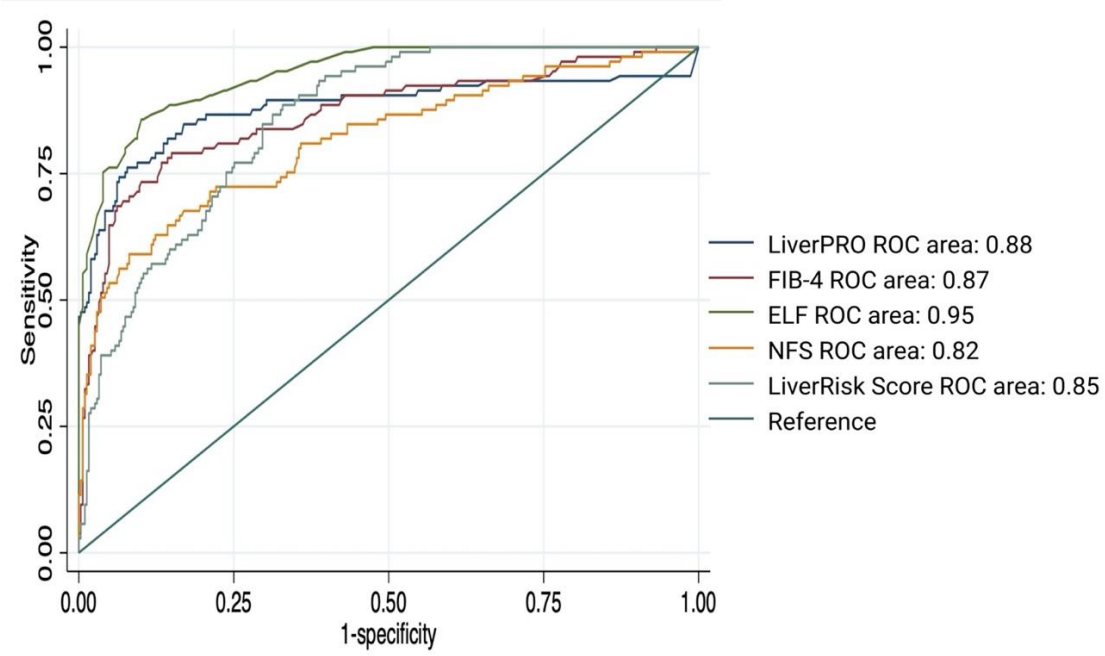
Figure 2A-D. Area under the receiver operating characteristic curve for LiverPRO in the development and DECIDE cohorts. A) Significant fibrosis (≥F2) in the development cohort (n=462); B) Advanced fibrosis (≥F3) in the development cohort (n=462); C) Transient elastography (TE ≥8 kPa) in the DECIDE cohort (n=6,468); D) Transient elastography (TE ≥12 kPa) in the DECIDE cohort (n=6,468). P-values tested using DeLong method. P-values tested using DeLong method.

A) Significant fibrosis (≥F2) in the development cohort



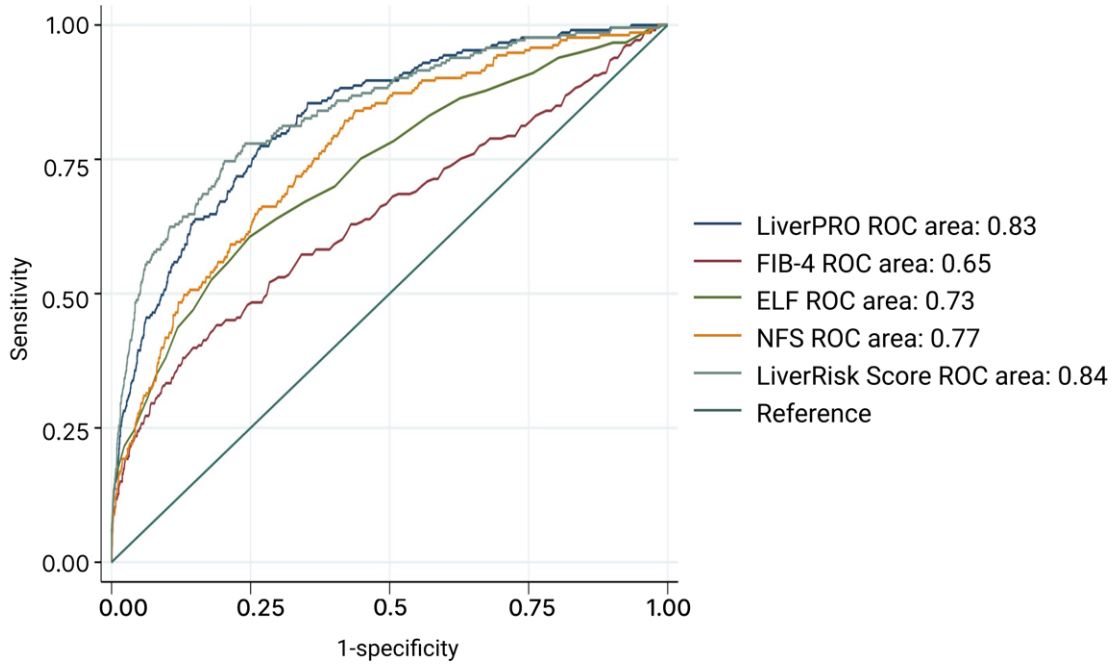
LiverPRO vs. FIB-4: p=0.0027, LiverPRO vs. ELF: p=0.3896, LiverPRO vs. NFS: p=0.0011, FIB-4 vs. ELF: p=0.0034 FIB-4 vs. NFS: p=0.0101, ELF vs. NFS: p=0.0000

B) Advanced fibrosis (≥F3) in the development cohort



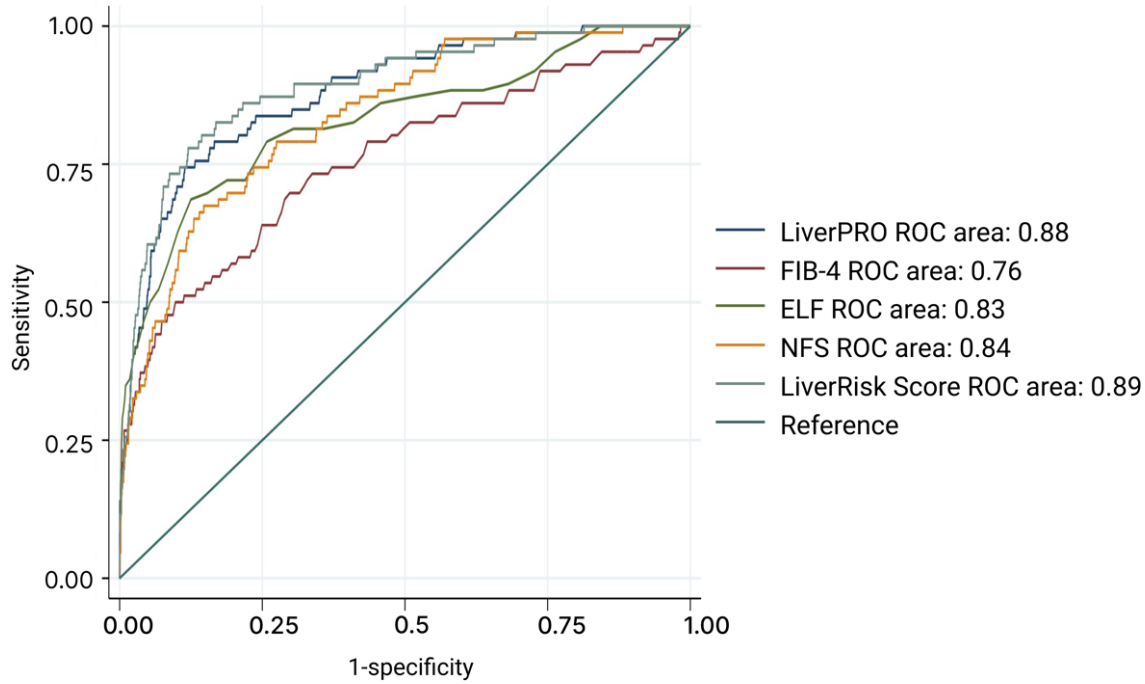
LiverPRO vs. FIB-4: p=0.5911, LiverPRO vs. ELF: p=0.0019, LiverPRO vs. NFS: p=0.2482, FIB-4 vs. ELF: p=0.0004, FIB-4 vs. NFS: p=0.1668, ELF vs. NFS: p=0.0001

C) Transient elastography (TE ≥8 kPa) in the DECIDE cohort



LiverPRO vs. FIB-4: p=0.0000, LiverPRO vs. ELF: p=0.0000, LiverPRO vs. NFS: p=0.0000, LiverRisk Score vs. LiverPRO: p=0.0000, FIB-4 vs. ELF: p=0.0001, FIB-4 vs. NFS: p=0.0000, ELF vs. NFS: p=0.0528

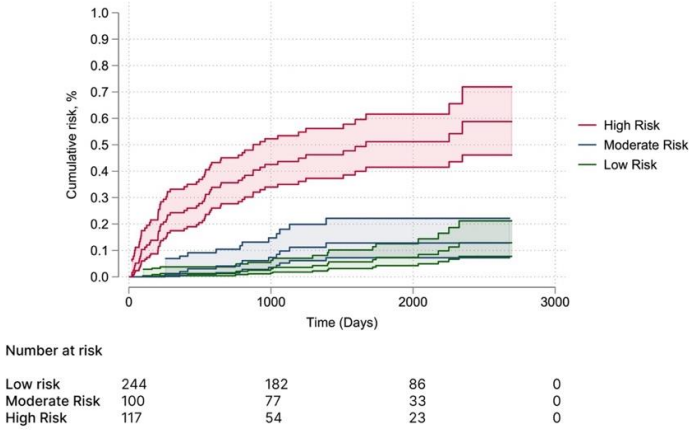
D) Transient elastography (TE ≥12 kPa) in the DECIDE cohort



LiverPRO vs. FIB-4: p=0.0000, LiverPRO vs. ELF: p=0.0611, LiverPRO vs. NFS: p=0.0116, LiverRisk Score vs. LiverPRO: p=0.0000, FIB-4 vs. ELF: p=0.0012, FIB-4 vs. NFS: p=0.0002, ELF vs. NFS: p=0.8071

Figure 3A-D. Cumulative incidence for liver-related events and liver-related mortality in the development cohort for LiverPRO and FIB-4

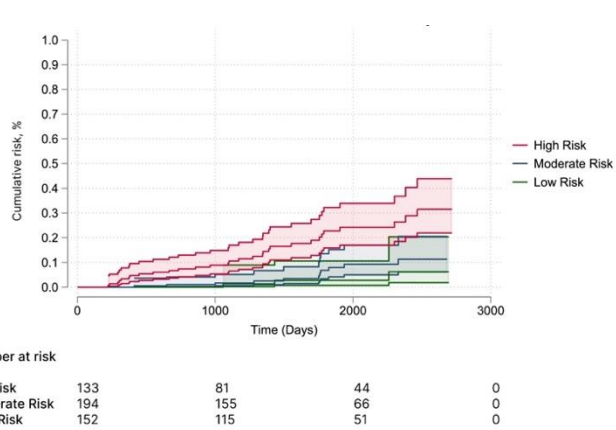
A) Liver-related events in the development cohort for LiverPRO



Number at risk

Low risk	244	182	86	0
Moderate Risk	100	77	33	0
High Risk	117	54	23	0

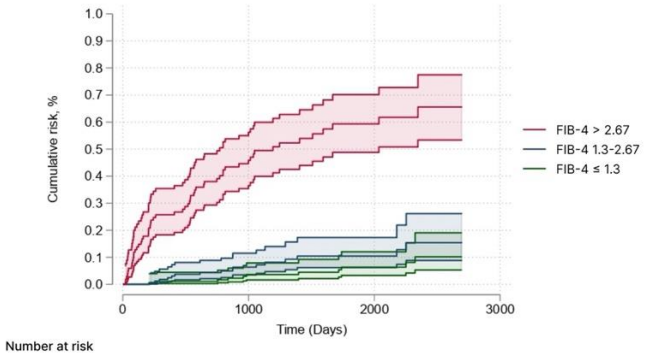
B) Liver-related mortality in the development cohort for LiverPRO



Number at risk

Low risk	133	81	44	0
Moderate Risk	194	155	66	0
High Risk	152	115	51	0

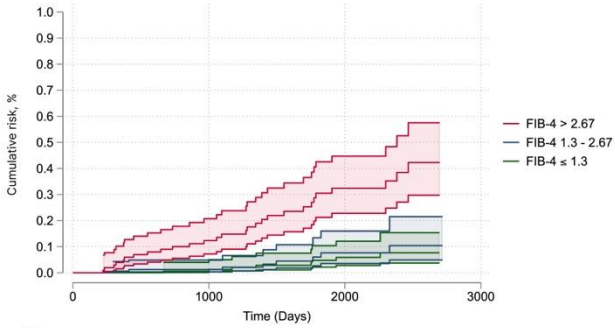
C) Liver-related events in the development cohort for FIB-4



Number at risk

FIB-4 ≤ 1.3	177	138	75	0
FIB-4 1.3-2.67	164	120	48	0
FIB-4 > 2.67	102	45	18	0

D) Liver-related mortality in the development cohort for FIB-4



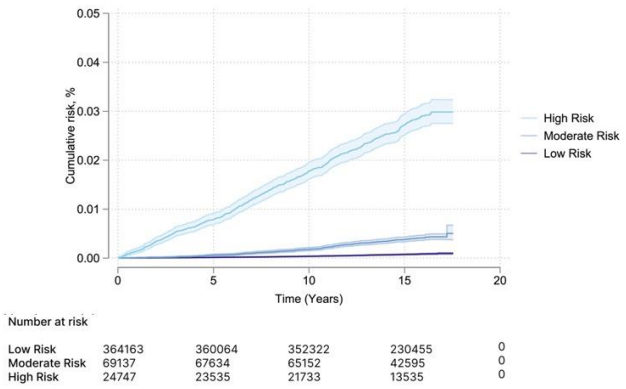
Number at risk

FIB-4 ≤ 1.3	177	143	76	0
FIB-4 1.3 - 2.67	164	125	50	0
FIB-4 > 2.67	102	74	35	0

Tests are compared using Log-rank test and p-values reported, and shaded areas are 95% confidence bands. A) Liver-related events in the development cohort for LiverPRO (p-value <0.001); B) Liver-related mortality in the development cohort for LiverPRO (p-value <0.001); C) Liver-related events in the development cohort for FIB-4 (p-value <0.001); D) Liver-related mortality in the development cohort for FIB-4 (p-value < 0.001).

Figure 4 A-F. Cumulative incidence for liver-related events and liver-related mortality in the UK Biobank cohort for LiverPRO, FIB-4, and LiverRisk score

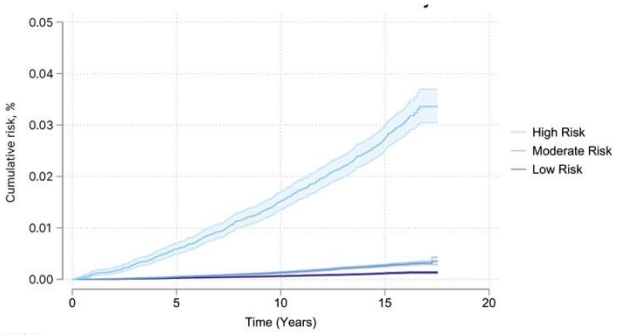
A) Liver-related events in the UK biobank for LiverPRO



Number at risk

Low Risk	364163	360064	352322	230455	0
Moderate Risk	69137	67634	65152	42595	0
High Risk	24747	23535	21733	13535	0

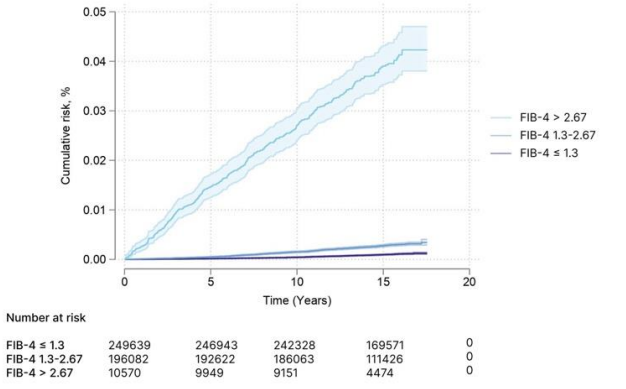
B) Liver-related mortality in the UK biobank for LiverPRO



Number at risk

Low risk	258717	256523	252120	165431	0
Moderate Risk	178298	174971	169057	110073	0
High Risk	20422	19297	17780	11022	0

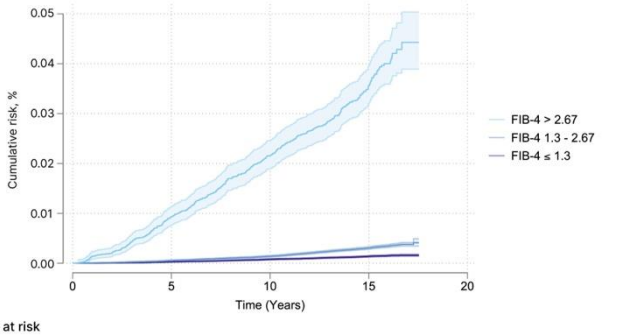
C) Liver-related events in the UK biobank for FIB-4



Number at risk

FIB-4 ≤ 1.3	249639	246943	242328	169571	0
FIB-4 1.3-2.67	196082	192622	186063	111426	0
FIB-4 > 2.67	10570	9949	9151	4474	0

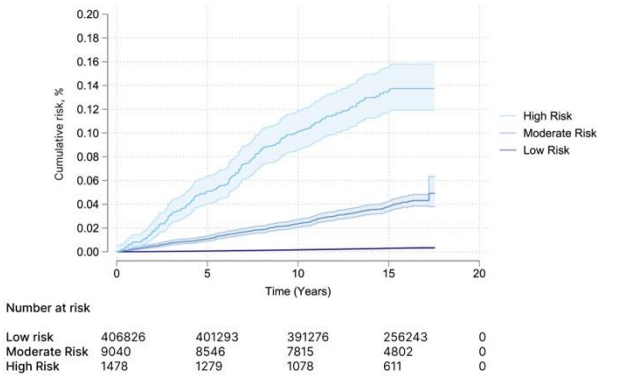
D) Liver-related mortality in the UK biobank for FIB-4



Number at risk

FIB-4 ≤ 1.3	249697	247023	242445	169706	0
FIB-4 1.3 - 2.67	196196	192794	186360	111659	0
FIB-4 > 2.67	10680	10133	9346	4593	0

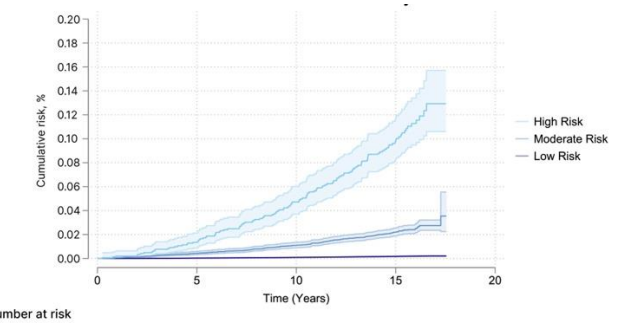
E) Liver-related events in the UK biobank for LiverRisk score



Number at risk

Low risk	406826	401293	391276	256243	0
Moderate Risk	9040	8546	7815	4802	0
High Risk	1478	1279	1078	611	0

F) Liver-related mortality in the UK biobank for LiverRisk score



Number at risk

Low risk	406826	401431	391581	256568	0
Moderate Risk	9040	8604	7925	4892	0
High Risk	1478	1331	1152	651	0

Tests are compared using Log-rank test and p-values reported, and shaded areas are 95% confidence bands. A) Liver-related events in the UK biobank for LiverPRO (p-value <0.0001); B) Liver-related mortality in the UK biobank for LiverPRO (p-value < 0.001); C) Liver-related events in the UK biobank for FIB-4 (p-value <0.0001); D) Liver-related mortality in the UK biobank for FIB-4 (p-value <0.0001); E) Liver-related events in the UK biobank for LiverRisk score (p-value <0.0001); F) Liver-related mortality in the UK biobank for LiverRisk score (p-value <0.001).