A Risk Score to Predict Long-Term Liver-Related Outcomes in the General Population

Miquel Serra-Burriel, PhD, Adrià Juanola, MD, Feliu Serra-Burriel, PhD, Prof. Maja Thiele, MD, Isabel Graupera, MD, Elisa Pose, MD, Guillem Pera, PhD, Prof. Ivica Grgurevic, MD, Llorenç Caballeria, MD, Salvatore Piano, MD, Laurens A. van Kleef, MD, Mathias Reichert, MD, Prof. Dominique Roulot, MD, Juan M Pericàs, MD, Prof. Jörn M Schattenberg, MD, Prof. Emmanuel A Tsochatztis, MD, Prof. Indra Neil Guha, MD, Montserrat Garcia-Retortillo, MD, Rosario Hernández-Ibáñez MD, Jordi Hoyo MD, Matilde Fuentes MD, Carmen Expósito MD, Alba Martínez MD, Patricia Such MD, Anita Madir MD, Sönke Detlefsen MD, Marta Tonon MD, Andrea Martini MD, Ann T Ma, MD, Judith Pich BSc, Eva Bonfill MSc, Marta Juan, Anna Soria MD, Marta Carol, PhD, Jordi Gratacós-Ginès, MD, Rosa M Morillas, MD, Pere Toran, MD, JM Navarrete, MD, Antoni Torrejón, MD, MD, Céline Fournier, MD, Anne Llorca, MD, Anita Arslanow, PhD, Prof. Harry J de Koning, PhD, Fernando Cuccietti, PhD, Prof. Michael Manns, MD, Prof. Philip N Newsome, MD, Rubén Hernáez, MD, Prof. Alina M Allen, MD, Prof. Paolo Angeli, MD, Prof. Robert J de Knegt MD, Prof. Tom H Karlsen, MD, Prof. Peter Galle, MD, Prof. Vincent Wai-Sung Wong, MD, Prof. Núria Fabrellas PhD, Prof. Laurent Castera MD, Prof. Aleksander Krag MD, Prof. Frank Lammert MD, Prof. Patrick S Kamath, MD, and Prof. Pere Ginès, MD for the LiverScreen Consortium investigators*

*The list of the LiverScreen Consortium investigators is shown in the appendix

Tables: 2

Figures: 4

Supplementary Tables: 5

Supplementary Figures: 16

Word Count: 3,365

Summary

Background

Liver cirrhosis is a major cause of death worldwide. Cirrhosis develops after a long asymptomatic period of fibrosis progression, with the diagnosis frequently occurring late when major complications or cancer develop. Currently, there is a lack of reliable tools for timely identification of subjects at risk of cirrhosis so as to allow for early intervention. We aimed to develop a novel score to identify subjects at risk for future liver-related outcomes.

Methods

The score was derived from an international prospective cohort of 6,357 subjects without known liver disease from general population, who underwent liver fibrosis assessment by transient elastography. The model's discriminatory accuracy and calibration were externally validated in two prospective cohorts including 8,369 subjects from general population. Moreover, prognostic value in the prediction of liver-related outcomes was ascertained in 416,200 participants without known liver disease with median follow-up of 12 years (UK Biobank cohort).

Findings

The score, composed of age, gender, and six standard laboratory variables, accurately predicted liver stiffness in development and external validation cohorts, and was superior to conventional serum biomarkers of fibrosis, as measured by AUC (0.83 95%CI[0.78 to 0.89] vs FIB4 (0.68 95%CI[0.61 to 0.75] at 10kPa). The score was effective in identifying subjects at risk of liver-related mortality and hospitalization, and liver cancer, thereby allowing stratification to different risk groups for liver-related outcomes. The hazard ratio for liver-related mortality in the high-risk group was 471 (95%CI 234 to 590) compared to the minimal risk group, and overall AUC of the score in predicting 10-year liver-related mortality was 0.90 95%CI(0.88 to 0.91) vs FIB4 0.84 95%CI(0.82 to 0.86).

Interpretation

A ©LiverRisk score based on simple parameters predicts liver fibrosis and future development of liver-related outcomes in the general population. The score may allow for stratification of individual subjects according to liver risk and thus guide preventive care.

INTRODUCTION

Liver cirrhosis accounts for 2.4% of yearly deaths worldwide and is associated with a significant economic burden for healthcare systems.[1] Notably, cirrhosis is the second cause of years of life lost in European countries[2]. Moreover, cirrhosis may lead to hepatocellular carcinoma, the incidence of which is increasing in many areas of the world.[3] Cirrhosis, characterized by diffuse hepatic fibrosis with nodular regeneration, is the final consequence of any chronic inflammatory process in the liver that may be caused by different factors, particularly hepatitis virus, alcohol, or metabolic syndrome, the latter currently known as non-alcoholic fatty liver disease (NAFLD). Persistent liver inflammation is clinically silent but may result in liver fibrosis, eventually leading to cirrhosis. Although this process takes years or decades, the diagnosis is generally made only at later stages when the disease becomes symptomatic and patients develop severe complications related to liver failure or portal hypertension that require multiple hospitalizations, or liver cancer.[2,4] The vast majority of these symptomatic patients die of liver disease unless liver transplantation is performed. Although the prevalence of cirrhosis due to hepatitis C virus infection is decreasing worldwide due to extremely effective oral antiviral drugs, that of NAFLD is increasing markedly, owing to the epidemics of obesity and type-2 diabetes mellitus.[2,4]

Early identification of individuals at risk for progressive fibrosis would enable lifestyle modifications or therapeutic interventions to prevent the development of cirrhosis, and would facilitate selection of patients for specialist referral. However, current non-invasive tools for identification of subjects in the population at risk for progressive hepatic fibrosis and, therefore, the long-term development of cirrhosis and liver-related death have significant limitations.[5] Techniques such as transient electrography that measure liver stiffness, a surrogate for hepatic fibrosis, are accurate, but application of elastography to population screening is limited by expense and lack of availability outside of specialist settings.[2,4] Risk scores based on

liver blood tests, such as fibrosis-4 index (FIB-4) or aspartate aminotransferase-AST to platelet ratio index (APRI) show some utility in predicting the long-term development of cirrhosis or liver-related death in the general population.[5] However, because these indices were designed for fibrosis assessment in patients with hepatitis C virus infection and high prevalence of fibrosis, their predictive accuracy for the general population is modest.[5,6]

Hence, there is an unmet medical need to develop more accurate tools using easily available laboratory or clinical variables for the identification of subjects at risk for the long-term development of cirrhosis, liver-related complications, and death. Such predictive tools would enable case-finding and individualized follow up for persons with progressive liver disease in primary care and other non-liver health care settings, before development of cirrhosis or its complications, and subsequent allow application of preventive measures such as weight loss in overweight/obese patients with NAFLD and alcohol rehabilitation in patients with high alcohol consumption [7,8]. Therefore, the aim of the current study was to develop a liver risk score to identify subjects at risk for future liver-related outcomes.

PATIENTS AND METHODS

The current study consists of two distinct parts. The aim of the first part was to develop and validate a diagnostic liver risk score ("LiverRisk score") in subjects from the general population which predicted individual values of liver stiffness by using a combination of standard demographic, clinical, and/or laboratory variables. The aim of the second part was to assess whether the LiverRisk score is useful for the prediction of future liver-related outcomes in individuals without known liver disease in the general population.

Patient population

Derivation cohort for the LiverRisk score

Patient-level data from seven independent prospective studies using transient elastography to assess liver stiffness were used in the development of the model aimed at predicting the presence of liver fibrosis. These studies include subjects from Denmark,[9] Hong Kong,[10] Germany,[11] France,[12] United Kingdom,[13] and Spain.[14,15]

Information on gender, age, alcohol consumption, body mass index (BMI), waist circumference, arterial pressure, diabetes mellitus, arterial hypertension, fasting glucose, creatinine, total cholesterol, HDL-cholesterol, triglycerides, aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma-glutamyl transferase (GGT), bilirubin, leukocyte levels, hemoglobin, and platelet count were available from these databases. The outcome of interest was a validated liver stiffness value (in kPa), as measured by transient elastography[5]. All quantitative measurements, including biomarkers and liver stiffness were standardised across all cohorts. A model was developed from this cohort to generate a LiverRisk score that was predictive of the measured liver stiffness value (in kPa), which estimates the presence of hepatic fibrosis[5].

Validation cohorts for the LiverRisk score

The LiverRisk score obtained from the derivation cohort was validated in two external cohorts. The first external validation cohort included participants in the Rotterdam Study[16], a population-based study of subjects older than 45 years who underwent liver stiffness measurement, while the second validation cohort included subjects participating in the LiverScreen Study, a European multicenter prospective diagnostic study also assessing the presence of liver fibrosis in the population using liver stiffness by transient elastography.[17]

Prognostic evaluation cohort of the LiverRisk score

The prognostic cohort was obtained from the UK Biobank dataset.[18] The UK Biobank is a large population-based cohort that includes over 500,000 individuals with baseline demographic, serologic, lifestyle, and genetic measurements initiated in 2007. UK Biobank collects information on all participants which includes baseline demographic, environmental, and lifestyle characteristics of all individuals, as well as information on hospitalizations and death from all participants. Data of death, including primary and secondary causes of death, are recorded from ICD-10 codes from the death registry and are updated two to three times every year. Data on hospitalizations are also based on ICD-10 codes and updated every year. Exclusion criteria for our evaluation included: diagnosis of liver disease before enrollment (n=3,471), diagnosis of viral hepatitis at baseline or at any point during follow-up (n=541), and incomplete laboratory variables (n=86,263). We conducted a complete case analysis only in the cohort without missing variables.

The evaluated outcomes included: liver-related mortality, first liver-related hospitalization, and incident liver cancer.[2] We also selected non-liver-related mortality, first non-liver-related hospitalization, and incident cancer as negative control outcomes. **Statistical analysis**

Model development

Variable selection in the development sample was performed using a recursive feature elimination (RFE) algorithm.[19] RFE is a technique that ranks the most relevant predictors in a dataset, by training models with and without all potential predictor combinations. Next, to determine the optimal number of variables we assessed the incremental gain in predictive performance associated with each variable and stopped at the inflexion point. After variable selection with RFE, we trained four statistical models with centered and scaled selected predictors, due to the different scales of the predictors and to ease the intercept of models to the mean liver stiffness, including: a linear regression model (LM),[20] quantile regression model (QR),[21] gradient boosting model (GBM),[22] and a random forest model (RF).[23] No further functional transformations of the predictors were used, and no interaction terms were included in the linear model. To assess the degree of potential over-fitting of each algorithm, we trained them using a 5-fold 5-repeat cross-validation procedure. The sample size considerations for model development are shown in supplementary Table S1.

Model evaluation

To assess the discriminatory accuracy of the developed model, in all three diagnostic cohorts, we used the area under the receiver-operating characteristics curve (AUROC) at 3 values of the LiverRisk score that estimate levels of fibrosis severity in population-based studies: ≥6, ≥9, and ≥15 kPa thresholds.[24-26] Using these cutoffs we categorized the subjects into 4 risk groups according to the predicted risk of liver fibrosis: minimal-risk group, (LiverRisk score values <6), low-risk group (LiverRisk score values from 6 to <9), medium-risk group (LiverRisk score values from 9 to <15), and high-risk group (LiverRisk score values ≥15). All models were compared to FIB4 and APRI scores, two methods used in clinical practice to assess liver fibrosis non-invasively (supplementary Table S2).[5] Confidence intervals were computed with bootstrapping with 2,000 random draws. To inspect the calibration of the predictive models, linear regression models between predicted and observed liver stiffness val-

ues were estimated with calibration intercept and slopes, and graphical representations were plotted.

Prognostic evaluation

For the prognostic evaluation of the models, we calculated the competing risks-adjusted (for non-liver-related events) cumulative incidence functions of liver-related outcomes (hospitalization, cancer incidence, and mortality) as a function of 4 different risk categories (minimal-risk, low-risk, medium-risk, and high-risk) according to the LiverRisk score. ICD10 codes used are shown in supplementary table S3. Cox regression models were also used to estimate the hazard ratios of different thresholds of the LiverRisk score. Several subgroup analyses were carried to assess the sensitivity of the scores with respect to different population characteristics. Analyses were carried for age groups, presence or absence of diabetes, obesity, alcohol consumption patterns, gender, and ethnicity. We also assessed the association between the continuous LiverRisk score and liver-related and non-liver-related 10-year mortality and hospitalizations with generalized additive models (GAM). All analyses were also performed to compare the performance of the LiverRisk score with that of FIB4 and APRI scores. All analyses were performed in R version 4.1.2.

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing, or the decision to submit the report. Ethical authorization was obtained to analyse all study cohorts. The UK Biobank study was approved by the North West Multi-Centre Research Ethics Committee and all participants provided written informed consent to participate in the UK Biobank study.

RESULTS

Derivation and validation of the LiverRisk score

We included a total of 14,726 subjects, 6,357 subjects in the derivation cohort, 4,370 in the first external validation cohort, and 3,999 in the second external validation cohort. The baseline characteristics of subjects included in the three cohorts are shown in Table 1.

In the derivation cohort, the four different models developed had very high accuracy in predicting liver stiffness either as continuous or categorical measurements using cutoff values of 6, 10, and 15 kPa (supplementary Figure S1 and supplementary Table S4). Findings were highly consistent in the two validation cohorts albeit accuracy was slightly lower compared to that of the derivation cohort (supplementary Table S5). Calibration results of the four models in the validation cohorts are shown in supplementary Figures S2 to S4. Out of the 4 models evaluated, the linear regression model (LM), from now on designated as ©LiverRisk score, was selected due to the better calibration and simpler model interpretation. Variables included in the ©LiverRisk score were age, gender, fasting glucose, cholesterol, AST, ALT, GGT, and platelet count. The accuracy of ©LiverRisk score in predicting liver stiffness was superior to that of standard non-invasive fibrosis tests, such as FIB-4 or APRI for the different cutoffs used (Table 2). The ©LiverRisk score can be calculated with an on-line calculator [https://liverriskscore.com**]**.

Association between LiverRisk score and liver-related mortality

A total of 416,200 subjects that met the inclusion criteria were included in the prognostic cohort (Table 1**)**. We calculated the ©LiverRisk score for each of the 416,200 subjects using their entry variables and analyzed its association with liver-related mortality, first liver-related hospitalization, and liver cancer during follow-up. During a median follow-up period

of 12 years, 28,627 of the 416,200 subjects died (6.9%), of whom 596 (2.1% of all deaths) died because of liver disease.

We estimated the competing risks-adjusted cumulative incidence of liver-related mortality for four groups (minimal-risk, low-risk, medium-risk, and high-risk group) according to selected cutoff values of ©LiverRisk score of 6, 10, and 15 as shown before (Figure 1, panel a). The proportion (and number) of subjects in these four groups was 86.4% (359,713 subjects), 12.7% (52,845), 0.8% (3,157), and 0.1% (485), respectively. There was a strong association between ©LiverRisk score groups and the probability of liver-related death, with subjects within the low, medium, and high-risk groups demonstrating a progressively higher probability of liver-related death at 12 years of follow-up compared to those in the minimal-risk group (figure 1a).

Figure 2 (panel a) shows the competing-risks adjusted hazard ratios of liver-related and non-liver-related mortality of all subjects divided into the risk groups. There was a progressive increase in hazard ratio of liver-related mortality according to risk groups, with subjects in the high-risk group having a hazard ratio (HR) of 437 (95%CI -347 to 641-) for liver-related mortality compared to subjects in the minimal risk group. The score was highly specific in predicting liver-related mortality, yet it was also associated with an increased hazard ratio of non-liver-related death, but the effect was lower compared to that of liver-related death (HR of 2.29 comparing high-risk and minimal-risk groups).

Ten-year liver-related mortality estimates increased markedly after the ©LiverRisk score reached a value of approximately 10, while non-liver-related mortality increased initially and then plateaued at around a ©LiverRisk score value of 20 (Figure 3). The significance of the ©LiverRisk score in predicting liver-related mortality persisted across different subpopula-

tions, such as age groups, alcohol consumption, diabetes mellitus, gender, ethnicity, obesity (supplementary figures S5 to S10, and supplementary table S6).

FIB-4 and APRI also predicted liver-related mortality in the cohort, but their accuracy was lower compared to that of the ©Liver Risk score (figure 4 and figures S11 and S12). ©Liver-Risk score also outperformed the fibrotic NASH index (FNI), a score that includes AST, HDL cholesterol, and HbA1c that has been reported to predict liver fibrosis in subjects with NAFLD (figure S13).

Association between liver risk score and first liver-related hospitalization and incident liver cancer

During a median follow-up of 12 years, 2,438 of the 416,200 subjects (0.59%) had at least one liver-related hospitalization. ©LiverRisk score groups were associated with progressively increased risk of liver-related hospitalization but not with risk of non-liver related hospitalization (Figure 1, panel b). The hazard ratios of liver-related hospitalization in the medium and high-risk groups were 47 (95% CI 42-53) and 126 (95% CI 102-154), respectively, compared to subjects in the minimal-risk group (Figure 2, panel b). The significance of the ©LiverRisk score in predicting first liver-related hospitalization persisted across different subpopulations categorized by age, alcohol consumption, diabetes mellitus, gender, ethnicity, and obesity (supplementary figures S14 to S19).

The incidence of liver cancer was also associated with ©LiverRisk score groups. Out of the whole cohort, 182 subjects (0.04%) developed hepatocellular carcinoma during a median follow-up of 8 years, with subjects in the high-risk group having a cumulative probability of 4.4% of developing liver cancer at 8 years of follow-up, while subjects in the two lower risk groups had a very small probability of incident liver cancer (minimal-risk group (0.009%), low risk 0.1%), and medium risk (1.0%) (Figure 1 panel c and figure 2 panel c). FIB-4 and

APRI also predicted liver-related hospitalization and incident liver cancer in the cohort, but their accuracy was lower compared to that of the ©Liver Risk score (figure 4, table S7, and figure S12).

DISCUSSION

The current study reports on a new score, the ©LiverRisk score, that predicts degree of liver stiffness and also future liver-related outcomes in an adult general population without known liver disease. The ©LiverRisk score is composed of eight variables that include age and gender as well as six laboratory variables (fasting glucose, cholesterol, AST, ALT, GGT, and platelet count), all of which are easily available in standard laboratory evaluations worldwide, and can be calculated with an on-line calculator. The ©LiverRisk score is reminiscent of scores widely used to assess risk profiles in chronic diseases, such as cardiovascular risk scores[27], and appears to be quite specific for liver-related outcomes.

The proposed ©LiverRisk score is effective in identifying subjects at risk for liver-related mortality and liver-related hospitalization as well as liver cancer and allows categorization of subjects of the population in four groups with markedly different risk of liver-related outcomes. As for liver-related mortality, only 0.04% of subjects in the minimal-risk group and 0.5% in the low-risk group died because of liver disease compared with 4.1% of the subjects in the medium-risk group, and 12.9% of those into the high-risk group. This corresponds to a hazard ratio of liver-related mortality of 471 in the high-risk group and 134 in the medium-risk group as compared to the minimal-risk group.

The accuracy of the ©LiverRisk score in predicting long-term liver-related outcomes was better than that of FIB-4 or APRI. This is probably related to the fact that these latter scores were derived from smaller cohorts of patients with chronic hepatitis C with high prevalence of liver fibrosis[5], whereas the ©LiverRisk score was derived from a larger, non-selected,

population-based cohort with low prevalence of liver fibrosis, reflective of the situation in the general population. The dependent variable used for development of ©LiverRisk score was liver stiffness assessed by transient elastography, a measurement that provides a good estimate of presence and severity of hepatic fibrosis.[5] The ©LiverRisk score was very accurate for diagnosis of increased liver stiffness in the derivation cohort as well as two independent validation cohorts, including a total of more than 14,000 subjects from the general population. Therefore, it is likely that the prognostic value of the ©LiverRisk score is related to its capacity to identify liver fibrosis early. Current evidence indicates that liver fibrosis is a strong predictor of liver-related complications and death, both in NAFLD and alcohol-associated liver disease.[28,29] Of note and at variance with other studies assessing the value of some scores in the prediction of future clinical events in cohorts of subjects with NAFLD,[30] our study was performed in population-based cohorts of adult subjects and therefore is not selective for any specific etiology of liver disease.

The ©LiverRisk score reported here is applicable for general use in clinical practice worldwide due to its simplicity, use of laboratory variables that are readily available, and relative low cost. The ©LiverRisk score may be used by general practitioners and nurses for opportunistic screening of liver fibrosis among subjects seen in primary care with metabolic risk factors for chronic liver disease or chronic alcohol consumption. This may allow subsequent correction of etiological factor(s), which may then prevent disease progression and improve prognosis. Besides, the ©LiverRisk score may be applied as a tool for population screening by automatically embedding the score into standard laboratory analyses performed for periodic controls in patients with chronic conditions, in hospitals or health centers, or in regular health check-ups. Hence, further studies are expected to explore the use of ©LiverRisk score in population screening. The score can also be used for risk prediction in individual subjects and maybe useful to empower individuals to change their lifestyle and behavior to decrease

the potential future risk of suffering from severe liver disease[7,8]. Finally, the ©LiverRisk sore can also be helpful to inform local policymakers and health authorities about liver diseases risks in the population for which they are responsible.

Limitations

In spite of the very large cohort size with long follow-up, the current study has some limitations. First, the prognostic value of the ©LiverRisk score was evaluated in a very large cohort but assessment was retrospective by calculating the value of the score for each subject at entry into the cohort and then assessing liver-related hospitalizations and liver-related death during follow-up through ICD-10 codes. Although the cohort meets relevant standards of quality with respect to data collection, the prognostic value of the ©LiverRisk score should ideally be tested with prospective collection of data. On the other hand, since the majority of subjects included in the different cohorts were of Caucasian origin, it remains to be established whether the current findings apply similarly to all ethnic groups.

In summary, we report the development and validation of a ©LiverRisk score that predicts future development of liver-related outcomes in the general population. The calculation of the ©LiverRisk score is based on simple demographic and laboratory parameters and can therefore be easily applicable to clinical practice in most countries of the world. The ©LiverRisk score may be useful for predicting risk in individual subjects and help them modify risk factors for liver disease as well as for screening for liver diseases at the population level. Future studies are needed to evaluate the impact of the use of this ©LiverRisk score and document cost effectiveness of screening, which may eventually help reduce the large burden of liver diseases in the world.

References

1 Sepanlou SG, Safiri S, Bisignano C, *et al.* The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2020; **5**: 245–66.

2 Karlsen TH, Sheron N, Zelber-Sagi S, *et al.* The EASL–Lancet Liver Commission: protecting the next generation of Europeans against liver disease complications and premature mortality. *Lancet* 2022; **399**: 61–116.

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

4 Ginès P, Krag A, Abraldes JG, Solà E, Fabrellas N, Kamath PS. Liver cirrhosis. *Lancet* 2021; **398**: 1359.

5 Berzigotti A, Tsochatzis E, Boursier J, *et al.* EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis – 2021 update. *J Hepatol* 2021; **75**: 659–89.

6 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* 2021; published online Dec 29. DOI:10.1016/J.CGH.2021.12.034.

7 Koutoukidis DA, Jebb SA, Tomlinson JB, Cobbold JF, Aveyard P. Association of weight changes with changes in histological features and blood markers in alcoholic steatohepatitis. *Clin Gastroenterol Hepatol* 2022; 20:e538.-e547.

8 Avanceña AL, Miller N, Uttal SE, Hutton DW, Mellinger JL. Cost-effectiveness of alcohol use treatments in patients with alcohol-related cirrhosis. *J Hepatol* 2021; **74**:1286-94.

9 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.

10 Wong VW-S, Chu WC-W, Wong GL-H, *et al.* Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. *Gut* 2012; **61**: 409–15.

11 Arslanow A, Baum C, Lammert F, Stokes CS. Nichtinvasive Früherkennung von Lebererkrankungen im Rahmen der betrieblichen Gesundheitsförderung. *Zentralbl Arbeitsmed Arbeitsschutz Ergon* 2017; **67**: 201–10.

12 Roulot D, Costes JL, Buyck JF, *et al.* Transient elastography as a screening tool for liver fibrosis and cirrhosis in a community-based population aged over 45 years. *Gut* 2011; **60**: 977–84.

13 Harman DJ, Ryder SD, James MW, *et al.* Direct targeting of risk factors significantly increases the detection of liver cirrhosis in primary care: a cross-sectional diagnostic study utilising transient elastography. DOI:10.1136/bmjopen-2014.

14 Caballería L, Pera G, Arteaga I, *et al.* High Prevalence of Liver Fibrosis Among European Adults With Unknown Liver Disease: A Population-Based Study. *Clin Gastroenterol Hepatol* 2018; **16**: 1138-1145.e5.

15 Fabrellas N, Hernández R, Graupera I, *et al.* Prevalence of hepatic steatosis as assessed by controlled attenuation parameter (CAP) in subjects with metabolic risk factors in primary care. A population-based study. *PLoS One* 2018; **13**: e0200656.

16 Hofman A, Breteler MMB, van Duijn CM, *et al.* The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007; **22**: 819–29.

17 Graupera I, Thiele M, Ma AT, et al. LiverScreen Project: study protocol for screening for liver fibrosis in the general population in European countries. *BMC Public Health* 2022 Jul 19;22(1):1385. doi: 10.1186/s12889-022-13724-6.

18 Sudlow C, Gallacher J, Allen N, *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**. DOI:10.1371/JOURNAL.PMED.1001779.

19 Chen X, Jeong JC. Enhanced recursive feature elimination. In: Sixth International Conference on Machine Learning and Applications (ICMLA 2007). IEEE, 2007: 429–35.

20 Su X, Yan X, Tsai C-L. Linear regression. *WIREs Computational Statistics* 2012; **4**: 275–94.

21 Yu K, Lu Z, Stander J. Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)* 2003; **52**: 331–50.

22 Culp M, Johnson K, Michailidis G. ada: An R Package for Stochastic Boosting. *J Stat Softw* 2007; **17**: 1–27.

23 Hastie T, Tibshirani R, Friedman J. Random Forests. 2009; : 587–604.

24 **Serra-Burriel M, Graupera I,** Torán P, et al. Transient elastography for screening fo liver fibrosis: Cost-effectiveness analysis from siix prospective cohorts in Europe and Asia. *J Hepatol* 2019; **71**: 1141–51.

25 de Franchis R, Bosch J, Garcia-Tsao G, *et al.* Baveno VII – Renewing consensus in portal hypertension. *J Hepatol* 2022; **76**: 959–74.

26 Bazerbachi F, Haffar S, Wang Z, et al. Range of normal liver stiffness and factors associated with increased stiffness measurements in apparently healthy individuals. *Clin Gastroenterol Hepatol* **2019**; 17: 54-64.

27 Goff DC, Lloyd-Jones DM, Bennett G, *et al.* 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation* 2014; **129**: 49–73.

28 Lackner C, Spindelboeck W, Haybaeck J, *et al.* Histological parameters and alcohol abstinence determine long-term prognosis in patients with alcoholic liver disease. *J Hepatol* 2017; **66**: 610–8.

29 Sanyal AJ, van Natta ML, Clark J, *et al.* Prospective Study of Outcomes in Adults with Nonalcoholic Fatty Liver Disease. *N Engl J Med* 2021; **385**: 1559–69.

30 Angulo P, Bugianesi E, Bjornsson ES, *et al.* Simple Noninvasive Systems Predict Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2013; **145**: 782-789.e4.

Figure list

Figure 1. Cumulative probability of a) liver and non-liver-related mortality; , b) liver and non-liver-related hospitalization, and c) liver and non-liver cancer in the 416,200 subjects from the prognostic cohort categorized into risk groups according to the ©LiverRisk score (minimal-risk: <6; low-risk: 6-10; medium-risk >10-15; and high-risk: >15). Shadowed areas represent 95% confidence intervals.

Figure 2. Hazard ratios (Cox proportional hazards) competing risks results of a) liver-related mortality and non-liver-related mortality; b) first liver-related hospitalization and first non-liver related hospitalization; and c) liver and non-liver cancer in the 416,200 subjects from the prognostic cohort categorized according to the ©LiverRisk score (minimal-risk: <6; low-risk: 6-10; medium-risk >10-15; high-risk: >15).

Figure 3. Ten-year mortality (top) and first hospitalization (bottom) estimates as a function of ©LiverRisk score; a) liver-related (red, continuous line) and non-liver-related (grey, discontinuous line).

Figure 4. Cumulative probability of liver-related mortality (panels a) and b)), liver-related hospitalization (panels c) and d)) and liver cancer (panels e) and f)) in the 416,200 subjects from the prognostic cohort categorized according to FIB-4 (panels a), c), and e)) and APRI scores values (panels b), d) and f)).

Appendix

LiverScreen Consortium investigators:

Marifé Alvarez, Peter Andersen, Paolo Angeli, Alba Ardèvol, Anita Arslanow, Luca Beggiato, Zahia Ben Abdesselam, Lucy Bennett, Bajiha Boutouria , Alessandra Brocca, M Teresa Broquetas, Llorenc Caballeria, Valeria Calvino, Judith Camacho, Aura Capdevila, Marta Carol, Laurent Castera, Marta Cervera, Fernando Cucchietti, Anna de Fuentes, Rob de Knegt, Sonke Detlefsen, Alba Diaz, José Diéguez Bande, Vanessa Esnault, Núria Fabrellas, Josep lluis Falco, Rosa Fernández, Celine Fournier, Matilde Fuentes, Peter Galle, Edgar García, Montserrat García-Retortillo, Esther Garrido, Pere Ginès, Rosa Gordillo Medina, Jordi Gratacós-Gines, Isabel Graupera, Ivica Grgurevic, Indra Neil Guha, Eva Guix, Rebecca Harris, Elena Hernández Boluda, Rosario Hernández-Ibañez, Jordi Hoyo, Arfan Ikram, Simone Incicco, Mads Israelsen, Marta Juan, Adria Juanola, Ralf Kaiser, Patrick S Kamath, Tom H Karlsen, Maria Kjærgaard, Harry J de Koning, Marko Korenjak, Aleksander Krag, Johanne Kragh Hansen, Marcin Krawczyk, Irina Lambert, Frank Lammert, Philippe Laboulaye, Simon Langkjær Sørensen, Cristina Laserna-Jiménez, Sonia Lazaro Pi, Elsa Ledain, Vincent Levy, Vanessa Londoño, Guirec Loyer, Anne Llorca, Ann T. Ma, Anita Madir, Michael Manns, Denise Marshall, M. Lluïsa Martí, Sara Martínez, Ricard Martínez Sala, Roser Masa Font, Jane Møller Jensen, Rosa M Morillas, Laura Muñoz, Ruth Nadal, Laura Napoleone, JM Navarrete, Phillip N Newsome, Vibeke Nielsen, Martina Pérez, Juan Manuel Pericas Pulido, Salvatore Piano, Judit Pich, Judit Presas Escobet, Elisa Pose, Katrine Prier Lindvig, Matthias Reichert, Carlota Riba, Dominique Roulot, Ana Belén Rubio, Maria Sánchez-Morata, Jörn Schattenberg, Feliu Serra-Burriel, Miquel Serra-Burriel, Louise Skovborg Just, Milan Sonneveld, Anna Soria, Christiane Stern, Patricia Such, Maja Thiele, Pere Toran, Antoni Torrejón, Marta Tonon, Emmanuel A Tsochatzis, Laurens van Kleef, Paulien van Wijngaarden, Vanessa Velázquez, Ana Viu, Susanne Nicole Weber, Tracey Wildsmith

Acknowledgements

We acknowledge the help of Beatriz Márquez in the preparation of the manuscript. This project has received funding from European Union's Horizon2020 programme under grant agreement No 847989. Also funded by Fondo de Investigación Sanitaria de Salud FIS PI18/01330 funded by Instituto de Salud Carlos III, Spanish Ministry of Economy, Industry, and Competitiveness and by the European Regional Development Fund (ERDF) (to NF).

Conflict of interest

Author contributions

MSB and PG conceived the idea for the study with input from FSB, AJ, MT, IG, EP, IGr, RJdK, LC, AK, FL, and PSK, and were responsible for the decision to submit the manuscript.

MSB designed the study, accessed and verfified the data, developed the database, undertook statistical analyses and interpretation, and drafted and revised the manuscript. FSB accessed and verified the data, developed the database, undertook statistical analyses, interpreted the data, and drafted and revised the manuscript. AJ, MT, IG, EP, GP, IG, LC, SP, LVK, MR, DR, JMP, JMS, ET, acquired data, interpreted the data, and drafted and revised the manuscript. ING, MGR, RH, JH, MF, CE, AM, PS, AM, SD, MT, AM, ATM, JP, EB, MJ, AS, MC, JGG, RMM, PT, JMN, AT, CF, AL, AA, HJdK, FC, MM, PNN, RH, AMA, PA, RJdK, THK, PG, VWWS, NF, LC, AK, FL, PSK interpreted data, contributed to manuscript drafting and revision. PG designed the study, accessed and verified the data, and drafted and revised the manuscript.

Data sharing statement Data from this manuscript can be requested by qualified researchers. Before the use of the data, proposals need to be approved by all partners of LiverScreen Consortium and a signed data sharing agreement will need to be executed. Approval will depend on the scientific value of the proposal, compatibility with the original patient consent, and data protection legislation.