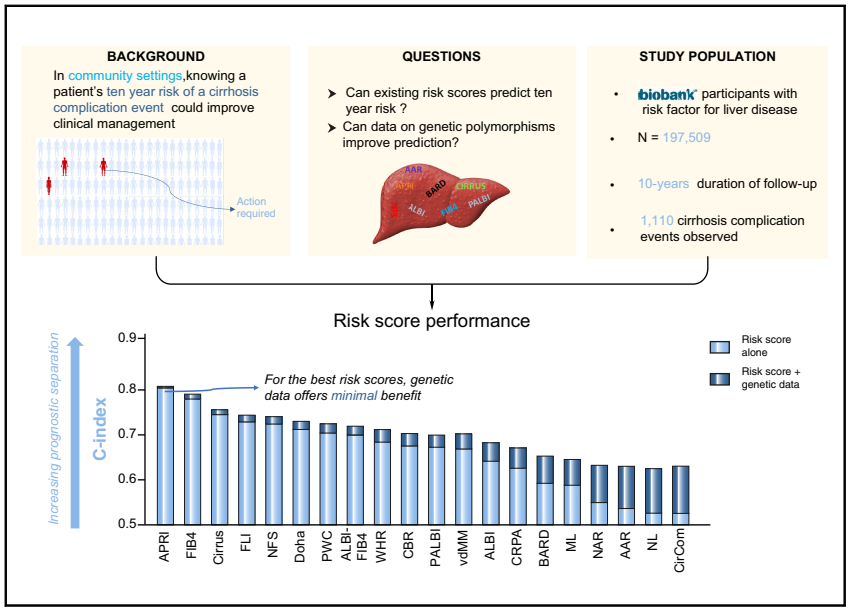
Research Article
Cirrhosis and Liver Failure

JOURNAL
OF HEPATOLOGY

Performance of routine risk scores for predicting cirrhosis-related morbidity in the community

Graphical abstract



Highlights

- Individualised 10-year risk of cirrhosis-related morbidity can be predicted in the community.
- The APRI score exhibited the best discriminative ability (C-index >0.80).
- 10-year cumulative incidence was 14.8% for individuals with APRI in the 99th percentile.
- Genetic risk scores were outperformed by more accessible alternatives.
- Genetic risk scores add little new prognostic information beyond what is already captured by APRI and FIB-4.

Authors

Hamish Innes, Joanne R. Morling, Stephan Buch, Victoria Hamill, Felix Stickel, Indra Neil Guha

Correspondence

Hamish.Innes@gcu.ac.uk (H. Innes).

Lay summary

New approaches are needed in community settings to reduce the late diagnosis of chronic liver disease. Thus, in a community cohort, we assessed the ability of 20 routine risk scores to predict 10-year risk of cirrhosis-related complications. We show that 2 routine risk scores in particular – "APRI" and "FIB-4" – could be repurposed to estimate an individual's 10-year risk of cirrhosis-related morbidity. Adding genetic risk factor information to these scores only modestly improved performance.

https://doi.org/10.1016/j.jhep.2022.02.022
 J. Hepatol. 2022, 77, 365–376

Research Article
Cirrhosis and Liver Failure

JOURNAL OF HEPATOLOGY

Performance of routine risk scores for predicting cirrhosis-related morbidity in the community

Hamish Innes[1,2,3,*], Joanne R. Morling[3,4,5], Stephan Buch[6], Victoria Hamill[1,2], Felix Stickel[7], Indra Neil Guha[4,5]

[1]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow, UK; [2]Public Health Scotland, Glasgow, UK; [3]Lifespan and Population Health, University of Nottingham, Nottingham, UK; [4]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, NG7 2UH, UK; [5]Nottingham Digestive Diseases Centre, School of Medicine, University of Nottingham, NG7 2UH, UK; [6]Medical Department 1, University Hospital Dresden, TU Dresden, Germany; [7]Department of Gastroenterology and Hepatology, University Hospital of Zurich, Switzerland

See Editorial, pages 293–295

Background & Aims: Models predicting an individual's 10-year risk of cirrhosis complications have not been developed for a community setting. Our objectives were to assess the performance of existing risk scores – both with and without genetic data – for predicting cirrhosis complications in the community.
Methods: We used a 2-stage study design. In stage 1, a systematic review was conducted to identify risk scores derived from routine liver blood tests that have demonstrated prior ability to predict cirrhosis-related complication events. Risk scores identified from stage 1 were tested in a UK Biobank subgroup, comprising participants with a risk factor for chronic liver disease (stage 2). Cirrhosis complications were defined as hospitalisation for liver cirrhosis or presentation with hepatocellular carcinoma. Discrimination of risk scores with and without genetic data was assessed using the Wolbers C-index, Harrell's adequacy index, and cumulative incidence curves.
Results: Twenty risk scores were identified from the stage-1 systematic review. For stage-2, 197,509 UK biobank participants were selected. The cumulative incidence of cirrhosis complications at 10 years was 0.58%; 95% CI 0.54-0.61 (1,110 events). The top performing risk scores were aspartate aminotransferase-to-platelet ratio index (APRI: C-index 0.804; 95% CI 0.788-0.820) and fibrosis-4 index (FIB-4: C-index 0.780; 95% CI 0.764-0.795). The 10-year cumulative incidences of cirrhosis complications for participants with an APRI score exceeding the 90th, 95th and 99th percentile were 3.30%, 5.42% and 14.83%, respectively. Inclusion of established genetic risk loci associated with cirrhosis added <5% of new prognostic information to the APRI score and improved the C-index only minimally (*i.e.* from 0.804 to 0.809).
Conclusions: Accessible risk scores derived from routine blood tests (particularly APRI and FIB-4) can be repurposed to estimate 10-year risk of cirrhosis morbidity in the community. Genetic data improves performance only minimally.

Lay summary: New approaches are needed in community settings to reduce the late diagnosis of chronic liver disease. Thus, in a community cohort, we assessed the ability of 20 routine risk scores to predict 10-year risk of cirrhosis-related complications. We show that 2 routine risk scores in particular – "APRI" and "FIB-4" – could be repurposed to estimate an individual's 10-year risk of cirrhosis-related morbidity. Adding genetic risk factor information to these scores only modestly improved performance.

Introduction

Chronic liver disease (CLD) is a global public health problem, causing more than a million deaths every year.[1] Most deaths are secondary to complications of cirrhosis, such as decompensated cirrhosis and primary liver cancer. One of the most notable hallmarks of CLD is that it is usually not diagnosed until patients present with complications from cirrhosis; at which point, treatment options are limited and prognosis is irreversibly poor.[2] In this scenario, earlier diagnosis of CLD could allow clinicians to act sooner to prevent liver related mortality.[3,4] However, this must be weighed against the potential for unnecessary intervention (*e.g.* referral to specialist care and further diagnostic follow-up) in the vast majority of patients with CLD who never go on to develop complications of cirrhosis. Thus, improving outcomes hinges on being able to manage patients in a way that is proportionate to their risk of severe disease. This requires effective risk stratification tools.

Models that predict individual risk of disease complications are becoming more common in primary care. Exemplars include QRISK3, which estimates a patient's risk of having a heart attack or stroke within the next 10 years.[5] Predicting complication risk is particularly useful where the number of individuals who go on to develop complication events is a small fraction of the total number of individuals with disease (as with CLD). Although there is no CLD equivalent to QRISK3 at present, we hypothesised that existing risk scores based on routine liver blood tests may be useful for predicting cirrhosis complication events in a

Keywords: Non-invasive tests; NITs; liver fibrosis; NAFLD; alcohol liver disease; prognosis; genetics; single nucleotide polymorphisms.
Received 2 November 2021; received in revised form 26 January 2022; accepted 14 February 2022; available online 7 March 2022
* Corresponding author. Address: Glasgow Caledonian University; Cowcaddens Road; G40BA; Glasgow, UK.
E-mail address: Hamish.Innes@gcu.ac.uk (H. Innes).
https://doi.org/10.1016/j.jhep.2022.02.022

community setting. Examples include the fibrosis-4 index (FIB-4), and the non-alcoholic fatty liver disease (NAFLD) fibrosis score, which show prognostic utility among patients attending specialist liver clinics.[6,7] However, the prognostic performance of such risk scores in a community setting – *i.e.* where the burden of CLD is at its greatest, but the incidence rate of severe disease is far lower – is not well characterised.

Another key area of uncertainty relates to the added value of genetic data for CLD risk stratification. Knowledge of the germline genetic risk factors influencing onset of cirrhosis has advanced appreciably in recent years.[8] Several genetic risk scores have now been developed, which show strong association with cirrhosis-related outcomes.[9,10] However, the incremental utility of genetic data, over and above cheaper and more readily accessible risk scores, remains unclear. In other words, we do not know whether genetic data provides any additional prognostic information beyond what is already captured by routine scores.

The primary objectives of this study were therefore to: i) quantify performance of existing fibrosis scores for identifying individuals at risk of cirrhosis complications; and ii) assess the degree to which the performance of these scores can be enhanced by adding data on genetic risk.

Materials and methods

Stage 1 (identifying candidate risk scores)

A systematic review was performed to identify risk scores with evidence of prognostic ability for cirrhosis complications. The search was designed to be sensitive rather than specific (*i.e.* to capture all potentially relevant risk scores accepting that some will be false positives). For eligible studies, we extracted information on: study participants (*i.e.* location, number of centres, setting); outcome definition; sample size; and model performance. Further details can be found in Appendix A.

Stage 2 (evaluating risk score performance)

Risk scores identified in stage 1 were then evaluated in the United Kingdom Biobank (UKB) cohort. UKB is a community cohort study of more than half a million individuals in the UK (N = 502,492). Participants were interviewed in May 2006 to July 2010 from 22 UKB assessment centres located throughout the UK. All individuals aged 40-69 years and living within 25 miles of an assessment centre (approximately 9 million persons in total) were sent an invitation letter for the study. During the interview, participants completed a comprehensive health questionnaire, a physical examination and donated biological specimens. Follow-up data on subsequent health outcome events are supplied through record linkage to UK mortality, hospital admission and cancer registries.[11] UKB has approval from the UK North West Multicentre Research Ethics committee. Informed consent was obtained from each participant.

Study population

Previous studies suggest focusing on patients with established risk factors for CLD optimises risk stratification in the community.[12–15] Thus, our study population was restricted to UKB participants with one or more of the following CLD risk factors at UKB interview (n = 238,585).

1) Alcohol consumption exceeding 14 units/week (the recommended threshold for safe drinking in UK[16]).
2) Abdominal obesity (waist-hip ratio >1.0 if male, and >0.9 if female).
3) General obesity (BMI >30).
4) Diagnosis of type 2 diabetes mellitus (T2DM).

A detailed description of the covariates used to define these risk factors is provided in Appendix B. We then excluded participants for any of the following:

a. Developed the primary outcome event prior to UKB interview (see following paragraph for definition);
b. Missing data for ≥1 risk score identified in our systematic review;
c. Missing genetic data for one or more of the 20 single nucleotide polymorphisms (see later paragraph for further information).

Primary outcome event

The primary outcome event was presentation with an incident (*i.e.* first time) cirrhosis complication event. We defined this as either a hospital admission for cirrhosis, death from cirrhosis, or presentation with hepatocellular carcinoma (HCC). Only events occurring during the first 10 years of follow-up were considered. Hospital admissions due to cirrhosis were identified using a validated set of ICD & Operations/Procedure (OPCS4) codes.[17] HCC was defined as the presence of an ICD10:C22.0 code within either a cancer registration or hospital admission record. See Table S1 for further details.

Calculating risk score values

Risk score values were largely derived from blood specimens collected at the participant's baseline interview. The specific UKB field IDs used to calculate risk score values are provided in Table 1. Laboratory methods adopted by UKB to generate biomarker data from blood specimens have been described previously.[18,19]

Statistical analyses

Definition of the 'at risk' period

All statistical analyses were underpinned by survival analysis methods. Participants were followed up from the date of their UKB interview through to the earliest of either: a) the date of cirrhosis complications (if at all); b) the date of death (if at all); c) the study completion date, or d) the 10-year follow-up date.

The study completion date relates to the date of hospital registry completion; 30-Jun-2020, 31-Oct-2016, and 1-Mar-2016, for participants in England, Scotland and Wales, respectively. The 10-year follow-up date was defined as 10 years after the date of UKB interview.

Competing risk perspective

The cumulative incidence of cirrhosis complications depends not only on the risk of cirrhosis complications itself, but also on the risk of competing events. Previous studies show that ignoring competing risk events leads to biased estimates of cumulative incidence.[20] Thus, in this study, non-cirrhosis-related mortality occurring before cirrhosis complications was treated as a competing risk event where this was possible. The definition of non-cirrhosis mortality is provided in Table S1.

Risk score discrimination

Discrimination is arguably the most fundamental aspect of risk score performance. In general, it is defined as the degree to

Table 1. Risk scores identified in stage 1 systematic review and the prognostic factors underpinning each score.

	Risk score																			
Prognostic factor (UKB field ID)	**AAR**	**ALBI**	**ALBI-FIB-4**	**APRI**	**BARD**	**CBR**	**CRPA**	**CirCom**	**Cirrus**	**DOHA**	**FIB-4**	**FLI**	**ML**	**NAR**	**NFS**	**NL**	**PALBI**	**PWC**	**vdMM**	**WHR**	**Total**
Platelet count (30080)			X	X					X	X	X				X		X	X	X		9
Aspartate aminotransferase (30650)	X		X	X	X					X	X				X				X		8
Albumin (30600)		X	X				X		X	X				X	X		X				8
Alanine aminotransferase (30620)	X		X		X						X				X				X		6
Bilirubin (30840)		X	X			X			X								X				5
Age (21022)			X								X				X				X		4
BMI (21001)					X							X			X						3
Waist circumference (48)												X								X	2
Type 2 diabetes mellitus (various fields*)					X										X						2
Lymphocyte count (30120)													X			X					2
Neutrophil count (30140)														X		X					2
Gamma glutamyl transferase (30730)												X									1
Creatinine (30700)									X												1
Mean corpuscular volume (30040)									X												1
Sodium (30530)									X												1
Total protein (30860)									X												1
Triglycerides (23407)												X									1
C-reactive protein (30710)							X														1
Prior hospital admission data (various fields*)								X													1
Cystatin (30720)						X															1
Monocyte count (30130)													X								1
Leukocyte count (30000)																		X			1
Hip circumference (49)																				X	1
Sex (31)																			X		1

Prognostic factors are listed in descending order of frequency. Fields used to infer type 2 diabetes mellitus are described in appendix B. Hospital admission data refers to in-patient hospital admission spells occurring prior to UKB enrolment.

AAR, aspartate aminotransferase-to-alanine aminotransferase ratio; ALBI, albumin-bilirubin score; ALBI-FIB-4, albumin-bilirubin fibrosis-4 index; APRI, aspartate aminotransferase to platelet ratio; BARD, BMI-AST ratio-diabetes model; CirCom, cirrhosis-specific comordbitiy score; Cirrus, cirrhosis using standard tests; CRPA, C-reactive protein-to-albumin ratio; CBR, cystatin-to-bilirubin ratio; DOHA, Doha score; FIB-4, fibrosis-4 index; FLI, fatty liver index; ML, monocyte-to-lymphocyte ratio; NAR, neutrophil-to-albumin ratio; NFS, non-alcohol fatty liver disease fibrosis score; NL, neutrophil-to-lymphocyte ratio; PALBI, platelet-albumin-bilirubin score; PWC, platelet-to-white cell count ratio; UKB, UK biobank; vdMM, van der Meer mortality; WHR, waist-hip ratio.

JOURNAL OF HEPATOLOGY

Table 2. Baseline characteristics and 10-year follow-up data for study population.

		Cirrhosis complication event (primary outcome)		Non-cirrhosis mortality (competing risk event)	
Characteristics (baseline)	**Participants, n (%)**	**Events, n (%)**	**10-year cumulative incidence, % (95% CI)**	**Events, n (%)**	**10-year cumulative incidence, % (95% CI)**
Age group, years					
<50	41,943 (21.2)	131 (11.8)	0.32 (0.27-0.38)	637 (6.3)	1.56 (1.44-1.68)
50-59	67,011 (33.9)	351 (31.6)	0.54 (0.49-0.60)	2,300 (22.7)	3.53 (3.39-3.68)
≥60	88,555 (44.8)	628 (56.6)	0.73 (0.67-0.78)	7,218 (71.1)	8.37 (8.19-8.56)
Sex					
Female	88,677 (44.9)	316 (28.5)	0.37 (0.33-0.41)	3,360 (33.1)	3.90 (3.78-4.03)
Male	108,832 (55.1)	794 (71.5)	0.75 (0.70-0.80)	6,795 (66.9)	6.41 (6.27-6.56)
Ethnicity					
White	170,242 (86.2)	942 (84.9)	0.63 (0.54-0.73)	1,277 (12.6)	4.80 (4.54-5.06)
Non-white	27,267 (13.8)	168 (15.1)	0.57 (0.53-0.61)	8,878 (87.4)	5.37 (5.26-5.48)
Townsend deprivation quintile					
Q1 (least deprived)	40,204 (20.4)	150 (13.5)	0.39 (0.33-0.45)	1,727 (17.0)	4.44 (4.24-4.65)
Q2	40,022 (20.3)	159 (14.3)	0.41 (0.35-0.48)	1,842 (18.1)	4.73 (4.52-4.94)
Q3	39,885 (20.2)	202 (18.2)	0.52 (0.45-0.59)	1,879 (18.5)	4.83 (4.62-5.05)
Q4	39,228 (19.9)	252 (22.7)	0.66 (0.58-0.74)	2,031 (20.0)	5.32 (5.10-5.55)
Q5 (most deprived)	37,923 (19.2)	347 (31.3)	0.94 (0.85-1.04)	2,663 (26.2)	7.21 (6.95-7.48)
Missing	247 (0.1)	0 (0.0)	\	13 (0.1)	\
Alcohol intake, units/week					
<15	76,400 (38.7)	450 (40.5)	0.60 (0.55-0.66)	4,433 (43.7)	5.97 (5.80-6.15)
15-49	108,887 (55.1)	423 (38.1)	0.40 (0.36-0.44)	4,844 (47.7)	4.57 (4.45-4.70)
50+	10,486 (5.3)	205 (18.5)	2.00 (1.74-2.29)	742 (7.3)	7.26 (6.77-7.77)
Missing	1,736 (0.9)	32 (2.9)	1.89 (1.32-2.62)	136 (1.3)	8.02 (6.79-9.37)
Type 2 diabetes					
No	182,445 (92.4)	863 (77.8)	0.49 (0.46-0.52)	8,692 (85.6)	4.90 (4.80-5.00)
Yes	15,064 (7.6)	247 (22.3)	1.68 (1.48-1.89)	1,463 (14.4)	9.99 (9.51-10.48)
BMI category, kg/m^2					
<30	109,049 (55.2)	474 (42.7)	0.45 (0.41-0.49)	5,336 (52.6)	5.02 (4.89-5.15)
≥30	88,460 (44.8)	636 (57.3)	0.74 (0.69-0.80)	4,819 (47.5)	5.62 (5.47-5.78)
All participants	197,509 (100.0)	1,110 (100.0)	0.58 (0.54-0.61)	10,155 (100.0)	5.29 (5.19-5.39)

which individuals who develop the outcome of interest have a higher risk score value *vs.* those who do not. In this study, we assessed discrimination of each risk score through a number of approaches.

First, we quantified the overall discriminative ability of each risk score using the C-index. In general, the C-index measures the proportion of all possible "participant-pairs" that are "concordant". A "participant pair" refers to a random selection of 2 individuals from the dataset, and this pair is said to be "concordant" if the individual with the higher risk score develops the outcome event of interest sooner than the individual with the lower risk score. C-index values usually range from 0.5 to 1.0, where a value of 0.5 indicates zero discrimination (*i.e.* no better than chance) and a value of 1.0 indicates perfect discrimination.[21] In this study, all risk scores were handled as continuous variables. We used the Wolbers-modified version of the C-index, which takes competing risk events into account.[22] In addition to the 10-year prediction horizon, we also calculated C-index values over a shorter time horizon of 5 years.

Second, we assessed if the C-index of each prediction model varied according to selected patient characteristics. The following characteristics were considered: sex, age group (40-49; 50-59 and 60+ years), deprivation, alcohol intake, obesity and type 2 diabetes. For deprivation, participants were grouped into quintiles based on the Townsend deprivation index. The Cochran's Q and the I^2 statistics were calculated to assess heterogeneity across strata.[23]

Third, we created "high" and "low" risk groups by dichotomising each risk score at a range of illustrative cut-off points. Five cut-off points were considered: a) 50th percentile; b) 80th percentile c) 90th percentile; d) 95th percentile; and e) 99th percentile. The 10-year cumulative incidence of cirrhosis complications was calculated in each high/low-risk group created. In general, the more discriminating a risk score, the more separation one would expect to see between the cumulative incidence in "high" and "low" risk categories. All cumulative incidence estimates were generated non-parametrically using the "stcompet" package in Stata version 17.

Improving discrimination by integrating genetic data

Next, we quantified the degree to which the discrimination of existing risk scores is improved by adding data on genetic risk for cirrhosis-related complications. Relevant genetic loci were identified by reviewing previous genetic association studies for alcohol-related liver disease, NAFLD or mixed aetiology cohorts (see Appendix C). Individualised data for selected loci were obtained from Version 3 of the UKB genetic dataset.[24] Genetic data were combined with each existing risk score using Cox regression modelling. Forty Cox models were fitted in total: 2 for each of the 20 existing risk scores (*i.e.* 1 model that included all genetic loci as independent variables, and 1 model that omitted them). All predictor variables were included as continuous variables to avoid information loss. Crucially, one cannot necessarily assume that the relationship between a given risk score and risk of cirrhosis complications is linear. Thus, we used Royston's multivariate fractional polynomial procedure[25] to identify the optimal functional relationship between each existing risk score and the outcome. Through this procedure, we were able to account for non-linear relationships – thus further reducing the potential for information loss when combining with genetic data. However, we constrained all such non-linear relationships to be monotonic (*i.e.* a first-order fractional polynomial[25]) to ensure

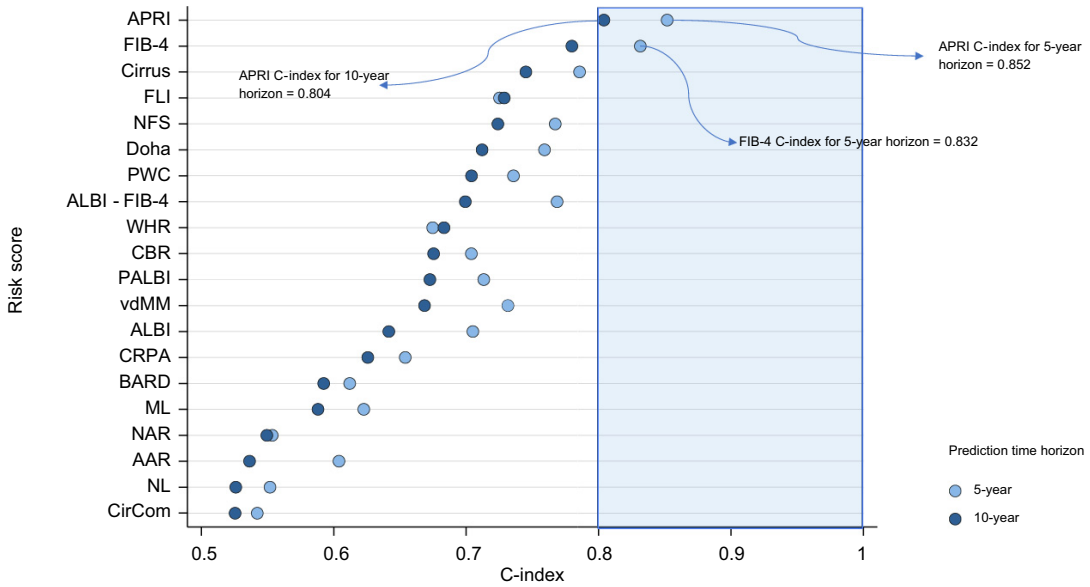


Fig. 1. Summary C-index estimate for risk scores over a 5- and 10-year prediction horizon. Ordered from top to bottom in descending C-index for 10-year horizon. All estimates are based on the Wolbers modification of the Harrell's C-index. Acronym expansions for risk scores can be found in the main text. The blue box denotes C-index values exceeding 0.80, which EASL guidelines on non-invasive tests suggest is the minimum level of discrimination needed for a model to be clinically useful (albeit the empirical rationale for this threshold is not clear). 95% CIs for all C-index estimates are available in Table S8. AAR, aspartate aminotransferase-to-alanine aminotransferase ratio; ALBI, albumin-bilirubin score; ALBI-FIB-4, albumin-bilirubin fibrosis-4 index; APRI, aspartate aminotransferase to platelet ratio; BARD, BMI-AST ratio-diabetes model; CirCom, cirrhosis-specific comordbitiy score; Cirrus, cirrhosis using standard tests; CRPA, C-reactive protein-to-albumin ratio; CBR, cystatin-to-bilirubin ratio; DOHA, Doha score; FIB-4, fibrosis-4 index; FLI, fatty liver index; ML, monocyte-to-lymphocyte ratio; NAR, neutrophil-to-albumin ratio; NFS, non-alcohol fatty liver disease fibrosis score; NL, neutrophil-to-lymphocyte ratio; PALBI, platelet-albumin-bilirubin score; PWC, platelet-to-white cell count ratio; vdMM, van der Meer mortality; WHR, waist-hip ratio.

Table 3. Genetic loci included in the risk score augmentation analysis.

SNP	Chr:Basepair position	Minor allele	Ref allele	Missing proportion	MAF	Nearest gene	Position	Phenotype
rs12904	1:155106697	A	G	0.000	0.411	*EFNA1*	UTR3	Fibrosis/cirrhosis
rs2642438	1:220970028	A	G	0.000	0.291	*MARC1.*	Exonic	Fibrosis/cirrhosis
rs708118	1:228201801	C	T	0.013	0.389	*WNT3A*	Intronic	HCC
rs5743836	3:52260782	G	A	0.000	0.162	*TLR9*(dist = 603)	Upstream	Hepatic encephalopathy
rs72613567	4:88231392	TA	T	0.000	0.270	*HSD17B13*	Intronic	Fibrosis/cirrhosis
rs2562582	5:36605360	C	T	0.050	0.179	*SLC1A3*(dist = 1097)	Intergenic	Hepatic encephalopathy
rs888655	5:72917439	A	G	0.003	0.273	*ARHGEF28*(dist = 4544)	Intergenic	Fibrosis/cirrhosis
rs11134977	5:175904141	C	T	0.014	0.448	*FAF2*	Intronic	Fibrosis/cirrhosis
rs9398804	6:126703390	A	T	0.039	0.445	*CENPW*	Intronic	Fibrosis/cirrhosis
rs7029757	9:132566666	A	G	0.010	0.090	*TOR1B*	Intronic	Fibrosis/cirrhosis
rs2792751	10:113940329	T	C	0.000	0.269	*GPAM*	Exonic	Fibrosis/cirrhosis
rs1799992	11:118957246	C	T	0.014	0.399	*HMBS*	Intronic	Fibrosis/cirrhosis
rs28929474	14:94844947	T	C	0.000	0.019	*SERPINA1*	Exonic	Fibrosis/cirrhosis
rs58542926	19:19379549	T	C	0.000	0.074	*TM6SF2*	Exonic	HCC; fibrosis/cirrhosis
rs187429064	19:19380513	G	A	0.009	0.011	*TM6SF2*	Exonic	Fibrosis/cirrhosis
rs15052	19:41813375	C	T	0.005	0.170	*HNRNPUL1*	UTR3	Fibrosis/cirrhosis
rs429358	19:45411941	C	T	0.000	0.154	*APOE*	Exonic	HCC; fibrosis/cirrhosis
rs313853	19:47287939	C	T	0.024	0.339	*SLC1A5*	UTR5	Hepatic encephalopathy
rs601338	19:49206674	G	A	0.000	0.499	*FUT2*	Exonic	Hepatic encephalopathy
rs641738	19:54676763	T	C	0.009	0.437	*TMC4*	Exonic	Fibrosis/cirrhosis
rs1883711	20:39179822	C	G	0.009	0.028	*MAFB*(dist = 134666)	Intergenic	Fibrosis/cirrhosis
rs738409	22:44324727	G	C	0.000	0.216	*PNPLA3*	Exonic	HCC; fibrosis/cirrhosis

All variants listed above are in linkage disequilibrium. Rs11134977 is used in place of rs374702773, which is not available in the UKB genetic dataset. Both MAF and the missing proportion relate to the full genetic dataset, comprising of 487,409 participants. UTR3 = 3 prime untranslated region; UTR5 = 5 prime untranslated region. For more information see Appendix C.
MAF, minor allele frequency; SNP, single nucleotide polymorphism.

clinical plausibility. All new risk scores integrating genetic data were derived from the linear predictor of the corresponding regression model. N.B. because of insufficient computational resources, it was not possible to perform Royston's multivariate fractional polynomial procedure within a Fine-Gray regression modelling context accounting for competing risks.

When assessing the added value of the genetic data, our general approach was to compare the performance of the risk model including genetic data to that of the equivalent model omitting genetic data. To build a composite picture, we did this in 3 ways. First, we calculated the Wolbers C-index for each risk score with genetic data and compared this to the Wolblers C-index for the original score omitting genetic data.

Second, we used Harrell's Adequacy Index to assess the percentage of new prognostic information provided by the genetic data.[26] The adequacy index is defined as: (LR_{RS}/$LR_{RS+genetic}$); where $LR_{RS+genetic}$ is the likelihood ratio statistic for the Risk score + genetic model (in Cox model), and LR_{RS} is the likelihood ratio statistic for the original risk score (in Cox model). One minus the adequacy index indicates the fraction of new prognostic information provided by the genetic data.[26]

Third, we created "high" and "low" risk groups by dichotomising each genetic risk score at the 50th; 80th; 90th; 95th and 99th percentiles. We calculated the cumulative incidence in each group and compared this informally to the equivalent estimate for the score omitting genetic data. As discussed earlier, a more discriminating score should translate into greater separation of the cumulative incidence between high- and low-risk groups.

Two sensitivity analyses were performed to assess if the added value of genetic data was altered when: a) restricting the study population to participants of White British ancestry; and b) genetic data was represented using a specific genetic risk score previously developed by Bianco *et al.*[10] White British ancestry was defined by the UKB core team using genetic and self-reported data (see UKB field ID:22006).

Relationship between risk score value and 10-year complication risk

The relationship between risk score value and 10-year cumulative incidence was assessed by calculating the 10-year cumulative incidence within 12 granular risk score groups. These groups were risk score deciles 1-9; 90-94th percentile; 95-98th percentile, and ≥99th percentile (*i.e.* the top 1%). Decile 1 refers to people whose risk score value is in the 0-9th percentile; decile 2 relates to a risk score value in the 10-19th percentile, and so on. The cumulative incidence was calculated for both cirrhosis-related complication events and the competing risk event (non-cirrhosis mortality). This analysis was only performed for the 3 risk scores with the greatest discriminative ability.

Results

Stage 1 (identification of candidate risk scores):

One thousand and eighty-three studies were retrieved from our systematic search, of which 32 met our inclusion/exclusion criteria. Most eligible studies included patients attending secondary care centres (27/32) (see Table S2). Twenty risk scores were identified in total, as enumerated in Table 1. The

formulae used to calculate all risk score values are outlined in Appendix D.

Stage 2 (derivation and characteristics of study population): Our study population was comprised of 197,509 individuals (Fig. S1). The mean age was 57.4 years with slightly more females (55%) than males (45%). In total, 88,460 participants (45%) had a BMI exceeding 30 kg/m^2 and 10,486 (5.3%) reported current alcohol intake exceeding 50 units/week (Table 1). Detailed descriptive statistics for each risk score, including values at specific percentile cut-off points, are shown in Table S3. Minor allele frequencies for the genetic loci of interest ranged from 1.8% to 29.3% (Table S4).

Primary outcome event

One thousand, one hundred and ten (0.56%) individuals presented with a cirrhosis complication event within 10 years of UKB interview. The most common type of event was a hospital admission for cirrhosis (982/1,110; 88.5%), followed by presentation with HCC (101/1,110) (Table S5-S6).

The median duration of follow-up was 10.00 years per participant (mean: 9.59 years). The 10-year cumulative incidence of cirrhosis complications was 0.58% (95% CI 0.54-0.61). 10,115 participants experienced the competing risk event, equating to a 10-year cumulative incidence of 5.29% (95% CI 5.19-5.39) (Table 2).

Risk score discrimination

All risk scores were associated with the outcome event, but discriminative ability varied widely. The best performing risk scores over a 10-year time horizon were: aspartate aminotransferase-to-platelet ratio index (APRI: C-index: 0.804; 95% CI 0.788-0.819); FIB-4 (C-index: 0.780; 95% CI 0.765-0.795); and cirrhosis using standard tests (Cirrus: 0.745; 95% CI 0.728-0.762). Discrimination was even greater over a 5-year time horizon (Fig. 1; Table S7). The 10-year cumulative incidence of cirrhosis complications for participants with an APRI score exceeding the 50th, 80th, 90th, 95th and 99th percentile was 0.96%, 1.98%, 3.30%, 5.42% and 14.83%, respectively (Table S8).

Discrimination did not vary significantly by age, sex, ethnicity, obesity, alcohol intake or deprivation quintile (Table S9). However, discrimination was higher for individuals with T2DM. For example, the C-index for FIB-4 was 0.820 (95% CI 0.794-0.847) for participants with T2DM, *vs.* 0.764 (95% CI 0.747-0.782) for those without (Cochran Q p <0.001; I^2 = 83.13%). A similar pattern was apparent for APRI and Cirrus (Table S9).

Improving discrimination with genetic data

Our genetic augmentation analysis incorporated 20 independent genetic loci, enumerated in Table 3. The C-index of the genetic risk factor model alone was 0.618 (0.601-0.635) and the regression parameters for this model are shown in Table S10. The fraction of new prognostic information provided by genetic data was greatest

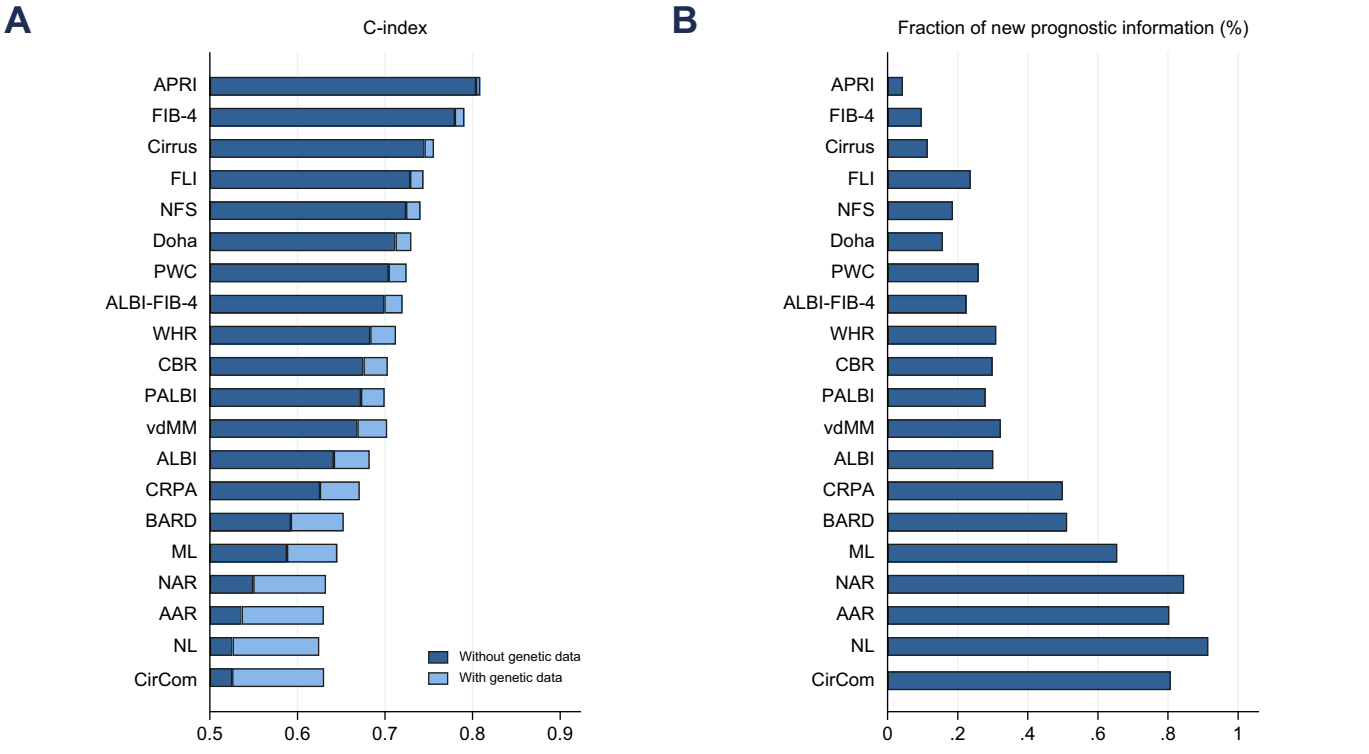


Fig. 2. Risk score improvement through addition of genetic risk data. Improvement is assessed in 2 ways. (A) Indicates the Wolbers C-index without genetic data *vs.* C-index with genetic data for each score. (B) Represents the fraction of new prognostic information provided by genetic data, calculated from Harrell's adequacy index. N.B. risk scores are arranged from top to bottom in order of descending C-index. AAR, aspartate aminotransferase-to-alanine aminotransferase ratio; ALBI, albumin-bilirubin score; ALBI-FIB-4, albumin-bilirubin fibrosis-4 index; APRI, aspartate aminotransferase to platelet ratio; BARD, BMI-AST ratio-diabetes model; CirCom, cirrhosis-specific comordbitiy score; Cirrus, cirrhosis using standard tests; CRPA, C-reactive protein-to-albumin ratio; CBR, cystatin-to-bilirubin ratio; DOHA, Doha score; FIB-4, fibrosis-4 index; FLI, fatty liver index; ML, monocyte-to-lymphocyte ratio; NAR, neutrophil-to-albumin ratio; NFS, non-alcohol fatty liver disease fibrosis score; NL, neutrophil-to-lymphocyte ratio; PALBI, platelet-albumin-bilirubin score; PWC, platelet-to-white cell count ratio; vdMM, van der Meer mortality; WHR, waist-hip ratio.

for scores with lower discriminative ability (*i.e.* cirrhosis-specific comordbitiy score [CirCom], neutrophil-to-lymphocyte ratio, and aspartate aminotransferase-to-alanine aminotransferase ratio) and weakest for scores with higher discriminative ability (*i.e.* APRI; FIB-4; Cirrus). For the CirCom score for example, the addition of genetic data added >70% of new prognostic information and the C-index was improved from 0.526 to 0.630. In contrast, genetic data added <5% of new prognostic information to the APRI score, and the C-index improvement was marginal (*i.e.* from 0.804 to 0.809) (Fig. 2; Table S11).

Equally, the separation in cumulative incidence between high/low-risk groups was only marginally improved when

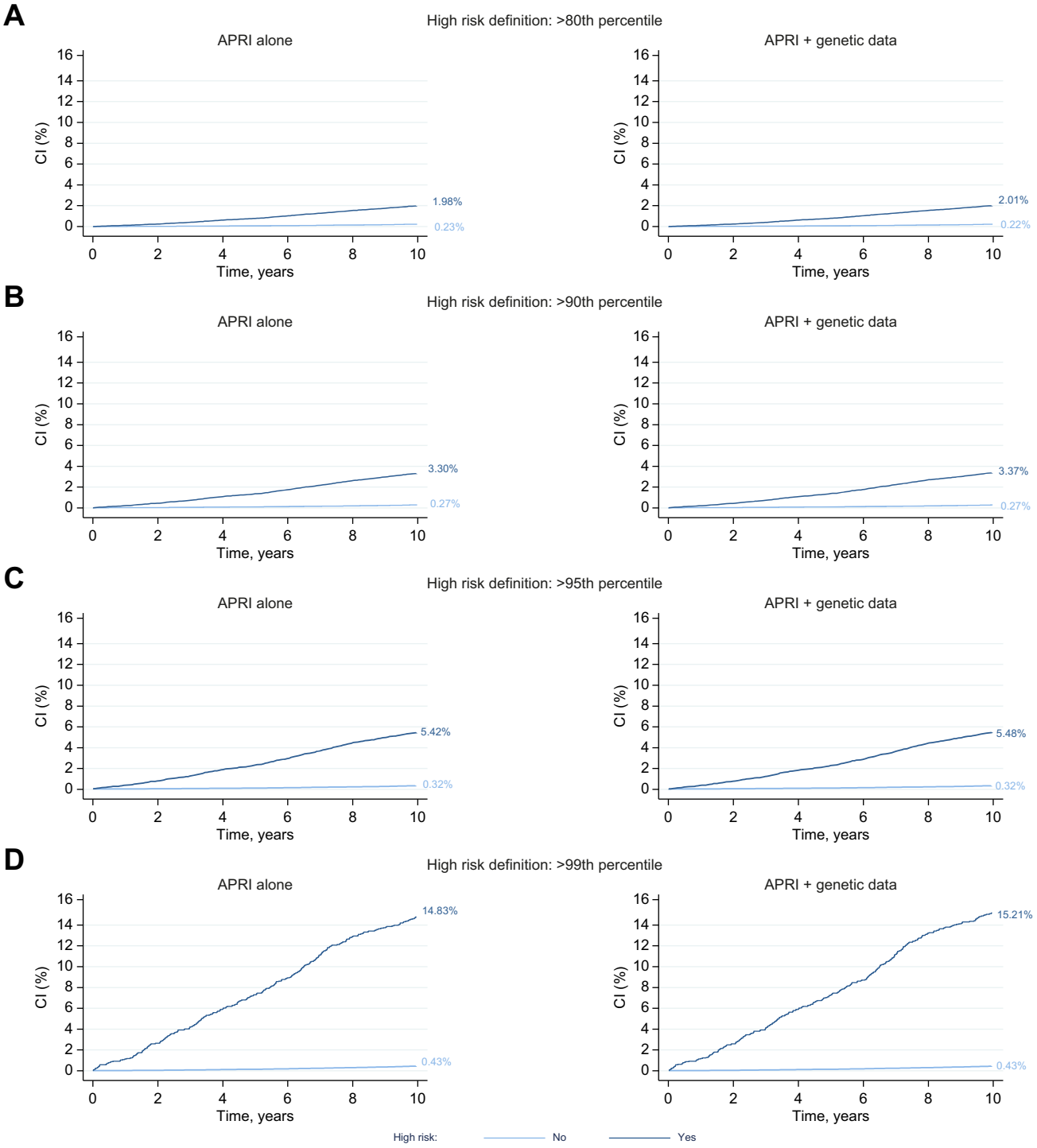


Fig. 3. Comparison of 10-year CI for high/low-risk participants, defined by original APRI *vs.* APRI + genetic data. If genetic data improves discrimination then this should translate into greater separation in cumulative incidence curves for the high- and low-risk groups. High/low-risk was defined according to illustrative percentile cut-off points. As an example, the 80th percentile definition means that individuals whose score was in the 80th percentile or greater (*i.e.* in the top 20%) were categorised as high risk, and the remainder were categorised as low risk. APRI, aspartate aminotransferase-to-platelet ratio index; CI, cumulative incidence.

genetic data was added to APRI and FIB-4. For example, based on the 99th percentile definition of high/low-risk, 10-year cumulative incidence was 14.83% (high risk) *vs.* 0.43% (low risk) for APRI alone. The equivalent separation for the APRI-genetic model was 15.21% (high risk) *vs.* 0.43% (low risk) (Fig. 3). Similarly, at the 99th percentile definition, the 10-year cumulative incidence was 12.36% (high risk) *vs.* 0.46% (low risk) for FIB-4 alone. The equivalent separation for the FIB-4-genetic model was 12.82% (high risk) and 0.45% (low risk) (Fig. S2).

The added value of genetic data was unchanged in all sensitivity analyses (Table S11).

The relationship between risk score value and 10-year complication risk

The cumulative incidence of cirrhosis complications did not increase linearly with risk score value. Rather an exponential relationship was observed (see Fig. 4; Table S12). Also, as risk score value increased, the cumulative incidence of a cirrhosis complication event either converged with or surpassed the cumulative incidence of the competing risk event.

Discussion

New precision medicine strategies are needed in community settings to reduce late diagnosis of CLD.[2–4] Hitherto studies have mainly focused on developing diagnostic algorithms for significant/severe liver fibrosis. Here, we explore the potential for bringing a "QRISK3" paradigm to CLD, where a patient's individualised risk of a complication event is directly estimated. Overall, our findings indicate that pre-existing risk scores derived from routine liver blood tests can be repurposed for estimating 10-year risk in the community. The APRI and FIB-4 scores showed the greatest potential for this repurposing. For example, the 10-year risk of a complication event for individuals whose APRI score surpassed the 99th percentile (14.83%) was ~25 times higher than the 10-year risk in the total study population (0.58%). C-index values for APRI and FIB-4 were either close to or exceeded 0.80, which EASL guidelines indicate is the minimum level of discrimination needed for a model to be clinically useful[27] (albeit, the empirical basis for this 0.80 threshold is not clear). The Cirrus score[28] also performed well and may hold a practical advantage over APRI and FIB-4 because it does not rely on aspartate aminotransferase measurement (which is not available in all settings). We also show that the discriminative ability was consistent across age group, sex, deprivation, BMI, and ethnicity categories. However, discrimination did vary according to T2DM status, which has also been observed elsewhere.[7]

The EASL-Lancet liver commission highlights the importance of shifting the diagnostic emphasis for CLD towards those at greatest risk of developing severe disease.[4] However, the current paradigm – *i.e.* focusing exclusively on a diagnostic rather than a prognostic perspective[29,30] – is ill-equipped to achieve this goal. First, fibrosis is a surrogate measure that may not fully capture the risk of developing severe morbidity (which is often what patients, clinicians and health systems are most interested in). Second, recommended thresholds for identifying fibrosis/cirrhosis in the community (*e.g.* FIB-4 >3.25 or enhanced liver fibrosis test >10.5[29,30]) are questionable insofar as they are extrapolated from secondary care where the spectrum of disease differs considerably from that in the community. Adapting these thresholds to a community context is an intractable challenge because liver biopsy (the gold standard for measuring fibrosis) is not ethically justifiable in a community cohort. Third, the inclination to avoid "missing" disease often requires the choice of low thresholds, thus necessitating specialist fibrosis tests on a large proportion of the population (many of whom will not ultimately develop cirrhosis morbidity). This places major constraints on resources and service capacity. For all these reasons, we believe that directly estimating individualised risk of a complication event will prove a useful adjunct to existing clinical management protocols/guidelines.[29,30] Crucially, we show it is feasible to predict 10-year risk using routinely collected data variables, *i.e.* without the need to adopt costly, proprietary or invasive biomarkers. Indeed, we suspect that APRI, FIB-4 and Cirrus scores could be automatically calculated for most individuals in primary care using prior test results stored digitally. A more nuanced/granular risk management system where "clinical outcomes" are front and centre would allow those at highest risk to be directed to specialist services in a timely manner. In contrast, low- and medium-risk patients could be managed with a proportionate use of specialist fibrosis tests (*i.e.* imaging or serum tests). Importantly, this would allow thresholds for clinical action (*i.e.* referral to specialist care) to be set by local

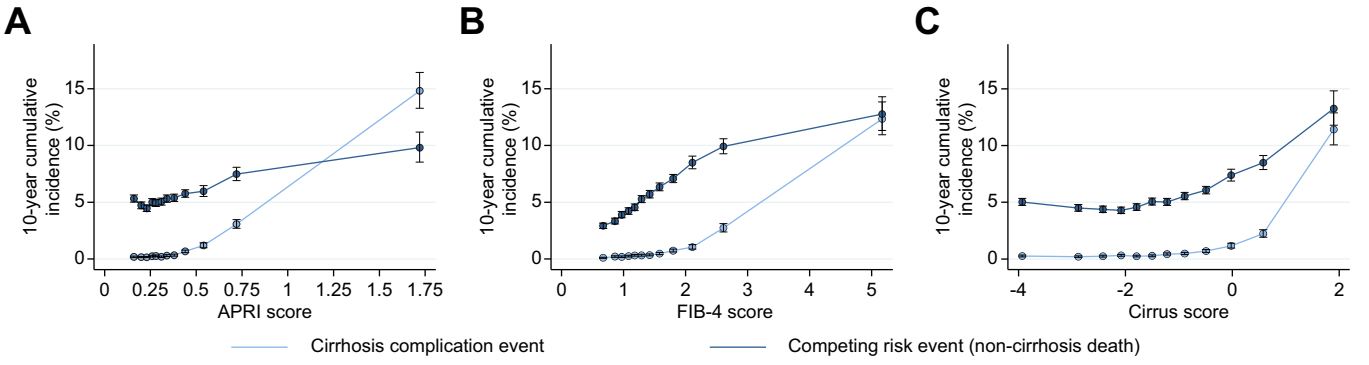


Fig. 4. Relationship between risk score value and 10-year cumulative incidence. The first 9 points moving from left to right represent the cumulative incidence within risk score deciles 1-9. Thus, the point on the furthest left represents the cumulative incidence for individuals whose risk score is in decile 1, the adjacent point represents decile 2 etc. However, the last 3 points moving from left to right represent the 90-94th, 95-98th and ≥99th percentile, respectively. Error bars represent 95% CIs. APRI, aspartate aminotransferase-to-platelet ratio index; Cirrus, cirrhosis using standard tests; FIB-4, fibrosis-4 index.

healthcare systems themselves, taking resourcing levels into account. Overall, therefore, our results may have important implications for existing clinical practice.

Our findings also lend an important perspective regarding the utility of genetic risk scores for predicting liver-related prognosis. Despite considerable interest of late in developing genetic scores for risk stratification,[9,10] the added value of genetic data relative to cheaper and more accessible biomarkers has not been adequately explored. Our results show that for the best performing risk scores (*e.g.* APRI, FIB-4 and Cirrus), the incremental benefit of genetic data was marginal. In the case of APRI for example, adding genetic data provided <5% new prognostic information and would only increase the C-index from 0.804 to 0.809. These modest gains are corroborated by our cumulative incidence curves for FIB-4 and APRI, which show comparable levels of prognostic separation with and without genetic data. Nevertheless, this conclusion is at odds with a recent study by De Vincentis *et al.*, who reported that genetic data *does* provides additional prognostic insight relative to existing scores.[31] Because this study was also underpinned by UKB data, our opposing interpretations must reflect methodological differences (*i.e.* "analytical flexibility"[32]). Two differences in our view are particularly salient. First, in the present analysis, all risk scores were handled as continuous variables and were combined using Royston's multivariate fractional polynomial procedure.[25] This was done to minimise information loss and is consistent with best practice guidelines.[21] Conversely, risk scores appear to have been categorised in De Vincentis *et al.*'s study, which, whilst analytically simpler, invariably leads to appreciable information loss and underestimation of a risk score's discriminative ability.[21] Second, although both studies were based on UKB data, there are sizeable differences in the exact participants included. In our study, we included all participants with a risk factor for CLD in order to mirror the target population in which these risk scores are likely to be applied in real world clinical practice. In contrast, De Vincentis included a substantial number of individuals (~60% of the cohort) with no risk factors for CLD. In combination, these methodological differences could account for our opposing conclusions. More generally, this highlights the non-trivial impact that analytical flexibility can have on scientific conclusions, as demonstrated recently by Botvinik-Nezer *et al.*[32]

This study has a number of limitations and caveats that warrant discussion. First, we were unable to validate risk scores in terms of their calibration performance. This was because the risk scores we assessed were not developed to predict future risk of cirrhosis complications. Thus, there is no equation we can validate relating specific risk score values to a 10-year predicted risk. Future studies are needed therefore to calibrate FIB-4 and APRI with respect to these prognostic outcomes. Ideally, this should be taken forward with large primary care datasets such as the QResearch cohort[5] and the UK clinical practice research data link. Nevertheless, we hope our own descriptive analysis, outlining the exponential relationship between APRI/FIB-4 score and 10-year cumulative incidence, will inform future efforts to calibrate these scores. Another limitation is that our study population was defined only in terms of risk factors for alcohol-related liver disease and NAFLD, and did not explicitly capture other forms of CLD such as viral hepatitis. However, NAFLD and alcohol account for the vast majority of cirrhosis cases in the UK and are more likely to be diagnosed late compared to other aetiologies.[2] These reasons justify our focus on these risk groups. Third, 26% of participants were excluded from our study population due to missing genotype or biomarker data. We did not use multiple imputation to impute missing values because this procedure is challenging to combine both with the multivariate fractional polynomial procedure and a competing risk perspective – both of which are critical aspects of this study. The main reasons for missing biomarker data (*e.g.* albumin or bilirubin) included the absence of a biological sample, an aliquot problem, or the reported value being outside the normal range at the time of measurement.[33] For the genetic data, some single nucleotide polymorphisms were set to missing if quality control checks were not passed. In addition, poor quality samples were excluded altogether using metrics such as extreme heterozygosity.[24] On the whole, data are likely to be missing completely at random and thus should not cause significant bias. Fourth, the UKB population is not representative of either a primary care population or the general UK population. UKB participants are more likely to be female, older in age and live in less socioeconomically deprived areas than non-participants.[34] Mortality rates and cancer incidence rates are lower in UKB than the general population. Fifth, the goal of our systematic review was intentionally limited in scope. We did not attempt to assess model performance through meta-analysis of the studies identified in our review. Instead, the systematic review was designed only to identify *potentially* relevant risk scores, which were then compared head-to-head in the powerful UKB dataset. Another caveat to note is that genetic data may hold practical advantages over routine liver risk scores, which may outweigh performance differences. One example is that genetic risk scores reflect lifetime risk (by virtue of being derived from immutable germline DNA); this may be an advantageous property in certain contexts. It should also be pointed out that the incremental benefit of genetic data is likely to increase with time, as new genetic risk factors for cirrhosis morbidity are uncovered and genetic risk scores refined. For example, existing genetic risk scores may be improved by incorporating rare pathogenic variants in candidate genes. A recent study by Pelusi *et al.* provides some support for this approach.[35] Finally, our findings regarding the added value of genetic data may not be generalisable to other clinical contexts – *e.g.* monitoring patients with compensated cirrhosis in secondary care. This question warrants further research.

In summary, we have performed a comprehensive comparative analysis of readily available risk scores for predicting cirrhosis complications in a community setting. Our results highlight the latent potential to estimate 10-year complication risk using inexpensive routine liver blood tests that are available in most non-specialist clinical settings. Moreover, for the top performing scores (*i.e.* APRI and FIB-4), we show that data on genetic risk factors for cirrhosis provide little additional prognostic information.

Abbreviations

APRI, aspartate aminotransferase-to-platelet ratio index; CirCom, cirrhosis-specific comordbitiy score; Cirrus, cirrhosis using standard tests; CLD, chronic liver disease; FIB-4, fibrosis-4 index; HCC, hepatocellular carcinoma; NAFLD, non-alcoholic fatty liver disease; T2DM, type 2 diabetes mellitus; UKB, UK biobank.

Financial support

HI is supported by a viral hepatitis fellowship from the Medical Research Foundation (grant ID: C0825). JRM receives salary

support from the UK Medical Research Council Clinician Scientist Award (MR/P008348/1). This work was also supported by grants from the Swiss National Funds (SNF no. 310030_169196) and the Swiss Foundation for Alcohol Research to FS.

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

a) Study concept: all authors; b) study design: all authors; c) acquisition of data: HI; d) resources: HI, JRM, ING; e) Statistical analysis: HI; f) drafting manuscript: HI, JRM, ING; g) critical revision of manuscript: all authors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data availability statement

The data used in this study are not publicly available, but can be acquired through successful application to the UK Biobank. For further information, see: https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access

Acknowledgements

We would like to thank the three anonymous Journal of Hepatology reviewers whose constructive criticism greatly improved the quality of this manuscript.

This research has been conducted using the UK Biobank resource: application number: 8764.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2022.02.022.

References

Author names in bold designate shared co-first authorship

[1] GBD 2017 Cirrhosis collaborators. 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;5:245–266.

[2] **Innes H**, **Morling JR**, Aspinall EA, Goldberg DJ, Hutchinson SJ, Guha IN. Late diagnosis of chronic liver disease in a community cohort (UK biobank): determinants and impact on subsequent survival. Public Helath 2020;187:165–171.

[3] Williams R, Aspinall R, Bellis M, Camps-Walsh G, Cramp M, Dhawan A, et al. Addressing liver disease in the UK: a blueprint for attaining excellence in health care and reducing premature mortality from lifestyle issues of excess consumption of alcohol, obesity and viral hepatitis. Lancet 2014;384:1953–1997.

[4] Karlsen TH, Sheron N, Zelber-Sagi S, Carrieri P, Dusheiko G, Bugianesi E, et al. The EASL-Lancet Liver Commission: protecting he next generation of Europeans against liver disease complications and premature mortality. Lancet 2022;399:61–116.

[5] Hippisley-Cox J, Couplan C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ 2017;357:J2099.

[6] Lee J, Vali Y, Boursier J, Spijker R, Anstee QM, Bossuyt P, et al. Prognostic accuracy of FIB-4 NAFLD fibrosis score and APRI for NAFLD-related events: a systematic review. Liver Int 2021;41:261–270.

[7] Younes R, Caviglia GP, Govaere O, Rosso C, Armandi A, Sanavia T, et al. Long-term outcomes and predictive ability of non-invasive scoring systems in patients with non-alcoholic fatty liver disease. J Hepatol 2021;75:786–794.

[8] Stickel F, Moreno C, Hampe J, Morgan MY. The genetics of alcohol dependence and alcohol-related liver disease. J Hepatol 2017;66:195–211.

[9] Whitfield JB, Schwantes-An TH, Darlay R, Aithal GP, Atkinson SR, Bataller R, et al. A genetic risk score and diabetes predicts development of alcohol-related cirrhosis in drinkers. J Hepatol 2022;76:275–282.

[10] Bianco C, Jamialahmadi O, Pelusi S, Baselli G, Dongiovanni P, Zanoni I, et al. Non-invasive stratification of hepatocellular carcinoma in non-alcoholic fatty liver using polygenic risk scores. J Hepatol 2021;74:775–782.

[11] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex disease of middle and old age. PLoS Med 2015;12:E1001779.

[12] Harman DJ, Ryder SD, James MW, Jelpke M, Ottey DS, Wilkes EA, et al. Direct targeting of risk factors significantly increases the detection of liver cirrhosis in primary care: a cross-sectional diagnostic study utilizing transient elastography. BMJ Open 2015;5:e007516.

[13] El-Gohary M, Moore M, Roderick P, Watkins E, Dash J, Reinson T, et al. Local care and treatment of liver disease (LOCATE) – a cluster-randomized feasibility study to discover, access and manage early liver disease in primary care. PLoS One 2018;13:e0208798.

[14] Caballeria L, Pera G, Arteaga I, Rodriguez L, Aluma A, Morillas RM, et al. High prevalence of liver fibrosis among Europeans adults with unknown liver disease: a population-based study. Clin Gastroenterol Hepatol 2018;16:1138–1145.

[15] Llop E, Iruzubieta P, Perello C, Carrillo CF, Cabezas J, Escudero MD, et al. High liver stiffness values by transient elastography related to metabolic syndrome and harmful alcohol use in a large Spanish cohort. United Eur Gastroenterol J 2021;9:892–903.

[16] Department of Health. UK Chief Medical Officers' alcohol guidelines review. Summary of the proposed new guidelines. January 2016. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/489795/summary.pdf. Accessed June 2021.

[17] Ratib S, Fleming KM, Crooks CJ, Walker AJ, West J. Causes of death in people with liver cirrhosis in England compared with the general population: a population-based cohort study. Am J Gastroenterol 2015;110:1149–1158.

[18] Fry D, Almond R, Moffat S, Gordon M, Singh P. UK Biobank project. Companion document to accompany serum biomarker data. Version 1.0. Date 11/3/2019. https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum_biochemistry.pdf. Accessed January 2022.

[19] Sheard SM, Nicholls R, Froggatt J. UK Biobank Haematology Data Companion Document. Date: 24/10/17. https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/haematology.pdf. Accessed January 2022.

[20] Jepsen P, Vilstrup V, Andersen PK. The clinical course of cirrhosis: the importance of multistate models and competing risks analysis. Hepatology 2015;62. 2922-302.

[21] Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–73.

[22] Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic models with competing risks. Methods and application to coronary risk prediction. Epidemiology 2009;20:555–561.

[23] Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Under reliance on I2 in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;8:79.

[24] Bycroft C, Freeman C, Petcova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature 2018;562:203–209.

[25] Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999;28:964–974.

[26] Harrell Jr FE. Regression modelling strategies. 2001. Chapter 9. ISBN 0172-7397.

[27] European Association for the 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;75:659–689.

[28] Hydes T, Moore M, Stuart B, Kim M, Su F, Newell C, 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.

[29] Newsome PN, Cramb R, Davison SM, Dillon JF, Foulerton M, Godfrey EM, et al. Guidelines on the management of abnormal liver blood tests. Gut 2018;67:6–19.

[30] National Institute for Health and Care Excellence (NICE). NICE Guideline. Non-alcoholic fatty liver disease (NAFLD): assessment and management. Available at: https://www.nice.org.uk/guidance/ng49. Accessed October 2021..
[31] De Vincentis A, Tavaglione F, Jamialahmadi O, Picardi A, Incalzi RA, Valenti L, et al. A polygenic risk score to refine risk stratification and prediction for severe liver disease by clinical fibrosis scores. Clin Gastroenterol Hepatol 2021. In press.
[32] Botvinik-Nexer R, Holzmeister F, Camerer CF, Dreber A, Huber J, Johannesson M, et al. Variability in the analysis of a single neuroimaging dataset by many teams. Nature 2020;582:84–88.
[33] UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications). Version 1.2. Available at: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/biomarker_issues.pdf. Accessed: December 2021.
[34] Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of sociodemographic and health-related characteristics of UK biobank participants with those of the general population. Am J Epidemiol 2017;186:1026–1034.
[35] Pelusi S, Baselli G, Pietrelli A, Dongiovanni P, Donati B, McCain MV, et al. Rare pathogenic variants predispose to hepatocellular carcinoma in non-alcoholic fatty liver disease. Sci Rep 2019;9:3682.