Performance of routine risk scores (with and without genetic data) for predicting cirrhosis morbidity in the community

AUTHORS: Hamish Innes; Joanne Morling; Stephan Buch; Victoria Hamill; Felix Stickel; Indra Neil Guha

Table of contents

Figure S1. Derivation of final sample

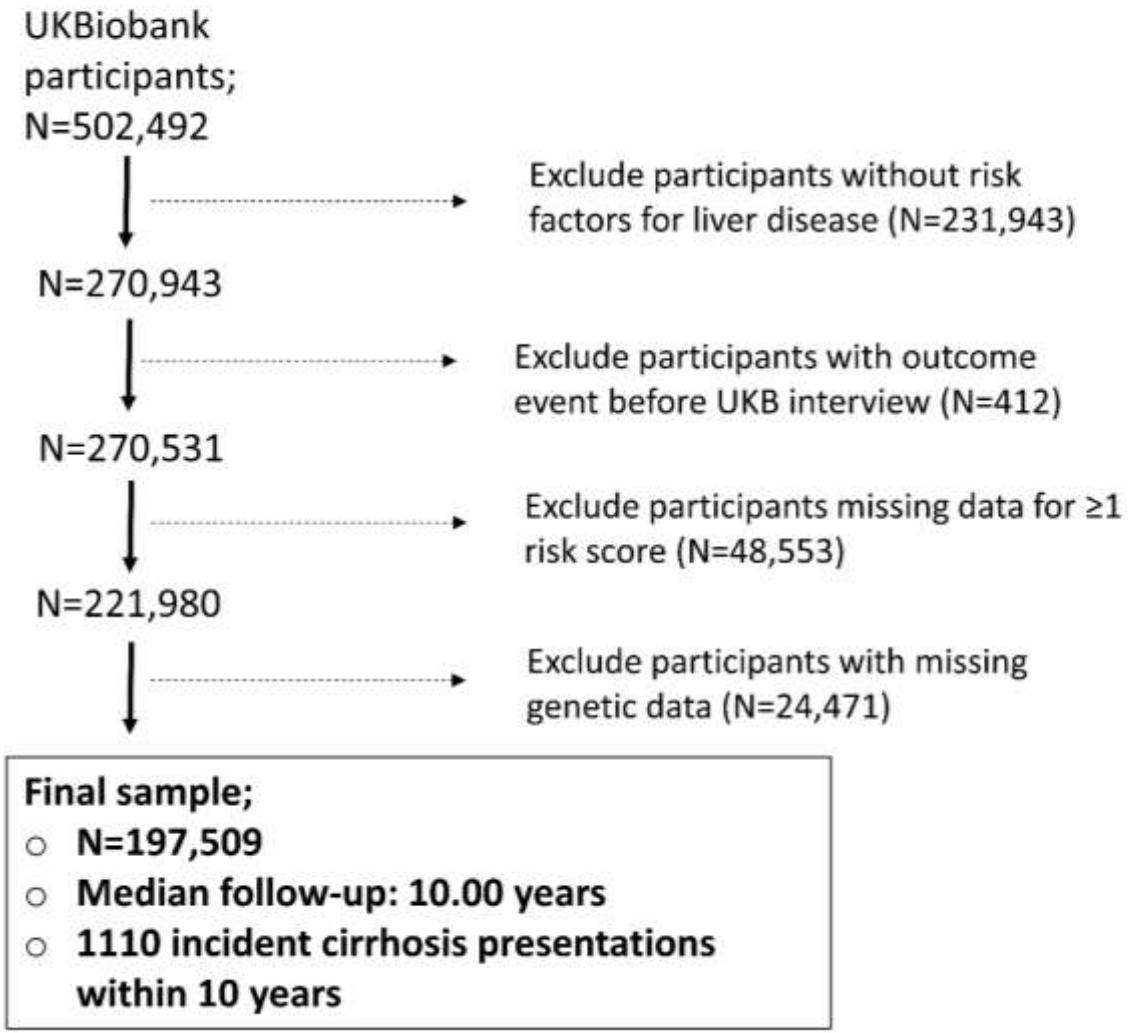
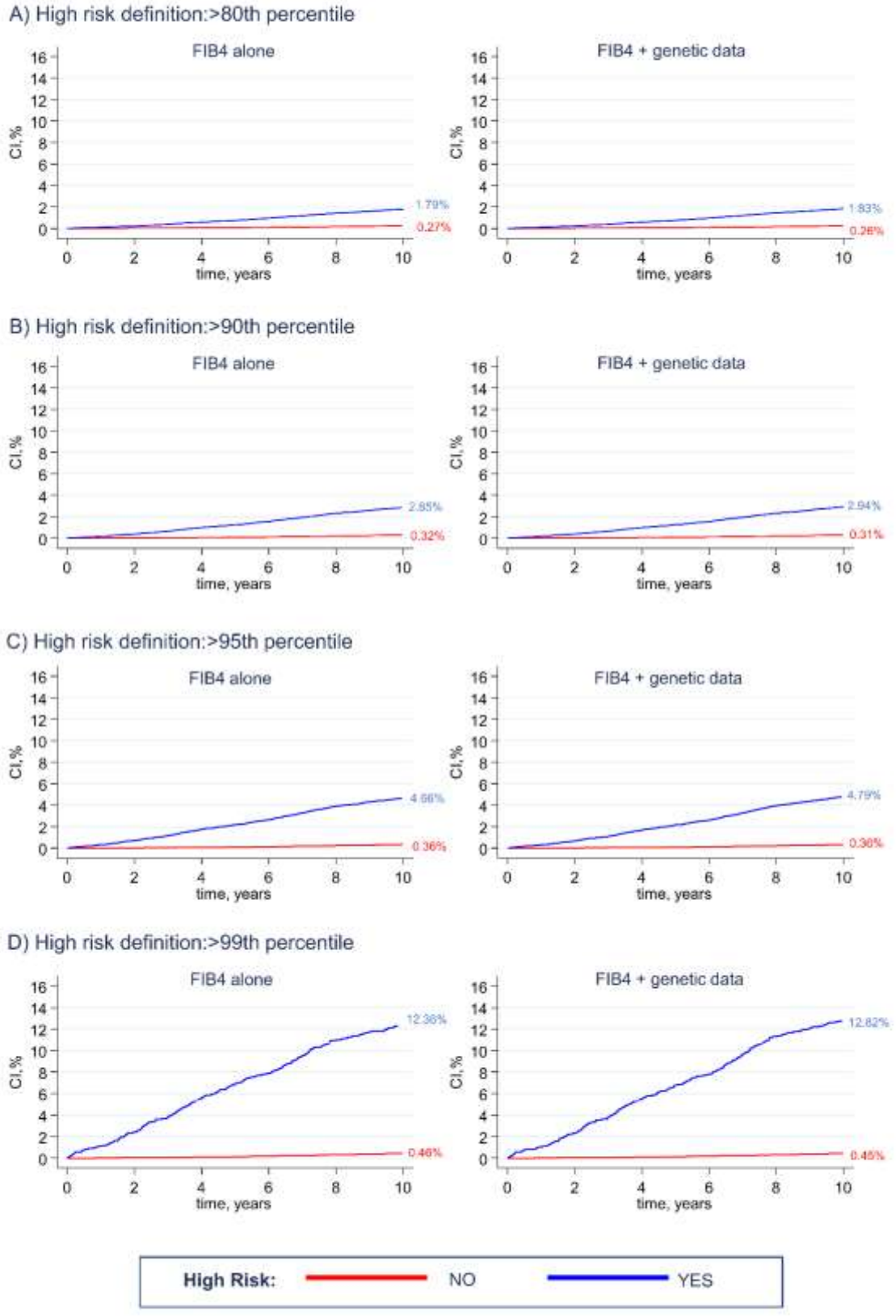


Figure S2: Comparison of ten year cumulative incidence (CI) for high/low risk participants, defined by original FIB4 versus FIB4 + genetic data



If genetic data improves discrimination then this should translate into greater separation in cumulative incidence curves between 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.

Table S1. ICD 10 and OPCS4 codes used to define cirrhosis-related complication events

Type of complication event	Health registry	Code type	Code	Code description
Cirrhosis-related hospital admission	Hospital admission	ICD 10	K70.3	Alcoholic cirrhosis of liver
			K71.7	Toxic liver disease with fibrosis and cirrhosis of liver
			K72.1	Chronic hepatic failure
			K74.4	Secondary biliary cirrhosis
			K74.5	Biliary cirrhosis, unspecified
			K74.6	Other and unspecified cirrhosis of liver
			K76.6	Portal hypertension
			I85.0; I859; I98.2; I98.3	Esophageal varices
			I86.4	Gastric varices
		OPCS4	J06.1	Tranjugular intrahepatic insertion of stent into portal vein
			J06.2	Transjugular intrahepatic insertion of stent graft into portal vein
			G10.4	Local ligation of varices of oesophagus
			G10.8	Other specified open operations on varices of oesophagus
			G10.9	Unspecified open operations on varices of oesophagus
			G14.4	Fibreoptic endoscopic injection sclerotherapy to varices of oesophagus
			G17.4	Endoscopic injection sclerotherapy to varices of oesophagus using rigid oesophagoscope
			G43.7	Fibreoptic endoscopic rubber band ligation of upper gastrointestinal tract varices
			T46.1*	Paracentesis abdominis for ascites
			T46.2*	Drainage of ascites not elsewhere specified
Hepatocellular carcinoma	Hospital admission or cancer registry	ICD 10	C22.0	Liver cell carcinoma
Cirrhosis-related death	Mortality	ICD 10	K70.3	Alcoholic cirrhosis of liver
			K71.7	Toxic liver disease with fibrosis and cirrhosis of liver
			K72.1	Chronic hepatic failure
			K74.4	Secondary biliary cirrhosis
			K74.5	Biliary cirrhosis, unspecified
			K74.6	Other and unspecified cirrhosis of liver
			K76.6	Portal hypertension
			I85.0; I859; I98.2	Esophageal varices
			I86.4	Gastric varices
			C22.0	Liver cell carcinoma

ICD-10 refers to International Classification of Disease version 10. OPCS4 refers to Operation/procedure codes version 4. A complication event was considered to be due to cirrhosis morbidity if any of the above codes were present in any diagnostic or cause of death position. However, the OPCS4:T461 and OPCS4:T462 (codes relating to ascites) codes are exceptions to this rule. Here, these codes were only considered to reflect cirrhosis morbidity if accompanied by at least one corroborating ICD code for chronic liver disease (i.e. ICD10: K70-K77). This qualification is necessary because ascites can have non-hepatic causes. N.B. non-cirrhosis mortality (the competing risk event) was defined as a death without any of codes indicated above for a cirrhosis-related death

Table S2: Detailed summary of eligible studies identified from the systematic review

N.B. Table S2 contains 36 data rows and 19 columns, and thus is supplied as a separate Excel file.

Table S3: descriptive statistics for each risk score, including values for specific cut off points

Risk score	Mean	SD	1st percentile	5th percentile	10th percentile	20th percentile	50th percentile	80th percentile	90th percentile	95th percentile	99th percentile
AAR	1.181	0.405	0.550	0.676	0.755	0.865	1.120	1.443	1.651	1.861	2.530
ALBI	-2.432	0.302	-3.082	-2.897	-2.800	-2.682	-2.448	-2.195	-2.044	-1.906	-1.628
ALBI_FIB4	-3.015	0.450	-3.923	-3.672	-3.536	-3.371	-3.041	-2.680	-2.463	-2.268	-1.861
APRI	0.334	0.337	0.131	0.167	0.189	0.219	0.295	0.406	0.494	0.599	0.987
BARD	2.383	0.819	0.000	1.000	1.000	2.000	2.000	3.000	3.000	3.000	4.000
CBR	2.100	0.410	1.453	1.599	1.684	1.794	2.036	2.348	2.571	2.812	3.452
CRPA	4.491	5.167	0.192	0.423	0.663	1.094	2.740	6.714	10.438	14.634	25.378
CirCom	0.058	0.323	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	2.000
Cirrus	-4.632	1.195	-7.674	-6.568	-6.089	-5.552	-4.606	-3.696	-3.204	-2.776	-1.806
DOHA	2.811	1.212	-0.507	0.739	1.295	1.895	2.904	3.781	4.215	4.570	5.338
FIB4	1.348	0.995	0.543	0.691	0.786	0.918	1.236	1.671	1.973	2.292	3.249
FLI	62.165	28.373	4.352	10.064	17.167	32.384	68.639	89.911	95.072	97.446	99.404
ML	0.270	0.242	0.083	0.130	0.151	0.180	0.247	0.333	0.400	0.467	0.667
NAR	0.964	0.330	0.402	0.529	0.602	0.699	0.916	1.197	1.381	1.557	1.987
NFS	-1.795	1.108	-4.358	-3.526	-3.137	-2.681	-1.820	-0.930	-0.417	0.048	1.041
NL	2.371	1.240	0.843	1.165	1.339	1.580	2.152	2.980	3.571	4.207	6.151
PALBI	12.269	1.797	7.830	9.430	10.106	10.871	12.281	13.692	14.457	15.126	16.490
PWC	37.660	11.381	16.678	21.944	24.787	28.437	36.329	45.981	52.042	57.800	71.549
WHR	0.906	0.086	0.710	0.758	0.788	0.830	0.911	0.980	1.016	1.041	1.098
vdMM	157.080	120.762	-130.552	-42.574	2.492	57.053	158.552	257.370	308.382	350.618	437.899

Aspartate aminotransferase to alanine aminotransferase ratio (AAR); Albumin-Bilirubin score (ALBI); Albumin-Bilirubin Fibrosis-4 index (ALBI-FIB4); Aspartate aminotransferase to platelet ratio (APRI); BMI-AST ratio-Diabetes model (BARD); Cirrhosis-specific comordbitiy score (CirCom); Cirrhosis using standard tests (CIRRUS); C-reactive protein to albumin ratio (CRPA); Cystatin to bilirubin ratio (CBR); Doha score (DOHA); Fibrosis-4 index (FIB4); fatty liver index (FLI); monocyte to lymphocyte ratio (ML); Neutrophil to albumin ratio (NAR); Non-alcohol fatty liver disease fibrosis score (NFS); Neutrophil to lymphocyte ratio (NL); Platelet-albumin-bilirubin score (PALBI); Platelet to white cell count ratio (PWC); van der Meer mortality score (vdMM); and the Waist-hip ratio (WHR).

Table S4: UK Biobank study population: genetic characteristics

Locus rsID	Minor allele	Gene	Minor allele frequency (%)	
			Total (N=197,509)	Cirrhosis complication event within 10 years (N=1110)
rs12904	A	*EFNA1*	40.8	41.5
rs2642438	A	*MARC1.*	29.4	27.7
rs708118	C	*WNT3A*	38.8	38.5
rs5743836	G	*TLR9* (dist=603)	16.2	17.1
rs72613567	TA	*HSD17B13*	27.2	25.7
rs888655	A	*ARHGEF28* (dist=4544)	27.6	27.3
rs11134977	C	*FAF2*	45.0	46.1
rs7029757	A	*TOR1B*	9.1	7.6
rs2792751	T	*GPAM*	26.9	28.6
rs1799992	C	*HMBS*	40.4	42.1
rs28929474	T	*SERPINA1*	1.9	3.2
rs58542926	T	*TM6SF2*	7.4	10.4
rs187429064	G	*TM6SF2*	1.1	1.4
rs15052	C	*HNRNPUL1*	17.3	18.5
rs429358	C	*APOE*	15.3	13.4
rs313853	C	*SLC1A5*	33.9	32.8
rs601338	G	*FUT2*	49.2	47.4
rs641738	T	*TMC4*	43.9	45.5
rs1883711	C	*MAFB* (dist=134666)	2.8	2.7
rs738409	G	*PNPLA3*	21.5	32.3

Table S5: Frequency of each type of cirrhosis complication event observed

Type of complication event *			Number of events (col %)
Hospital admission for cirrhosis	HCC presentation	Cirrhosis-related death	
No	No	Yes	33 (3.0)
No	Yes	No	95 (8.6)
Yes	No	No	972 (87.6)
Yes	No	Yes	4 (0.4%)
Yes	Yes	No	6 (0.5%)

*refers to complication events observed at index presentation

Table S6: Patients with a hospital admission due to cirrhosis morbidity (N=982): Frequency of specific ICD and OPCS4 codes in the index admission record.

Code	Code description	Frequency (col %)
ICD10: K70.3	Alcoholic cirrhosis of liver	147 (15.0)
ICD10: K71.7	Toxic liver disease with fibrosis and cirrhosis of liver	≤5 (<5.0)
ICD10: K72.1	Chronic hepatic failure	≤5 (<5.0)
ICD10: K74.4; K745	Secondary biliary cirrhosis	30 (3.1%)
ICD10: K74.6	Other and unspecified cirrhosis of liver	312 (31.8)
ICD10: K76.6	Portal hypertension	294 (29.9)
ICD10: I85.0; I859; I98.2; I98.3; I86.4	Esophageal varices	311 (31.7)
OPCS4: J06.1; J062	Hospital procedures relating to portal hypertension	≤5 (<5.0)
OPCS4: G10.4; G10.8; G10.9; G14.4: G17.4: G43.7	Hospital procedures relating to oesophageal varices	37 (3.7)
OPCS4: T46.1; T46.2	Hospital procedures relating to ascites	117 (11.9)
TOTAL		982 (100.0)

N.B. column percentages can add up to more than 100% because individuals typically have multiple codes within a single hospital admission record.

Table S7: C-index estimate for risk scores over a ten and five year time horizon

Risk score	Ten-year time horizon			Five-year time horizon		
	Estimate	95% CI		Estimate	95% CI	
		Lower	Upper		Lower	Upper
APRI	0.804	0.788	0.820	0.852	0.829	0.874
FIB4	0.780	0.764	0.795	0.832	0.808	0.855
Cirrus	0.745	0.728	0.762	0.786	0.758	0.813
FLI	0.729	0.714	0.744	0.725	0.700	0.751
NFS	0.724	0.707	0.741	0.767	0.739	0.795
Doha	0.712	0.694	0.730	0.759	0.730	0.788
PWC	0.704	0.688	0.720	0.736	0.709	0.762
ALBI-FIB4	0.699	0.682	0.717	0.769	0.741	0.796
WHR	0.683	0.668	0.698	0.675	0.650	0.700
CBR	0.675	0.659	0.692	0.704	0.677	0.731
PALBI	0.673	0.654	0.691	0.713	0.683	0.744
vdMM	0.669	0.652	0.685	0.731	0.704	0.759
ALBI	0.642	0.624	0.659	0.705	0.676	0.734
CRPA	0.626	0.609	0.642	0.654	0.628	0.680
BARD	0.592	0.576	0.609	0.612	0.585	0.639
ML	0.588	0.570	0.606	0.623	0.592	0.653
NAR	0.550	0.532	0.567	0.554	0.523	0.584
AAR	0.536	0.518	0.555	0.604	0.574	0.634
CirCom	0.526	0.508	0.544	0.552	0.520	0.584
NL	0.526	0.517	0.534	0.542	0.526	0.558

Table S8: Ten-year cumulative incidence of cirrhosis complications according to risk threshold.

Risk score	C-index	High risk: >50th percentile		High risk: >80th percentile		High risk: >90th percentile		High risk: >95th percentile		High risk: >99th percentile	
		High risk	10-year CI (%)	High risk	10-year CI (%)	High risk	10-year CI (%)	High risk	10-year CI (%)	High risk	10-year CI (%)
APRI	0.804	No	0.199	No	0.225	No	0.274	No	0.322	No	0.433
		Yes	0.955	Yes	1.984	Yes	3.303	Yes	5.423	Yes	14.830
FIB4	0.780	No	0.210	No	0.272	No	0.324	No	0.362	No	0.458
		Yes	0.943	Yes	1.792	Yes	2.850	Yes	4.657	Yes	12.359
Cirrus	0.745	No	0.253	No	0.306	No	0.348	No	0.390	No	0.465
		Yes	0.901	Yes	1.661	Yes	2.645	Yes	4.137	Yes	11.689
FLI	0.729	No	0.239	No	0.346	No	0.414	No	0.471	No	0.548
		Yes	0.916	Yes	1.506	Yes	2.047	Yes	2.603	Yes	3.494
NFS	0.724	No	0.268	No	0.332	No	0.382	No	0.426	No	0.509
		Yes	0.886	Yes	1.556	Yes	2.329	Yes	3.451	Yes	7.300
Doha	0.712	No	0.310	No	0.338	No	0.366	No	0.409	No	0.480
		Yes	0.844	Yes	1.534	Yes	2.474	Yes	3.778	Yes	10.209
PWC	0.704	No	0.262	No	0.372	No	0.437	No	0.480	No	0.539
		Yes	0.892	Yes	1.398	Yes	1.835	Yes	2.434	Yes	4.396
ALBI-FIB4	0.699	No	0.289	No	0.363	No	0.414	No	0.446	No	0.500
		Yes	0.865	Yes	1.434	Yes	2.044	Yes	3.076	Yes	8.190
WHR	0.683	No	0.279	No	0.416	No	0.472	No	0.515	No	0.562
		Yes	0.875	Yes	1.222	Yes	1.521	Yes	1.765	Yes	2.088
CBR	0.675	No	0.314	No	0.409	No	0.459	No	0.499	No	0.549
		Yes	0.841	Yes	1.249	Yes	1.638	Yes	2.056	Yes	3.334
PALBI	0.673	No	0.350	No	0.381	No	0.417	No	0.454	No	0.524
		Yes	0.804	Yes	1.362	Yes	2.021	Yes	2.915	Yes	5.809
vdMM	0.669	No	0.325	No	0.413	No	0.457	No	0.495	No	0.542
		Yes	0.828	Yes	1.232	Yes	1.653	Yes	2.135	Yes	4.073
ALBI	0.642	No	0.361	No	0.437	No	0.472	No	0.498	No	0.538
		Yes	0.793	Yes	1.138	Yes	1.527	Yes	2.076	Yes	4.480
CRPA	0.626	No	0.378	No	0.468	No	0.515	No	0.535	No	0.567
		Yes	0.777	Yes	1.018	Yes	1.142	Yes	1.384	Yes	1.568
ML	0.588	No	0.461	No	0.488	No	0.508	No	0.527	No	0.561
		Yes	0.694	Yes	0.942	Yes	1.218	Yes	1.561	Yes	2.231
NAR	0.550	No	0.472	No	0.534	No	0.537	No	0.553	No	0.573
		Yes	0.683	Yes	0.753	Yes	0.943	Yes	1.045	Yes	0.987
AAR	0.536	No	0.538	No	0.532	No	0.530	No	0.539	No	0.567
		Yes	0.617	Yes	0.760	Yes	1.006	Yes	1.310	Yes	1.579
NL	0.526	No	0.531	No	0.538	No	0.542	No	0.552	No	0.571
		Yes	0.624	Yes	0.736	Yes	0.897	Yes	1.067	Yes	1.246

Aspartate aminotransferase to alanine aminotransferase ratio (AAR); Albumin-Bilirubin score (ALBI); Albumin-Bilirubin Fibrosis-4 index (ALBI-FIB4); Aspartate aminotransferase to platelet ratio (APRI); BMI-AST ratio-Diabetes model (BARD); Cirrhosis-specific comordbitiy score (CirCom); Cirrhosis using standard tests (CIRRUS); C-reactive protein to albumin ratio (CRPA); Cystatin to bilirubin ratio (CBR); Doha score (DOHA); Fibrosis-4 index (FIB4); fatty liver index (FLI); monocyte to lymphocyte ratio (ML); Neutrophil to albumin ratio (NAR); Non-alcohol fatty liver disease fibrosis score (NFS); Neutrophil to lymphocyte ratio (NL); Platelet-albumin-bilirubin score (PALBI); Platelet to white cell count ratio (PWC); van der Meer mortality score (vdMM); and the Waist-hip ratio (WHR).

Table S9: C-index for APRI; FIB4 and Cirrus risk scores, according to selected characteristics

Subgroup		APRI		FIB-4		Cirrus	
		C-index (95%CI)	heterogeneity statistics	C-index (95%CI)	heterogeneity statistics	C-index (95%CI)	heterogeneity statistics
Age group, years	<50	0.787(0.739-0.835)	Q=2.85(P=0.24);Isq=29.74	0.753(0.706-0.800)	Q=3.09(P=0.21);Isq=35.28	0.683(0.629-0.737)	Q=7.58(P=0.02);Isq=73.61
	50-59	0.820(0.794-0.847)		0.790(0.763-0.818)		0.765(0.737-0.793)	
	≥60	0.793(0.771-0.814)		0.762(0.740-0.784)		0.735(0.712-0.759)	
Gender	Female	0.806(0.777-0.835)	Q=1.08(P=0.30);Isq=7.44	0.766(0.737-0.796)	Q=0.12(P=0.73);Isq=0.00	0.731(0.699-0.763)	Q=0.15(P=0.70);Isq=0.00
	Male	0.787(0.768-0.807)		0.773(0.754-0.791)		0.738(0.718-0.759)	
Ethnicity	White British ancestry	0.807(0.767-0.847)	Q=0.03(P=0.87);Isq=0.00	0.799(0.761-0.838)	Q=1.15(P=0.28);Isq=12.75	0.757(0.712-0.801)	Q=0.39(P=0.53);Isq=0.00
	Other	0.804(0.786-0.821)		0.776(0.759-0.793)		0.741(0.723-0.760)	
Deprivation quintile	Q1 (least deprived)	0.808(0.765-0.851)	Q=1.05(P=0.90);Isq=0.00	0.771(0.728-0.813)	Q=1.75(P=0.78);Isq=0.00	0.739(0.692-0.786)	Q=3.29(P=0.51);Isq=0.00
	Q2	0.796(0.755-0.837)		0.762(0.719-0.805)		0.728(0.680-0.776)	
	Q3	0.797(0.760-0.834)		0.791(0.758-0.824)		0.717(0.675-0.760)	
	Q4	0.817(0.786-0.849)		0.793(0.760-0.826)		0.762(0.728-0.796)	
	Q5 (most deprived)	0.801(0.773-0.829)		0.782(0.755-0.810)		0.751(0.722-0.780)	
Alcohol intake, units/week	<15	0.796(0.770-0.823)	Q=2.63(P=0.27);Isq=24.08	0.777(0.752-0.801)	Q=2.16(P=0.34);Isq=7.21	0.732(0.703-0.762)	Q=4.41(P=0.11);Isq=54.67
	15-49	0.811(0.775-0.847)		0.796(0.759-0.832)		0.778(0.743-0.813)	
	≥50	0.704(0.576-0.831)		0.704(0.576-0.831)		0.704(0.576-0.831)	
Type 2 diabetes	No	0.791(0.773-0.810)	Q=5.01(P=0.03);Isq=80.04	0.764(0.746-0.782)	Q=10.13(P=0.00);Isq=90.12	0.730(0.710-0.750)	Q=6.97(P=0.01);Isq=85.65
	Yes	0.830(0.802-0.859)		0.819(0.790-0.848)		0.781(0.749-0.813)	
BMI category, Kg/m^2	<30	0.790(0.765-0.816)	Q=1.62(P=0.20);Isq=38.42	0.772(0.748-0.796)	Q=1.65(P=0.20);Isq=39.44	0.734(0.707-0.761)	Q=1.07(P=0.30);Isq=6.32
	≥30	0.811(0.792-0.831)		0.793(0.773-0.812)		0.752(0.731-0.774)	

APRI=asparate aminotransferase to platelet ratio index; Fib-4=fibrosis 4 index; Cirrus=cirrhosis using standard tests score; Q=Cochran Q statistic; P=p-value for Cochran Q statistic; Isq=I^2 statistic.

Table S10: Cox regression coefficients for genetic-only model

Locus rsID	Gene	Effect allele	Ref allele	Full cohort (main analysis)		White British ancestry subset (sensitivity analysis)	
				HR (95%CI)	P	HR (95%CI)	P
rs12904	*EFNA1*	A	G	1.03 (0.94-1.12)	0.53	1.04 (0.95-1.14)	0.43
rs2642438	*MARC1.*	A	G	0.92 (0.84-1.01)	0.07	0.94 (0.85-1.04)	0.26
rs708118	*WNT3A*	C	T	0.99 (0.91-1.07)	0.75	0.98 (0.89-1.07)	0.64
rs5743836	*TLR9* (dist=603)	G	A	1.07 (0.96-1.20)	0.22	1.10 (0.98-1.24)	0.10
rs72613567	*HSD17B13*	TA	T	0.92 (0.84-1.01)	0.09	0.88 (0.79-0.98)	0.02
rs888655	*ARHGEF28* (dist=45	A	G	0.98 (0.89-1.08)	0.70	0.98 (0.89-1.08)	0.69
rs11134977	*FAF2*	C	T	1.04 (0.96-1.14)	0.30	1.01 (0.92-1.11)	0.84
rs7029757	*TOR1B*	A	G	0.81 (0.69-0.95)	0.01	0.77 (0.65-0.92)	0.004
rs2792751	*GPAM*	T	C	1.09 (0.99-1.19)	0.07	1.08 (0.98-1.20)	0.12
rs1799992	*HMBS*	C	T	1.07 (0.99-1.17)	0.11	1.09 (0.99-1.19)	0.08
rs28929474	*SERPINA1*	T	C	1.74 (1.37-2.21)	5.0×10^{-6}	1.75 (1.36-2.26)	1.2×10^{-5}
rs58542926	*TM6SF2*	T	C	1.45 (1.27-1.67)	7.9×10^{-8}	1.47 (1.27-1.70)	3.1×10^{-7}
rs187429064	*TM6SF2*	G	A	1.36 (0.96-1.92)	0.09	1.33 (0.91-1.93)	0.14
rs15052	*HNRNPUL1*	C	T	1.09 (0.98-1.21)	0.12	1.11 (0.99-1.25)	0.07
rs429358	*APOE*	C	T	0.86 (0.76-0.97)	0.02	0.88 (0.77-1.00)	0.05
rs313853	*SLC1A5*	C	T	0.95 (0.87-1.04)	0.23	0.95 (0.86-1.04)	0.26
rs601338	*FUT2*	G	A	0.93 (0.85-1.01)	0.07	0.91 (0.83-1.00)	0.05
rs641738	*TMC4*	T	C	1.07 (0.98-1.16)	0.14	1.06 (0.97-1.16)	0.21
rs1883711	*MAFB* (dist=134666)	C	G	0.95 (0.73-1.23)	0.68	1.00 (0.76-1.31)	1.00
rs738409	*PNPLA3*	G	C	1.75 (1.60-1.91)	6.0×10^{-35}	1.76 (1.60-1.94)	1.3×10^{-30}

Table S11: Risk score improvement through addition of genetic risk data.

N.B. Table S11 contains 73 data rows and thus is supplied as a separate Excel file.

Table S12. Relationship between specific risk score value and ten year cumulative incidence

Risk score	Risk score group	Mean Risk Score value	# Complication events	# Competing risk events (non cirrhosis death)	Ten-year cumulative incidence (%)					
					Cirrhosis complications			Competing risk event		
					Estimate	Lower	Upper	Estimate	Lower	Upper
APRI	Decile 1	0.16	33	1017	0.18	0.12	0.25	5.32	5.01	5.64
	Decile 2	0.20	30	907	0.16	0.11	0.22	4.72	4.42	5.02
	Decile 3	0.23	29	857	0.15	0.11	0.22	4.46	4.18	4.76
	Decile 4	0.26	46	957	0.24	0.18	0.32	5.00	4.69	5.31
	Decile 5	0.28	51	945	0.26	0.20	0.35	4.93	4.63	5.24
	Decile 6	0.31	36	969	0.19	0.13	0.26	5.04	4.74	5.36
	Decile 7	0.34	55	1021	0.29	0.22	0.38	5.31	5.00	5.64
	Decile 8	0.38	62	1036	0.33	0.25	0.42	5.38	5.07	5.71
	Decile 9	0.44	128	1108	0.67	0.56	0.78	5.76	5.43	6.09
	Decile 10	0.54	114	574	1.19	0.98	1.42	5.97	5.51	6.46
	≥95th percentile	0.72	237	575	3.07	2.70	3.47	7.48	6.90	8.08
	≥99th percentile	1.72	289	190	14.82	13.28	16.44	9.81	8.53	11.18
FIB4	Decile 1	0.67	19	558	0.10	0.06	0.15	2.92	2.69	3.16
	Decile 2	0.86	39	635	0.21	0.15	0.28	3.33	3.08	3.59
	Decile 3	0.97	35	749	0.18	0.13	0.26	3.89	3.63	4.18
	Decile 4	1.08	47	810	0.25	0.18	0.32	4.22	3.94	4.51
	Decile 5	1.18	59	873	0.31	0.24	0.40	4.55	4.26	4.85
	Decile 6	1.29	61	1012	0.32	0.25	0.41	5.27	4.96	5.59
	Decile 7	1.42	65	1094	0.34	0.27	0.43	5.69	5.37	6.03
	Decile 8	1.58	89	1222	0.47	0.38	0.57	6.34	6.01	6.70
	Decile 9	1.80	141	1365	0.73	0.62	0.86	7.08	6.73	7.45
	Decile 10	2.11	101	820	1.04	0.86	1.26	8.49	7.95	9.06
	≥95th percentile	2.61	212	769	2.73	2.38	3.11	9.92	9.27	10.60
	≥99th percentile	5.17	242	248	12.35	10.94	13.85	12.77	11.33	14.30
Cirrus	Decile 1	-3.93	49	964	0.26	0.19	0.34	5.03	4.73	5.35
	Decile 2	-2.88	37	858	0.20	0.14	0.27	4.50	4.21	4.80
	Decile 3	-2.42	47	839	0.25	0.18	0.33	4.38	4.09	4.67
	Decile 4	-2.08	62	823	0.32	0.25	0.41	4.29	4.01	4.58
	Decile 5	-1.79	47	877	0.25	0.19	0.33	4.57	4.28	4.87
	Decile 6	-1.50	53	971	0.28	0.21	0.36	5.06	4.76	5.38
	Decile 7	-1.22	82	969	0.43	0.34	0.53	5.03	4.73	5.34
	Decile 8	-0.89	91	1067	0.48	0.39	0.58	5.54	5.22	5.87
	Decile 9	-0.49	135	1166	0.70	0.59	0.83	6.07	5.74	6.41
	Decile 10	-0.02	111	711	1.16	0.96	1.39	7.39	6.87	7.92
	≥95th percentile	0.58	173	654	2.23	1.92	2.58	8.49	7.88	9.12
	≥99th percentile	1.90	223	256	11.42	10.06	12.88	13.26	11.79	14.82

APPENDIX A: Systematic review

The PICOTS system was used to frame our eligibility and exclusion criteria [1]; as follows:

Population: Studies based on either: a) adults with a risk factors for CLD (i.e. obesity, type 2 diabetes, excess alcohol intake, viral hepatitis) or b) adults with a CLD diagnosis.

Intervention/Model: A risk score was defined as any type of risk metric derive from two or more variables. We excluded risk scores that could not be calculated using core data from the UK biobank (i.e. data variables available for <80% of participants).

We also excluded scores for biliary specific disease (i.e. primary sclerosing cholangitis and primary biliary cirrhosis) which are not intended to be extrapolated to other CLD aetiologies.

Comparison: Risk scores with prognostic ability against cirrhosis-related events were selected. Prognostic ability was defined as a statistically significant association or concordance index >0.55.

Outcome: All cirrhosis complication events were included as eligible prognostic outcomes. For example, incident decompensated cirrhosis, liver-related mortality and liver cancer. All-cause mortality was included as an outcome if the study comprised patients with established CLD (i.e secondary care cohort study). Our focus was on prognostic outcomes, and thus performance of risk scores in a diagnostic context (i.e. detection of clinically significant fibrosis) was not considered.

Timing: model prediction over all time horizons were considered.

Setting: All settings were considered (i.e. secondary care, primary care and community settings).

The search was executed on 6th June 2021 using the PubMed platform. The following search terms were used.

CLD complications:

(("decompensated cirrhosis"[tiab]) OR ("advanced liver disease"[tiab]) OR ("cirrhosis complication*"[tiab]) OR ("severe liver disease"[tiab]) OR ("severe liver event"[tiab]) OR ("liver morbidity"[tiab]) OR ("liver-related morbidity"[tiab]) OR ("liver mortality"[tiab]) OR ("liver-related mortality"[tiab]))

AND

Prognostic studies:

((predict[tiab]) OR (progn*[tiab]) OR ("risk prediction"[tiab]) OR ("risk score"[tiab]) OR ("risk calculation"[tiab]) OR ("risk assessment"[tiab]) OR ("c-statistic"[tiab]) OR (discrimination[tiab]) OR (calibration[tiab]) OR (AUC[tiab]) OR ("area under the curve"[tiab]) OR ("area under the receiver operator curve"[tiab]))

The following risk scores identified in our search were excluded as they could not be calculated for UKB participants (.i.e. one or more components of the score was unavailable):

Lok index; NAFLD activity score; MELD; MELD sodium; Child Pugh score; Hepascore; Fibrometer; INR to albumin ratio; CAGE-B; SAGE-B; PAGE-B; CLIF-C AD; Fibrotest; Tapper et al encephalopathy risk score (Hepatology; 2018;68:1498-1507); Liver fat score; Enhanced liver fibrosis test (ELF)

REFERENCE

[1] Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ. 2017;356:I6460.

APPENDIX B: COVARIATES USED TO DEFINE THE STUDY POPULATION

BMI was determined from each participant's height and weight at the time of their assessment visit. Standing height was measured via the Seca202 height measure, and body weight was measured from the Tanita BC-418 MA body composition analysis. [1] Information on alcohol intake was elicited through a computer-assisted touchscreen system at UKB interview. Participants were asked to report their average alcohol intake per week/month in terms of the number of: glasses of red wine (Field IDs: 1568, 4407), glasses of champagne/white wine (Field IDs: 1578, 4418), pints of beer/cider (Field IDs: 1588, 4429), measures of spirits (Field IDs: 1598, 4440), glasses of fortified wine (Field IDs: 1608, 4451), and glasses of "other" types of alcoholic drinks (Field IDs: 5364, 4462). Non-weekly and occasional drinkers were asked to report consumption in an "average month" to generate more reliable estimates for infrequent drinkers. For each participant, we calculated the average number of alcohol units consumed per week, assuming there are 2 units (16g) of pure alcohol in a pint of beer/cider; 1.5 units (12g) in a glass of red wine, champagne, white wine, fortified wine, and "other" alcoholic drink; and 1 unit (8g) in a measure of spirits. These conversions are comparable to those used in the Health Survey for England methods protocol [2]. Although UKB participants were asked about diagnosis of diabetes mellitus

(UKB Field ID: 2443), they were not asked specifically about T2DM. Thus, we inferred T2DM status by taking all individuals who reported a diabetes diagnosis (UKB Field ID: 2443), and excluding those with evidence of non-type2 diabetes. Evidence of non-type 2 diabetes was based on either: 1) self-reported type 1 diabetes in UKB nurse interview; OR; 2) hospital admission for type 1 diabetes (ICD10: E10); OR 3) self-reported gestational diabetes (UKB field ID: 4041).

REFERENCES:

[1] UK Biobank. Protocol for a large-scale prospective epidemiological resource. Available at: https://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf [Accessed April 20, 2020]

[2] Osborne B, Cooper V. Health Survey for England. 2017 adult health related behaviours. Version 2. 2019. ISBN: 978-1-78734-255-2.

APPENDIX C: IDENTIFYING RELEVANT GENETIC RISK LOCI

Relevant genetic loci were identified by reviewing previous genetic association studies for alcohol-related liver disease, NAFLD or mixed aetiology cohorts

We considered 22 independent genetic loci that have been associated with either cirrhosis progression, cirrhosis complications, or HCC. To limit attrition in our final study population, we omitted 2/21 loci with a high missing proportion in the UKB genetic dataset (>3%; see Table 3). Thus, our final analysis considered 20 loci in total, enumerated in the table below.

Table 3: Genetic loci included in risk score augmentation analysis

SNP	Chr:Basepair position	Minor allele	Ref allele	Missing proportion	MAF	Nearest gene	Position	Phenotype	Source (Reference)
Variants included									
rs12904	1:155106697	A	G	0.000	0.411	*EFNA1*	UTR3	Fibrosis/cirrhosis	[3]
rs2642438	1:220970028	A	G	0.000	0.291	*MARC1.*	exonic	Fibrosis/cirrhosis	[2-4]
rs708118	1:228201801	C	T	0.013	0.389	*WNT3A*	intronic	HCC	[10]
rs5743836	3:52260782	G	A	0.000	0.162	*TLR9* (dist=603)	upstream	Hepatic encephalopathy	[8]
rs72613567	4:88231392	TA	T	0.000	0.270	*HSD17B13*	intronic	Fibrosis/cirrhosis	[1-5]
rs2562582	5:36605360	C	T	0.050	0.179	*SLC1A3* (dist=1097)	intergenic	Hepatic encephalopathy	[8]
rs888655	5:72917439	A	G	0.003	0.273	*ARHGEF28* (dist=4544	intergenic	Fibrosis/cirrhosis	[3]
rs11134977	5:175904141	C	T	0.014	0.448	*FAF2*	intronic	Fibrosis/cirrhosis	[5]
rs9398804	6:126703390	A	T	0.039	0.445	*CENPW*	intronic	Fibrosis/cirrhosis	[3]
rs7029757	9:132566666	A	G	0.010	0.090	*TOR1B*	intronic	Fibrosis/cirrhosis	[3]
rs2792751	10:113940329	T	C	0.000	0.269	*GPAM*	exonic	Fibrosis/cirrhosis	[6]
rs1799992	11:118957246	C	T	0.014	0.399	*HMBS*	intronic	Fibrosis/cirrhosis	[3]
rs28929474	14:94844947	T	C	0.000	0.019	*SERPINA1*	exonic	Fibrosis/cirrhosis	[1-5]
rs58542926	19:19379549	T	C	0.000	0.074	*TM6SF2*	exonic	HCC; fibrosis/cirrhosis	[1-5, 10]
rs187429064	19:19380513	G	A	0.009	0.011	*TM6SF2*	exonic	Fibrosis/cirrhosis	[6]
rs15052	19:41813375	C	T	0.005	0.170	*HNRNPUL1*	UTR3	Fibrosis/cirrhosis	[8]
rs429358	19:45411941	C	T	0.000	0.154	*APOE*	exonic	HCC; fibrosis/cirrhosis	[3,6,7]
rs313853	19:47287939	C	T	0.024	0.339	*SLC1A5*	UTR5	Hepatic encephalopathy	[8]
rs601338	19:49206674	G	A	0.000	0.499	*FUT2*	exonic	Hepatic encephalopathy	[8]
rs641738	19:54676763	T	C	0.009	0.437	*TMC4*	exonic	Fibrosis/cirrhosis	[9]
rs1883711	20:39179822	C	G	0.009	0.028	*MAFB* (dist=134666)	intergenic	Fibrosis/cirrhosis	[3]
rs738409	22:44324727	G	C	0.000	0.216	*PNPLA3*	exonic	HCC; fibrosis/cirrhosis	[1-5, 10]
Relevant variants not included*									
rs2562582	5:36605360	C	T	0.050	0.179	*SLC1A3* (dist=1097)	intergenic	Hepatic encephalopathy	[8]
rs9398804	6:126703390	A	T	0.039	0.445	*CENPW*	intronic	fibrosis/cirrhosis	[3]

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.

REFERENCES:

[1] Buch S, Stickel F, Trepo E, Way M, Herrmann A, Nischalke HD, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet. 2015;47:1443-8.

[2] Innes H, Buch S, Hutchinson S, Guha IN, Morling JR, Barnes E, et al. Genome-wide association study for alcohol-related cirrhosis identified risk loci in MARC1 and HNRNPUL1. Gastroenterology. 2020;159:1276-1289.

[3] Emdin CA, Haas M, Ajmera V, Simon TG, Homburger J, Neban C, et al. Association of genetic variation with cirrhosis: A multi-trait genome-wide association and gene-environment interaction study. Gastroenterology. 2021; 160:1620-1633.

[4] Emdin CA, Haas ME, Khera AV, Aragam K, Chaffin M, Klarin D, Hindy G, et al. A missense variant in mitochondrial amidoxime reducing component 1 gene and protection against liver disease. PLoS Genet. 2020;16:e1008629.

[5] Schwantes-An TH, Darlay R, Mathurin P, Masson S, Liangpunsakul S,Mueller S, et al. Genome-wide association study and meta-analysis on alcohol-associated cirrhosis identifies genetic risk factors. Hepatology. 2021;73:1920-1931.

[6] Jamialahmadi O, Mancina RM, Ciociola E, Tavaglione F, Luukkonen PK, Baselli G, Malvestiti F, et al. Exome-wide association study on alanine aminotransferase identifieds sequence variants in the GPAM and APOE associated with fatty liver disease. Gastroenterology. 2021;160:1634-1646.

[7] Innes H, Nischalke HD, Guha IN, Weiss KH, Irving W, Gotthardt D, et al. The rs429358 locus in apolipoprotein E is associated with hepatocellular carcinoma in patients with cirrhosis. Hepatol Commun. 2021; doi: 10.1002/hep4.1886.

[8] Gil-Gomez A, Ampuero J, Rojas A, Gallego-Duran R, Munoz-Hernandez R, Rico MC, et al. Development and validation of a clinical-genetic risk score to predict hepatic encephalopathy in patients with liver cirrhosis. Am J Gastroenterol. 2021;116:1238-1247.

[9] Teo K, Abeysekera KWM, Adams L, Aigner E, Anstee QM, Banales JM, et al. rs641738 C>T near MBOAT7 is associated with liver fat, ALT and fibrosis in NAFLD: A meta analysis. J Hepatol. 2021;74:20-30.

[10] Trepo E, Caruso S, Yang J, Imbeaud S, Couchy G, Bayard Q, et al. Common genetic variation in alcohol-related hepatocellular carcinoma: a case-control genome-wide association study. Lancet Oncol. 2021. doi: 10.1016/S1470-2045(21)00603-3.

APPENDIX D: RISK SCORE FORMULA

Risk scores were calculated using the formulas below:

- Neutrophil count to albumin ratio (NAR)=neutrophil cell count [10^9 cells/L]/albumin[g/dL]
- Aspartate aminotransferase to platelet ratio index (APRI)=((aspartate aminotransferase[U/L]/aspartate aminotransferase upper limit of normal)/platelet count [10^9 cells/L])*100
- Fibrosis 4 index (FIB4)=(age[years]*aspartate aminotransferase[U/L])/(platelet count [10^9 cells/L]*sqrt(alanine aminotransferase[U/L]))
- Cirrhosis using standard tests (Cirrus)=-8.415+(-0.222*albumin[g/L]) +(-0.011*creatinine [µmol/L])+(0.016*bilirubin[µmol/L])+(0.084*mean corpuscular volume [femtolitres])+(-0.017*platelet count[10^9 cells/L])+(0.153*total protein[g/L])
- Albumin bilirubin score (ALBI)=log10(bilirubin[µmol/L])+(albumin[g/L]* -0.085)

- Platelet albumin bilirubin score (PALBI)= (log10(bilirubin[umol/L)*2.02)-(0.37*(log10(bilirubin[umol/L)^2))-(0.04*albumin[g/L])-(3.48*log10(platelet count[10^9 cells/L])+(1.01*(log10(platelet count[10^9 cells/L]^2))
- ALBI-FIB4=(ALBI*1.331)+(FIB4*0.165)
- NAFLD fibrosis score=1.675+(0.037*age[years])+(0.094*BMI[kg/m2])+(1.13*diabetes[1=Yes; 0=No])+(0.99*AST:ALT ratio[U/L])-(0.013*platelet count[10^9 cells/L])-(0.66*albumin[g/dL])
- Fatty liver index (FLI)= exp((0.953*ln(triglycerides[mg/dl]))+(0.139*BMI[kg/m2])+(0.718*ln(GGT[U/L]))+(0.053*waist circumference[cm])-15.745)/ 1+exp((0.953*ln(triglycerides[mg/dl])+(0.139*BMI[kg/m2])+(0.718*ln(GGT[U/L]))+(0.053*waist circumference[cm])-15.745)
- C-reactive protein to albumin ratio (CRPA)= albumin[g/dL]/c-reactive protein[mg/L]
- BARD score=(1* BMI>28[1=yes; 0=no])+(2* AST/ALT ratio>0.8[1=Yes; 0=No])+(1*diabetes[1=Yes; 0=No])
- DOHA score= 8.5-(0.2*albumin[g/dl])+(0.01*aspartate aminotransferase[U/L])-(0.02*platelet count[10^9 cells/L])
- Cystatin to bilirubin ratio=(1.593*cystatin[mg/L])+(0.068*bilirubin[µmol/L)
- Van der Meer mortality score (vdMM)=6*age[years])-(platelet count [10^9 cells/L])+(258.8*log10(AST:ALT ratio))+(64.5*male gender[1=yes; 0=no])
- Monocyte count to lymphocyte ratio (ML)=monocyte[10^9 cells/L]/lymphocyte[10^9 cells/L]
- Neutrophil count to lymphocyte ratio (NLR)=neutrophil[10^9 cells/L]/lymphocyte[10^9 cells/L]
- Platelet to white cell count ratio (PWC)=platelet count[10^9 cells/L]/white cell count[10^9 cells/L]
- Waist hip ratio (WHR)= waist circumference [cm]/hip circumference [cm]
- Aspartate aminotransferase to alanine aminotransferase ratio (AAR) = aspartate aminotransferase [IU/L] / Alanine aminotransferase [IU/L]

Please note the following:

1) Some scores include laboratory prognostic factors whose values are benchmarked against an “upper limit of normal” (ULN) or a “lower limit of normal” (LLN). One example of this is the APRI score. To accommodate this, we used the 90th percentile

observed in the entire UKB population to define the upper limit of normal; conversely, the 10th percentile was used to define the lower limit of normal.

2) The algorithm for the CirCom score is complex and cannot be expressed in a single formula. Thus, please consult the Figure 1 schematic in the original paper for details of how this risk score is calculated. [1]
3) Although the original Cirrus score is based on seven prognostic factors, only six of these were available in the UK biobank (i.e. serum sodium was unavailable). [2] Thus, we tested a modified version of six-variable version of Cirrus. The coefficients for this modified version were sent to us by the creators of the Cirrus score to support this study.

REFERENCES:

[1] Jepsen P, Vilstrup H, Lash TL. Development and validation of a comorbidity scoring system for patients with cirrhosis. Gastroenterology. 2014; 146:147-156.

[2] 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.

APPENDIX E: TRIPOD CHECKLIST

Section/Topic		Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2-3
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3-4
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	3
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	4-5
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	4
	5b	Describe eligibility criteria for participants.	4-5
	5c	Give details of treatments received, if relevant.	NA
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	5

	6b	Report any actions to blind assessment of the outcome to be predicted.	None
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	5; Appendix D
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	None
Sample size	8	Explain how the study size was arrived at.	4-5
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	5;13
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	5; Appendix D
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	6-8
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	7-8
Risk groups	11	Provide details on how risk groups were created, if done.	6; 8
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	NA
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Figure S1
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	9
	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model performance	16	Report performance measures (with CIs) for the prediction model.	Figure 1; Table S11
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	Table S12; Figure 2
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	12-14
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	NA
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	10-14
Implications	20	Discuss the potential clinical use of the model and implications for future research.	10-14
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	No such materials provided
Funding	22	Give the source of funding and the role of the funders for the present study.	1

We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.