Performance of models to predict hepatocellular carcinoma risk among UK patients with cirrhosis and cured hepatitis C infection

Hamish Innes, Peter Jepsen, Scott McDonald, John Dillon, Victoria Hamill, Alan Yeung, Jennifer Benselin, April Went, Andrew Fraser, Andrew Bathgate, M. Azim Ansari, Stephen T Barclay, David Goldberg, Peter C Hayes, Philip Johnson, Eleanor Barnes, William Irving, Sharon Hutchinson, Indra Neil Guha

Table of contents

APPENDIX A

PROVENANCE OF HCC PREDICTION MODELS ASSESSED IN THIS STUDY

The non-genetic HCC prediction models assessed were developed from patients in China (aMAP), Canada (THRI), US (VHA model), and France (ANRS CO12 model) (Table 1). The number of patients in the model development dataset varied 10-fold; from 720 in the ANRS CO12 model, to 7689 for the US VHA model. The aMAP model was distinct insofar as the development dataset was made up entirely of hepatitis B patients who were mostly without cirrhosis (81% pre-cirrhotic). The four models use overlapping prognostic factors, with age, gender, platelet count and albumin all included in at least 2/4 models. Reported discrimination performance ranged from C-index 0.72 (ANRS-CO12 model) to C-index 0.82 (aMAP).

The genetic risk models were based on loci associated with increased liver steatosis. Specifically: rs738409:G in *PNPLA3* (both models); rs58542926:T in *TM6SF2* (both models); rs641837:T in *MBOAT7* (Dongiovanni et al model); rs1260326:T in *GCKR* (Dongiovanni et al model) and rs72613567:T in *HSD17B13* (Gellert-Kristensen et al model). Variants were weighted differently in each model (Table S1). Dongiovanni et al score used a data driven approach, where each variant was weighted according to the magnitude of association with liver fat content in the Dallas Heart study. By contrast, the Gellert-Kristensen et al score assumed nominally equal weightings for all variants. C-index values for discrimination were not reported by the authors of these genetic models.

Fig. S1. Derivation of final sample.

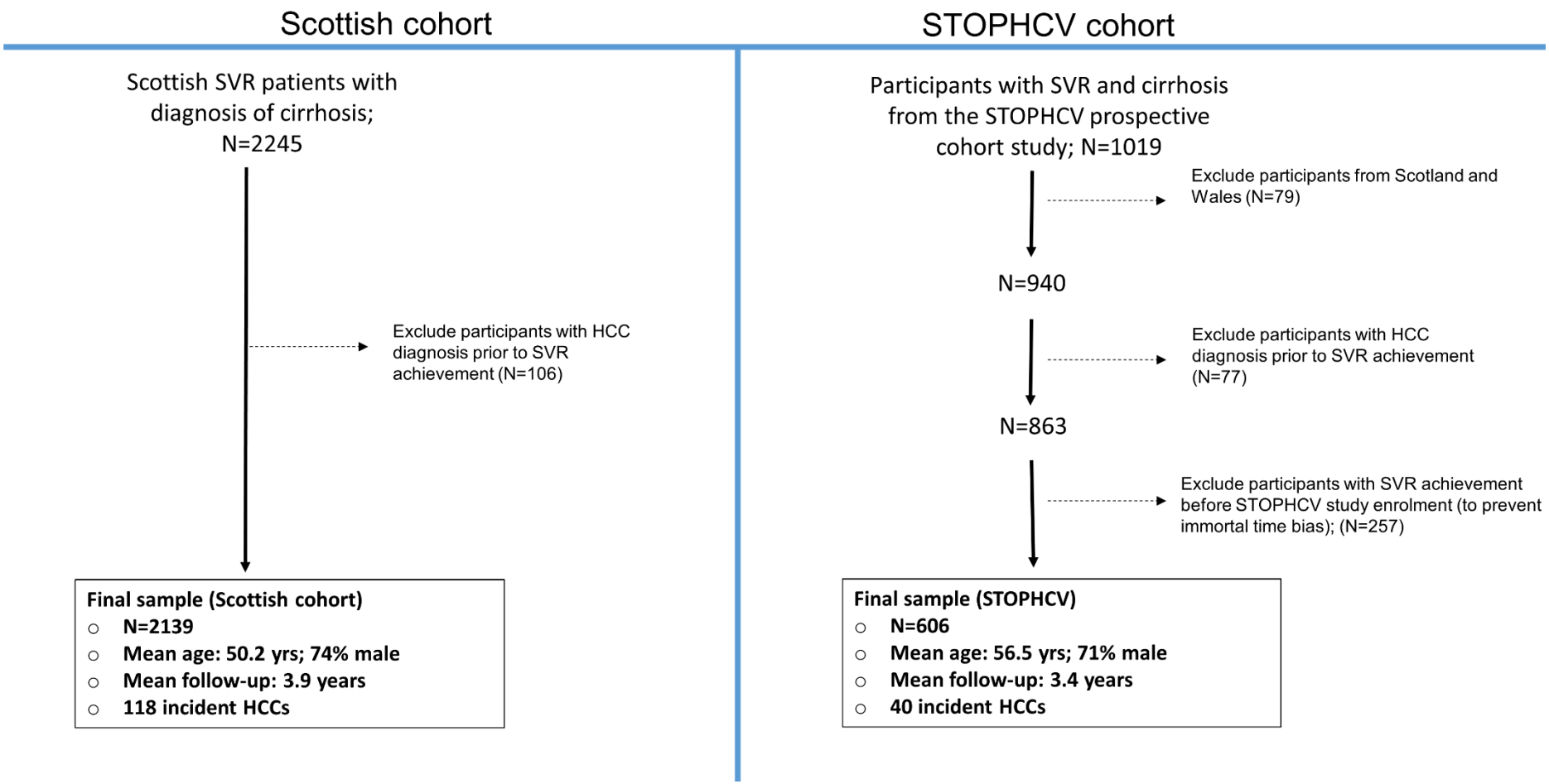
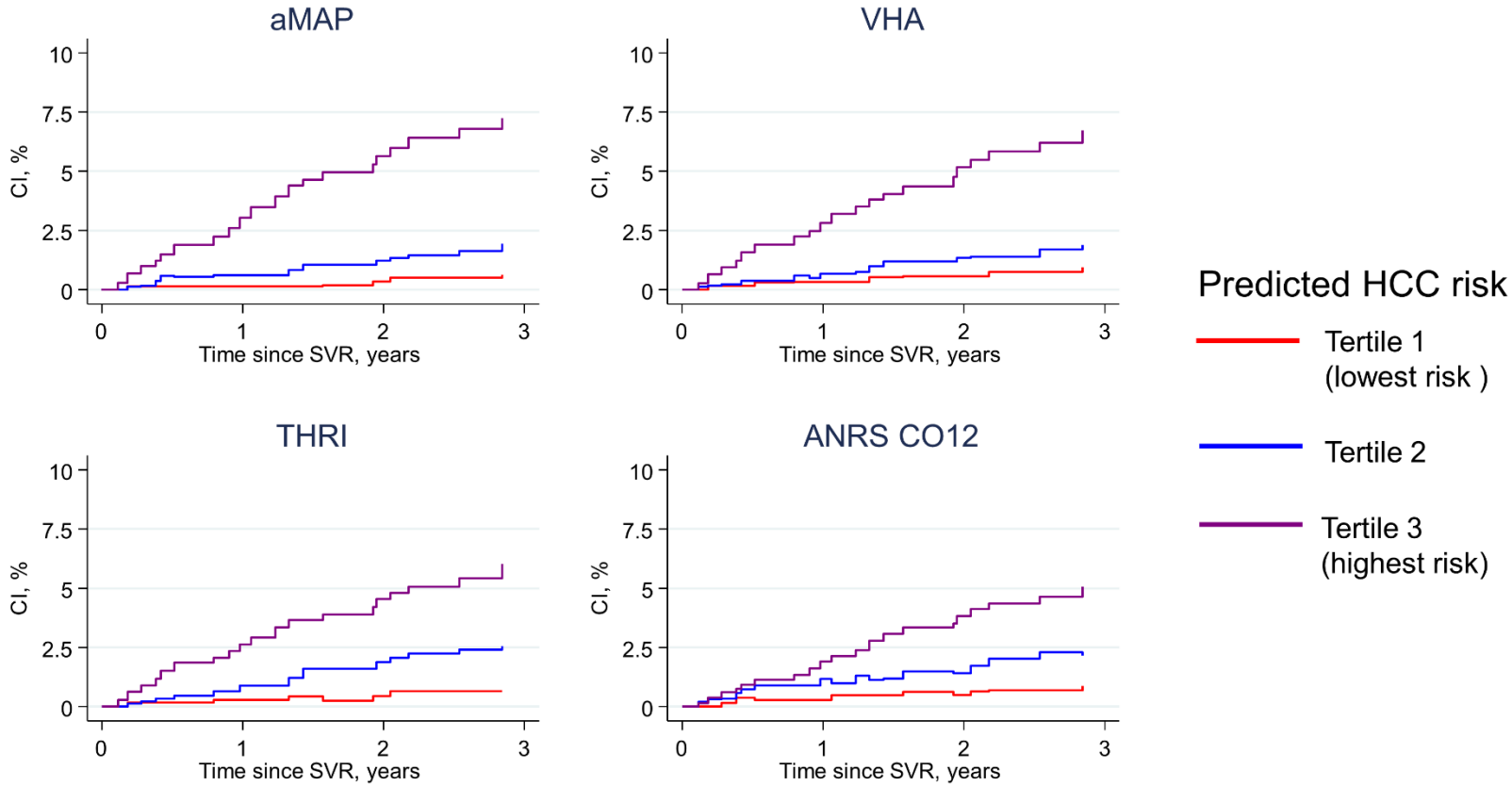
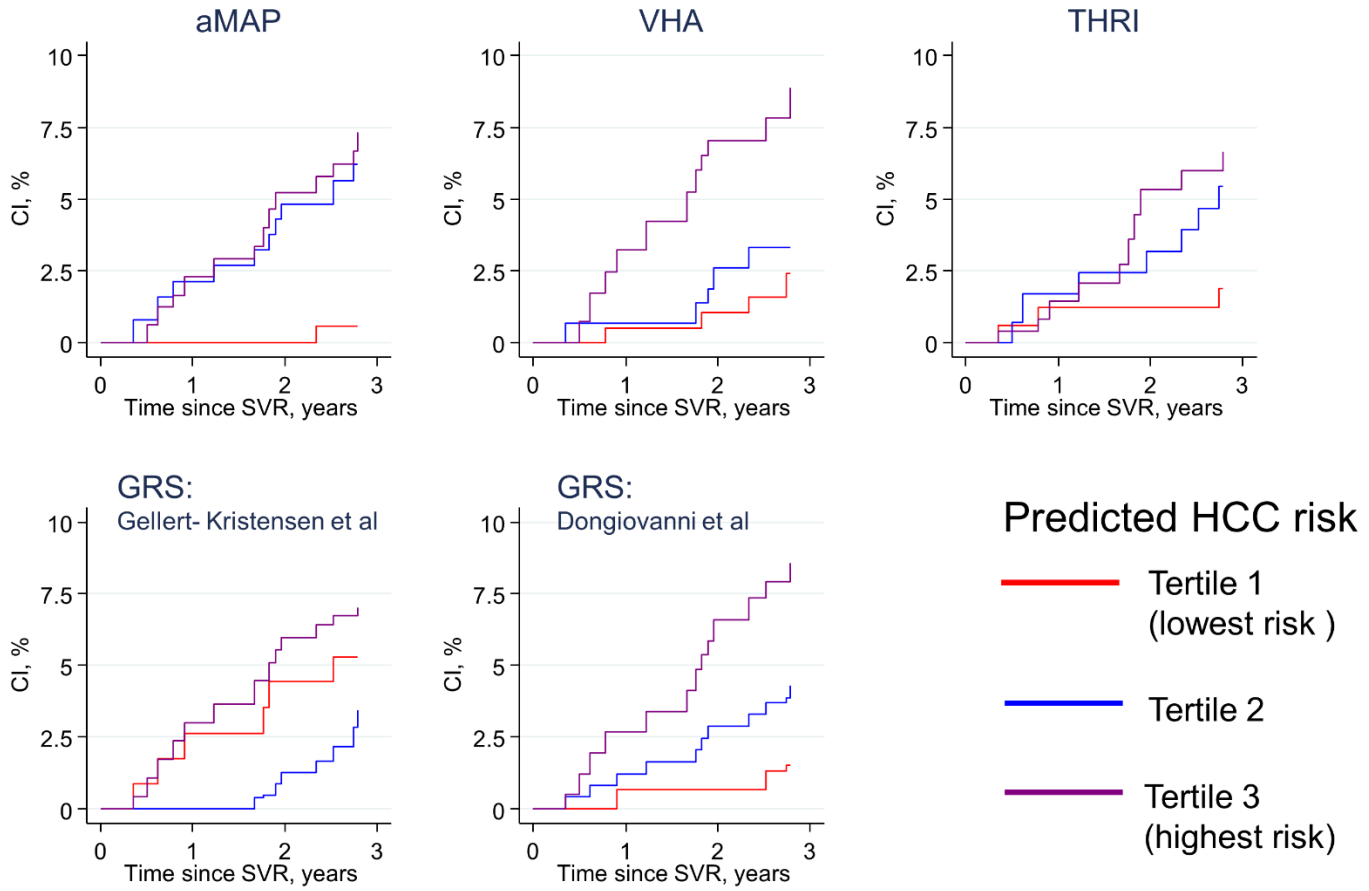


Fig. S2. Cumulative incidence (CI) of HCC according to predicted HCC risk in Scottish cohort.



Risk tertiles refer to three groups: a) those whose prediction is in the 33rd percentile or lower; b) those in the 33rd -67th percentile and c) those in the 68th-100th percentile. Cumulative incidence curves are generated non-parametrically with non-HCC mortality treated as a competing risk event.

Fig. S3. Cumulative incidence (CI) of HCC according to predicted HCC risk in STOP-HCV cohort.



GRS=Genetic Risk Score. Risk tertiles refer to three groups: a) those whose prediction is in the 33rd percentile or lower; b) those in the 33rd -67th percentile and c) those in the 68th-100th percentile. Cumulative incidence curves are generated non-parametrically with non-HCC mortality treated as a competing risk event.

Fig. S4. Concordance index (C-index) in Scottish cohort according to subgroup.

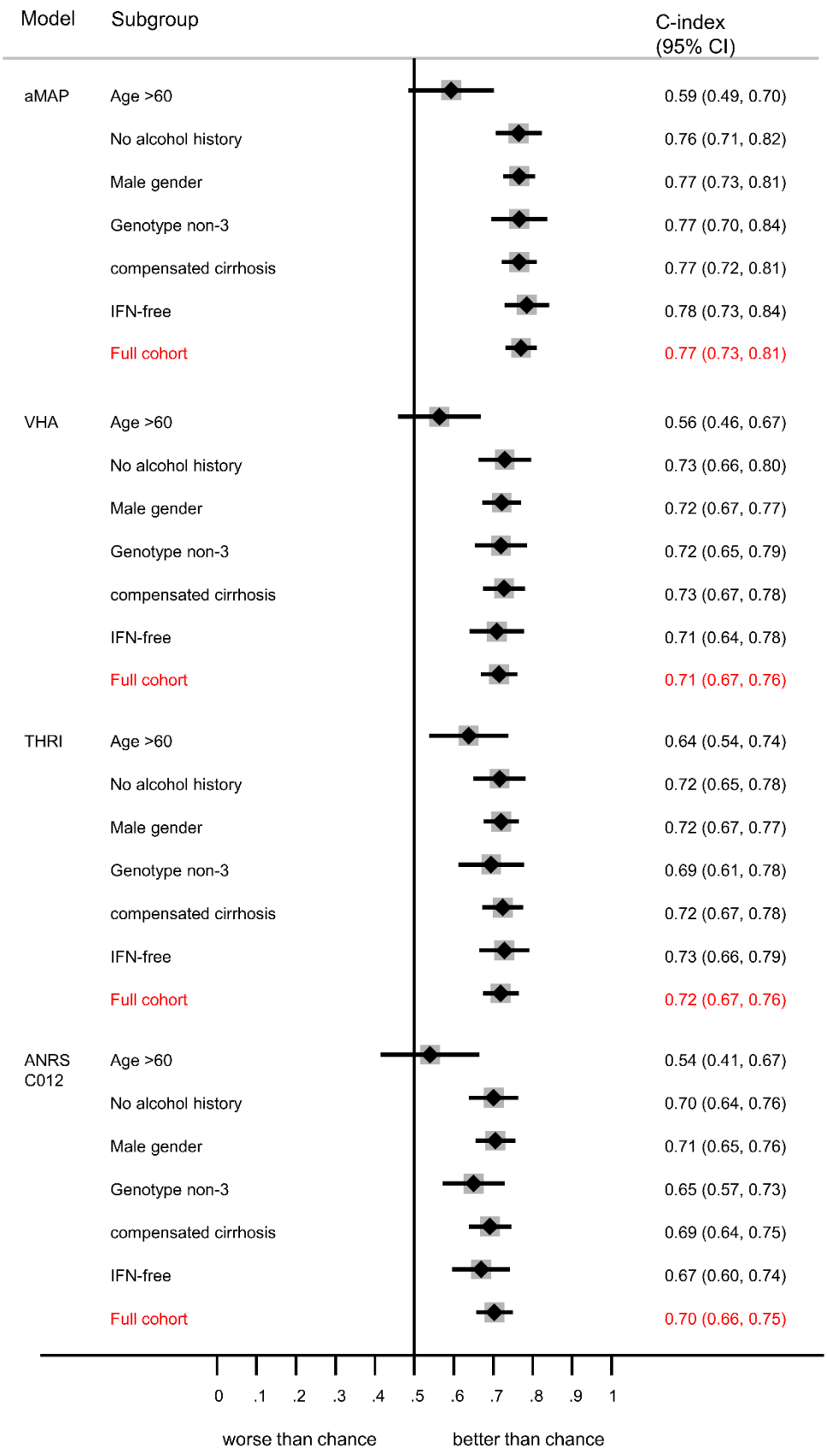


Fig. S5. Concordance index (C-index) in STOP-HCV cohort according to subgroup.

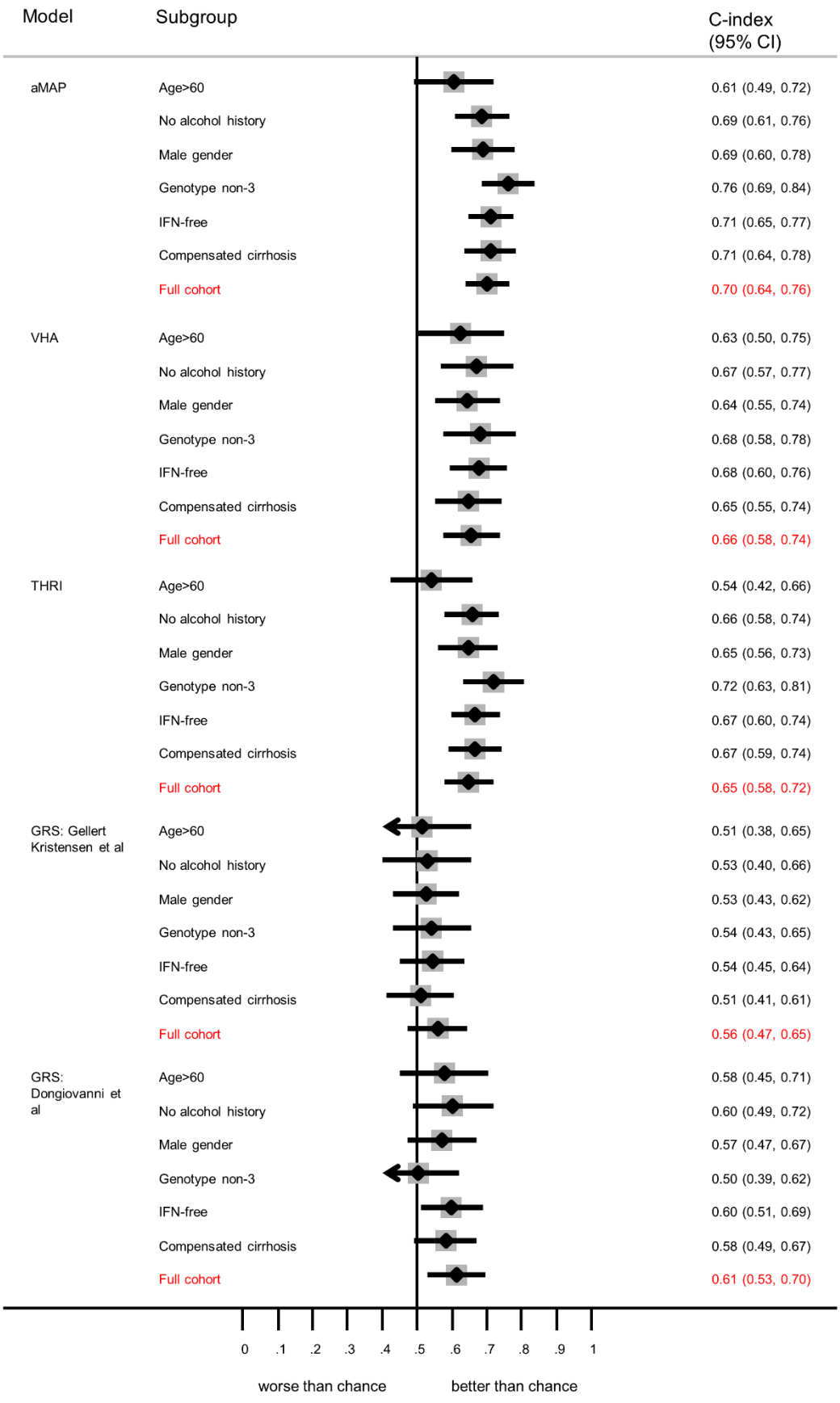
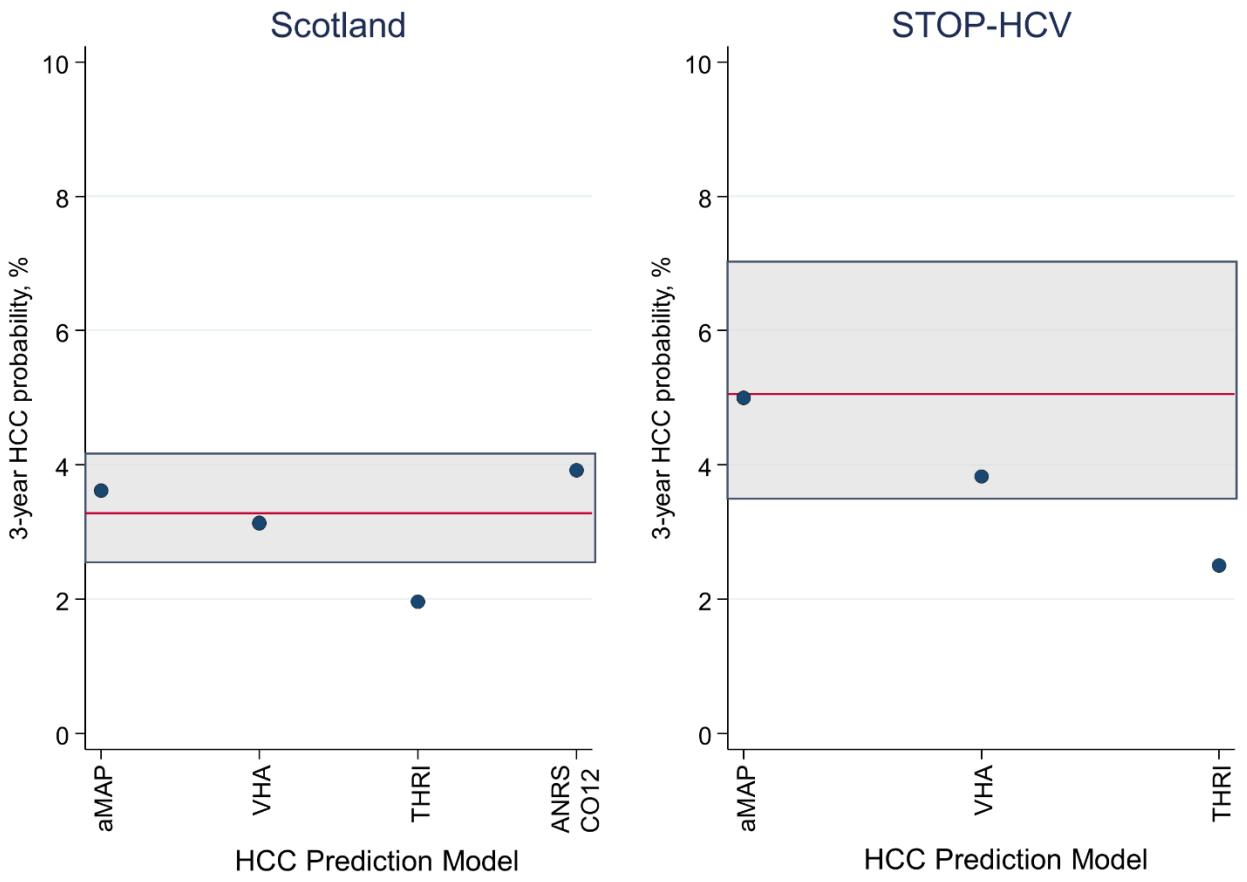


Fig. S6. Observed and predicted 3-year HCC probability in the Scottish and STOP-HCV cohort.



The probability predicted by each model is indicated by the blue circles. The observed 3-year probability of HCC and the 95% uncertainty interval is indicated by the red line and the grey box, respectively.

Table S1: Overview of the non-genetic HCC models assessed in this study.

HCC Model	Year published	Journal	Country	Setting	Characteristics of the development cohort								
					Sample size	Mean age at baseline	HCV GT non-3	Mean duration of follow-up	Number of HCC events	% with cirrhosis	% HCV Aetiology	Prognostic factors selected	Discrimination Performance (Harrell C-index)
aMAP	2020	J hepatol	China	Liver clinic	3688	38 YEARS	NA	3.6 years	95	19%	0% (100% HBV)	age, gender, albumin, bilirubin, platelet count	0.82
VA cirrhosis SVR model	2018	J hepatol	US	Liver clinic	7689	61.5 YEARS	90%	2.5 years	344	100%	100%	Age; race; platelet count; albumin; AST; ALT	0.74‡
Toronto HCC Risk Index (THRI)	2018	J hepatol	Canada	Liver clinic	2079	53.9 YEARS	NK	~6 years**	226	100%	42.5%	Age, etiology, gender, platelet count	0.74
ANRS CO12 CirVir	2017	Hepatology	France	Liver clinic	720	56.2 YEARS	86%	4.3 years	103	100%	100%	age, alcohol, platelet count, GGT, SVR	0.72

Development cohort refers to the dataset used to "train" the model

** this is the mean duration of follow-up for non-HCC cases. The mean follow-up duration for the entire cohort was not indicated in the original paper

‡ Gonen and Heller k-statistic as opposed to Harrell C-index.

Table S2. HCC genetic risk models assessed in this study: overview

Genetic Risk Model	Year published	Journal	Genetic risk variants selected	Method of variant selection	Weighting	Association with incident HCC	C-index
GRS#1 Dongiovanni, et al[1]	2018	Hepatology	rs738409_G; rs58542926_T; rs641738_T; rs1260326_T	Variants selected for known association with steatosis, cirrhosis and HCC	Weighted according to association with liver fat in the Dallas Heart study.	Degasperi et al [2] indicate that, among patients with HCV cirrhosis, individuals with score value >0.597 had more than twice the risk of HCC occurrence (HR:2.30;95% CI: 1.03-5.11; p=0.04) versus individuals with a score<=0.597.	Not reported
GRS#2: Gellert-Kristensen, et al[3]	2020	Hepatology	rs738409_G; rs58542926_T; rs72613567_T	Variants selected for known association with steatosis	Value of 1 assigned to each risk variant - all variants are thus weighted equally	Gellert-Kristensen et al [3] indicate that the score is significantly associated with HCC incidence in participants of the UKBiobank and Copenhagen general population studies. Individuals with a score of 5 or 6 had 29 (95%CI: 17-51) times the odds of HCC versus individuals with a score of zero.	Not reported

[1] Dongiovanni P, Stender S, Pietrelli A, Mancina RM, Cespiati A, Petta S, et al. Causal relationship of hepatic fat with liver damage and inuslin resistance in non-alcoholic fatty liver. Journal of Internal Medicine. 2018;283:356-370.
[2] Degasperi E, Galmozzi E, Pelusi S, D'ambrosio R, Soffredini R, Borghi M. et al. Hepatic fat – genetic risk score predicts hepatocellular carcinoma in HCV cirrhotic patients treated
[3] Gellert-Kristensen H, Richardson TG, Davey-Smith G, Nordestgaard BG, Tybjaerg-Hansen A, Stender S. Combined effecto of PNPLA3, TM6SF2 and HSD17B13 variants on risk of

Table S3: Discriminative ability of HCC prediction models in Scottish and STOPHCV cohorts, in terms of the Concordance index.

	HCC prediction model					
C Index	aMAP	VHA	THRI	ANRS C012	Gellert-Kristensen GRS	Dongiovanni GRS
Scottish cohort						
Wolbers C-index	0.771 (0.731-0.810)	0.715 (0.668-0.761)	0.719 (0.673-0.764)	0.703 (0.656-0.749)	NA	NA
Harrell C-index	0.783 (0.743-0.823)	0.722 (0.674-0.771)	0.723 (0.676-0.770)	0.711 (0.662-0.759)	NA	NA
STOPHCV cohort						
Wolbers C-index	0.701 (0.638-0.764)	0.657 (0.576-0.737)	0.648 (0.577-0.718)	NA	0.559 (0.473-0.645)	0.613 (0.530-0.695)
Harrell C-index	0.708 (0.645-0.771)	0.662 (0.582-0.742)	0.649 (0.579-0.719)	NA	0.560 (0.473-0.648)	0.617 (0.534-0.699)