11-1-2017

A Point System to Forecast Hepatocellular Carcinoma Risk Before and After Treatment Among Persons with Chronic Hepatitis C

Jian Xing

Philip R. Spradling

Anne C. Moorman

Scott D. Holmberg

Eyasu H. Teshale

See next page for additional authors

Recommended Citation
Dig Dis Sci 2017; 62(11):3221-3234.

Authors

Jian Xing, Philip R. Spradling, Anne C. Moorman, Scott D. Holmberg, Eyasu H. Teshale, Loralee B. Rupp, Stuart C. Gordon, Mei Lu, Joseph A. Boscarino, Mark A. Schmidt, Connie M. Trinacty, and Fujie Xu

Dig Dis Sci (2017) 62:3221–3234
DOI 10.1007/s10620-017-4762-0

ORIGINAL ARTICLE

A Point System to Forecast Hepatocellular Carcinoma Risk Before and After Treatment Among Persons with Chronic Hepatitis C

Jian Xing[1] · Philip R. Spradling[1] · Anne C. Moorman[1] · Scott D. Holmberg[1] · Eyasu H. Teshale[1] · Loralee B. Rupp[2] · Stuart C. Gordon[2] · Mei Lu[2] · Joseph A. Boscarino[3] · Mark A. Schmidt[4] · Connie M. Trinacty[5] · Fujie Xu[1] · for the Chronic Hepatitis Cohort Study (CHeCS) Investigators

Received: 24 April 2017 / Accepted: 11 September 2017 / Published online: 30 September 2017

Abstract

Background Risk of hepatocellular carcinoma (HCC) may be difficult to determine in the clinical setting.

Aim Develop a scoring system to forecast HCC risk among patients with chronic hepatitis C.

Methods Using data from the Chronic Hepatitis Cohort Study collected during 2005–2014, we derived HCC risk scores for males and females using an extended Cox model with aspartate aminotransferase-to-platelet ratio index (APRI) as a time-dependent variables and mean Kaplan–Meier survival functions from patient data at two study sites, and used data collected at two separate sites for external validation. For model calibration, we used the Greenwood–Nam–D'Agostino goodness-of-fit statistic to examine differences between predicted and observed risk.

Results Of 12,469 patients (1628 with a history of sustained viral response [SVR]), 504 developed HCC; median follow-up was 6 years. Final predictors in the model included age, alcohol abuse, interferon-based treatment response, and APRI. Point values, ranging from −3 to 14 (males) and −3 to 12 (females), were established using hazard ratios of the predictors aligned with 1-, 3-, and 5-year Kaplan–Meier survival probabilities of HCC. Discriminatory capacity was high (*c*-index 0.82 males and 0.84 females) and external calibration demonstrated no differences between predicted and observed HCC risk for 1-, 3-, and 5-year forecasts among males (all p values >0.97) and for 3- and 5-year risk among females (all p values >0.87).

Conclusion This scoring system, based on age, alcohol abuse history, treatment response, and APRI, can be used to forecast up to a 5-year risk of HCC among hepatitis C patients before and after SVR.

Keywords Hepatocellular carcinoma · Risk · Prediction · Score · Hepatitis C

Abbreviations

CHeCS	Chronic Hepatitis Cohort Study
KM	Kaplan–Meier
APRI	Aspartate aminotransferase-to-platelet ratio index
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
ROC	Receiver operator curve
SVR	Sustained viral response

See Acknowledgments for complete list of investigators.

Disclaimer The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Electronic supplementary material The online version of this article (doi:10.1007/s10620-017-4762-0) contains supplementary material, which is available to authorized users.

✉ Philip R. Spradling
pspradling@cdc.gov

[1] Division of Viral Hepatitis, National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), Centers for Disease Control and Prevention (CDC), Mailstop G37, 1600 Clifton Road NE, Atlanta, GA 30333, USA

[2] Henry Ford Health System, Detroit, MI, USA

[3] Center for Health Research, Geisinger Health System, Danville, PA, USA

[4] The Center for Health Research, Kaiser Permanente-Northwest, Portland, OR, USA

[5] The Center for Health Research, Kaiser Permanente-Hawaii, Honolulu, HI, USA

Introduction

In the United States, hepatocellular carcinoma (HCC) is the fifth and ninth leading cause of cancer-related death among males and females, respectively; most US HCC cases are attributable to infection with hepatitis C virus (HCV) [1, 2]. Overall 5-year survival among patients with HCC is <20% [3]. However, several potentially curative interventions exist for patients with early stage disease; these include liver transplantation, liver resection, and radiofrequency ablation [2].

Identification of HCV-infected patients who should undergo regular surveillance for HCC and the timing for implementation of such surveillance are not always clear. Neither a patient nor provider may be aware of the presence or the point of onset of cirrhosis and, accordingly, the need for HCC surveillance. Ideally, surveillance should target patients at high risk of HCC; surveillance of those at low risk is cost-ineffective and may lead to needless, and potentially harmful, interventions [1]. Among patients in the earlier stages of liver disease, risk of HCC could be periodically assessed and effective treatment could be instituted before patients transit into a phase of higher cancer risk. Clinical guidelines from all major world liver disease societies recommend imaging studies such as ultrasound every 6 months as the standard of care for identifying HCV-infected cirrhotic patients with early stage HCC when the disease may be more amenable to therapy [4–7].

A number of HCC prediction models have been developed, many of which were limited to HCV-infected patients with established cirrhosis. Since the development of many of these models, serum biomarkers such as the aspartate aminotransferase-to-platelet ratio index (APRI) and FIB-4 have been shown to correlate well with biopsy as a means of identifying severe fibrosis and cirrhosis [8, 9]. As these biomarkers are calculated from routinely obtained blood tests, they may be applied to virtually all patients and assessed serially during follow-up.

Older age, low platelet count, and the presence of cirrhosis have been identified as the predominant characteristics that identify HCV-infected patients at risk of HCC, although several other factors may refine the ascertainment of risk (e.g., race, family history, esophageal varices, smoking, and serum albumin, aminotransferase, and alpha fetoprotein levels) [10–16]. Using these factors, some classification systems have been developed to stratify patients into categories of low, intermediate, and high risk, according to estimates of magnitude and timeframes of risk [10–12, 14, 15]. In general, these models were developed using baseline patient characteristics that were fixed and unchanged over time; most apply strictly to patients with untreated HCV infection. It remains challenging, therefore, to determine the evolving degree of HCC risk over the course of follow-up, particularly after eradication of infection. Development of a forecasting tool that accommodates to changing conditions over time, during the course of a patient's follow-up, before and after sustained viral response (SVR), could inform clinician–patient discussions regarding the need to initiate HCC surveillance or, for those never treated, HCV treatment itself [17].

The objective of our study was to develop a simple point system to forecast HCC risk among HCV-infected patients that could be readily applied (and reapplied during follow-up visits) in the clinic. To this end, we developed risk estimates using the same methods employed to develop the Framingham Study risk score functions, which enable clinicians to forecast cardiovascular disease risk using a points-based scoring system derived from the values of several routinely ascertained clinical parameters [18].

Methods

Study Population: Derivation of Development and Validation Cohorts

We used data collected from patients with chronic hepatitis C enrolled in the Chronic Hepatitis Cohort Study (CHeCS), a multicenter observational study whose composition and criteria for inclusion have been summarized previously [19]. These data were accessed via electronic health records and administrative data (supplemented with individual chart review by trained data abstractors) collected during 2005 through 2014 from persons aged ≥18 years at four sites: Geisinger Health System, Danville, PA; Henry Ford Health System, Detroit, MI; Kaiser Permanente-Northwest, Portland, OR; and Kaiser Permanente, Honolulu, Hawaii. Chronic HCV infection was confirmed based on one or more positive HCV RNA tests before antiviral treatment. Data collected included patient demographics, medical encounters, receipt of and response to HCV antiviral therapy, and laboratory and biopsy results. Patients with human immunodeficiency virus or hepatitis B virus coinfection were excluded from the analysis. The parent study protocol was reviewed and approved by an institutional review board at each participating site.

We set out to develop separate scoring systems for males and females because (1) of the disparity in HCC occurrence among males and females, (2) nearly two-thirds of the CHeCS hepatitis C cohort is male, and (3) the Framingham scoring system for cardiovascular disease, upon which our approach was based, uses separate scoring systems for males and females. Accordingly, we separated our pooled cohort according to sex and used two of the

CHeCS study sites (Geisinger Health System and Henry Ford Health System) for model development and the two other CHeCS sites (Kaiser Permanente-Northwest and Hawaii) as the validation cohort.

Selection and Description of Potential HCC Risk Factors

For inclusion in the prediction model, we selected the following independent variables known to be or plausibly associated with HCC: age (as a continuous variable); history of alcohol use disorder (yes/no); history of diabetes mellitus (yes/no); body mass index (<25, 25–29, ≥30 kg/m^2); interferon treatment response (yes/no for never treated, IFN failure, and SVR); and APRI score (yes/no for six time-dependent dummy variables, corresponding to progressively higher score increments). We did not include other predictors (e.g., HCV genotype and estimated duration of HCV infection) for which we lacked complete data.

With regard to fibrosis biomarkers, we chose to use APRI rather than FIB-4 as the latter incorporates age in its calculation. In the event that age and fibrosis score appeared as discrete predictors in the final model, we wanted to ensure that our fibrosis measure did not include patient age.

For each patient, the presence of comorbid diabetes was based on ICD-9 codes for diabetes without chronic complications: 250.0–250.3, 250.7, 250.00–250.03, 250.07, and with chronic complications: 250.4–250.6, 250.04–250.06; alcohol use disorder was based on ICD-9 codes of 291.0–291.9, 303.00–303.03, 303.90–303.93, 305.00–305.03, 980, 291.81, 291.82, or 291.89.

APRI scores were calculated from laboratory data using the following formula [20]:

$$\text{APRI} = \frac{\text{AST(IU/L)/AST upper limit of normal (IU/L)}}{\text{Platelet count } (10^9/\text{L})} \times 100$$

AST and platelet counts had to have been collected within 7 days of each other. We calculated each patient's average APRI values for each calendar year. To minimize the possible effect of temporary fluctuations in the component laboratory values that might have been obtained during acute illness, APRI measurements were excluded if these laboratory results were collected during hospitalization. We also excluded from the analysis patients who were missing data that prevented calculation of a single APRI.

Our main outcome was an occurrence of HCC. We searched for HCC in tumor registry records, according to CDC's National Program of Cancer Registries (for Danville and Portland sites) and the National Cancer Institute's Surveillance, Epidemiology, and End Results program (for Detroit and Honolulu sites) collaborative data collection standards. Such cancers diagnosed during the follow-up period were included as HCC cases in the model.

HCC cases and dates of diagnoses (by calendar year) were ascertained during each patient's follow-up period; those whose diagnosis occurred before their initial APRI score were excluded. Hepatitis C treatment was included in the analysis if the first treatment start date occurred between the date of the initial APRI score and the end of follow-up; persons with a history of SVR achieved before the first APRI score in the dataset were excluded. Among cohort patients who received HCV antiviral treatment, SVR status was based on the availability of one or more negative HCV RNA tests at least 12 weeks post-treatment or on clinician determination abstracted from the patient's medical chart.

Survival Analysis Dataset

Baseline for each patient with an established diagnosis of HCV infection commenced on the date of the first calculable APRI after January 1, 2005. Follow-up continued until the date of HCC diagnosis; for patients in whom HCC did not occur, data were right-censored at the date of death, or the date of the earliest occurring of liver transplant (based on ICD-9 diagnosis or procedure codes of 996.82, 50.5, 50.51, 50.59, 47,135, or 47,136), 2 years following the last APRI measurement, or on December 31, 2014, if none of the aforementioned events occurred.

Prognostic Model Selection and Assessment of Discrimination

We selected our prognostic model using best subsets regression to identify predictors of importance and calculated the Harrell's *c*-index [21–23], a measure of the model's discriminatory accuracy (i.e., the model's capacity to distinguish subjects at high vs. low risk of HCC) that is analogous to the area under the receiver operator characteristic (ROC) curve [23, 24]. A *c*-index of 0.5 indicates no discriminatory accuracy, and an index of 1.0 represents perfect discrimination. To examine whether the *c*-index required adjustment, we performed bootstrapping with replacement using 80% of original dataset on 100 subdatasets of the same sample size generated from the original dataset.

We estimated the optimism of our prognostic model selection and validation, as suggested by Harrell [25], with the macros "Harrell_Optimism_Cox" [21] and "SURVCSTD" [22]. Initially including all predictors (i.e., age, alcohol use disorder history, diabetes history, history of HCV antiviral treatment, body mass index, and APRI) in

Table 1 Characteristics of hepatitis C virus-infected patients with and without hepatocellular carcinoma, Chronic Hepatitis Cohort Study, 2005–2014

Variables	Male				Female			
	Total $N = 7372$	Development ($n = 4741$)	Validation ($n = 2631$)	p value	Total $N = 5097$	Development ($n = 3413$)	Validation ($n = 1684$)	p value
Duration of follow-up								
N	7372	4741	2631		5097	3413	1684	
Median	6.3	5.9	6.9		6.5	6.1	7.3	
Range	0.5–10.0	0.5–10.0	0.5–10.0		0.5–10.0	0.5–10.0	0.5–10.0	
Mean (SE)	6.2 (0.03)	6.0 (0.04)	6.7 (0.06)	<0.001	6.4 (0.04)	6.2 (0.05)	6.9 (0.07)	<0.001
Person-year follow-up								
Total		28,346.7	17,506.3			21,016.2	11,615.2	
Post-SVR		2790.6	2454.0			2451.0	1710.2	
Number of clinic visits[a] per person-year during follow-up								
N	7146	4546	2600		4968	3302	1666	
Median	6.2	5.3	7.3		6.8	5.9	8.4	
Range	0.1–380.1	0.1–120.9	0.3–380.1		0.1–368.9	0.1–123.9	0.2–368.8	
Mean (SE)	10.4 (0.2)	9.5 (0.2)	12.1 (0.4)	<0.001	11.2 (0.3)	9.5 (0.2)	14.5 (0.7)	<0.001
Age (years) at baseline								
N	7372	4741	2631		5097	3413	1684	
Median	53	53	53		51	51	51	
Range	14–95	15–95	14–91		16–90	16–90	17–89	
Mean (SE)	52 (0.1)	51 (0.2)	52 (0.2)	0.05	49 (0.2)	48 (0.2)	51 (0.2)	<0.001
HCC								
No	6985 (94.8)	4492 (94.7)	2493 (94.8)	1.000	4980 (97.7)	3334 (97.7)	1646 (97.7)	1.000
Yes	387 (5.2)	249 (5.3)	138 (5.2)		117 (2.3)	79 (2.3)	38 (2.3)	
HCC post-SVR	33 (0.4)	15 (0.3)	18 (0.7)	0.04	10 (0.2)	6 (0.2)	4 (0.2)	0.15
HCC post-IFN failure	127 (1.7)	91 (1.9)	36 (1.4)		43 (0.8)	24 (0.7)	19 (1.1)	
HCC never treated	227 (3.1)	143 (3.0)	84 (3.2)		64 (1.3)	49 (1.4)	15 (0.9)	
SVR status								
SVR	914 (12.4)	529 (11.2)	385 (14.6)	<0.001	714 (14.0)	445 (13.0)	269 (16.0)	<0.001
IFN failure	1071 (14.5)	748 (15.8)	323 (12.3)		665 (13.0)	493 (14.4)	172 (10.2)	
Never treated	5387 (73.1)	3464 (73.1)	1923 (73.1)		3718 (72.9)	2475 (72.5)	1243 (73.8)	
Race								
White	4669 (63.3)	2832 (59.7)	1837 (69.8)	<0.001	3308 (64.9)	2130 (62.4)	1178 (70.0)	<0.001
Black	1639 (22.2)	1523 (32.1)	116 (4.4)		1114 (21.9)	1032 (30.2)	82 (4.9)	
Asian	200 (2.7)	56 (1.2)	144 (5.5)		188 (3.7)	52 (1.5)	136 (8.1)	
Other/ unknown	864 (11.7)	330 (7.0)	534 (20.3)		487 (9.6)	199 (5.8)	288 (17.1)	
Insurance (93 missing)								
Medicaid	730 (10.0)	511 (10.8)	219 (8.6)	<0.001	744 (14.8)	538 (15.8)	206 (12.7)	<0.001
Medicare only	1517 (20.8)	1517 (32.0)			1023 (20.3)	1023 (30.0)		
Medicare plus	854 (11.7)	263 (5.5)	591 (23.3)		500 (9.9)	157 (4.6)	343 (21.2)	
Private	3802 (52.2)	2074 (43.7)	1728 (68.1)		2342 (46.5)	1272 (37.3)	1070 (66.1)	
None	376 (5.2)	376 (7.9)			423 (8.4)	423 (12.4)		

Table 1 continued

Variables	Male				Female			
	Total $N = 7372$	Development ($n = 4741$)	Validation ($n = 2631$)	p value	Total $N = 5097$	Development ($n = 3413$)	Validation ($n = 1684$)	p value
BMI (kg/m^2)								
<25	2584 (35.1)	1924 (40.6)	660 (25.1)	<0.001	2110 (41.4)	1502 (44.0)	608 (36.1)	<0.001
25–30	2547 (34.5)	1447 (30.5)	1100 (41.8)		1365 (26.8)	875 (25.6)	490 (29.1)	
>30	2241 (30.4)	1370 (28.9)	871 (33.1)		1622 (31.8)	1036 (30.4)	586 (34.8)	
Diabetes[b]								
No	6712 (91.0)	4316 (91.0)	2396 (91.1)	1.000	4719 (92.6)	3146 (92.2)	1573 (93.4)	0.12
Yes	660 (9.0)	425 (9.0)	235 (8.9)		378 (7.4)	267 (7.8)	111 (6.6)	
Alcohol abuse[b]								
No	5218 (70.8)	3459 (73.0)	1759 (66.9)	<0.001	4210 (82.6)	2901 (85.0)	1309 (77.7)	<0.001
Yes	2154 (29.2)	1282 (27.0)	872 (33.1)		887 (17.4)	512 (15.0)	375 (22.3)	
Baseline APRI (all patients)								
<1.0	4664 (63.3)	3013 (63.6)	1651 (62.8)	0.02	3664 (71.9)	2398 (70.3)	1266 (75.2)	<0.001
1.0–2.0	1353 (18.4)	829 (17.5)	524 (19.9)		755 (14.8)	520 (15.2)	235 (14.0)	
>2.0	1355 (18.4)	899 (19.0)	456 (17.3)		678 (13.3)	495 (14.5)	183 (10.9)	
APRI at SVR								
N	898	513	385		702	434	268	
<1.0	748 (83.3)	427 (83.2)	321 (83.4)		597 (85.0)	375 (86.4)	222 (82.8)	
1.0–2.0	111 (12.4)	62 (12.1)	49 (12.7)		67 (9.5)	34 (7.8)	33 (12.3)	
>2.0	39 (4.3)	24 (4.7)	15 (3.9)	0.825	38 (5.4)	25 (5.8)	13 (4.9)	0.1364
Last APRI (all patients)								
N	7372	4741	2631		5097	3413	1684	
<1.0	4777 (64.8)	3072 (64.8)	1705 (64.8)		3651 (71.6)	2411 (70.6)	1240 (73.6)	
1.0–2.0	1280 (17.4)	839 (17.7)	441 (16.8)		760 (14.9)	525 (15.4)	235 (14.0)	
>2.0	1315 (17.8)	830 (17.5)	485 (18.4)	0.4344	686 (13.5)	477 (14.0)	209 (12.4)	0.0820
Follow-up post-SVR (years)								
N	898	513	385		702	434	268	
Median	5.77	5.00	6.60		6.08	5.63	6.76	
Range	0.79–10.00	0.79–10.00	0.80–10.00		0.76–10.00	0.81–10.00	0.76–10.00	
Mean (SE)	5.84 (0.10)	5.44 (0.13)	6.37 (0.15)	<0.001	5.93 (0.11)	5.65 (0.14)	6.38 (0.18)	0.003
Baseline HCV RNA level								
N	5800	4087	1713		4006	2928	1078	
<500,000	1924 (33.2)	1426 (34.9)	498 (29.1)		1587 (39.6)	1293 (44.2)	294 (27.3)	
≥500,000	3876 (66.8)	2661 (65.1)	1215 (70.9)	<0.001	2419 (60.4)	1635 (55.8)	784 (72.7)	<0.001
Baseline ALT[c]								
N	7273	4662	2611		5036	3368	1668	
Normal	1061 (14.6)	762 (16.3)	299 (11.5)		448 (8.9)	313 (9.3)	135 (8.1)	
1–2× ULN[b]	2100 (28.9)	1438 (30.8)	662 (25.4)		1231 (24.4)	832 (24.7)	399 (23.9)	
>2× ULN	4112 (56.5)	2462 (52.8)	1650 (63.2)	<0.001	3357 (66.7)	2223 (66.0)	1134 (68.0)	0.2539
Baseline platelet count[d]								
N	7266	4657	2609		5040	3369	1671	
Normal	5223 (71.9)	3256 (69.9)	1967 (75.4)		4098 (81.3)	2688 (79.8)	1410 (84.4)	
<Normal	2043 (28.1)	1401 (30.1)	642 (24.6)	<0.001	942 (18.7)	681 (20.2)	261 (15.6)	<0.001
Liver biopsy								
No								

Table 1 continued

Variables	Male				Female			
	Total $N = 7372$	Development ($n = 4741$)	Validation ($n = 2631$)	p value	Total $N = 5097$	Development ($n = 3413$)	Validation ($n = 1684$)	p value
Yes (METAVIR stage)								
F0–F1	465 (33.4)	343 (47.7)	122 (18.2)		424 (44.4)	294 (57.8)	130 (29.1)	
F2–F3	615 (44.2)	211 (29.3)	404 (60.1)		387 (40.5)	130 (25.5)	257 (57.5)	
F4	311 (22.4)	165 (22.9)	146 (21.7)	<0.001	145 (15.2)	85 (16.7)	60 (13.4)	<0.001

ALT alanine aminotransferase, *APRI* aspartate aminotransferase-to-platelet ratio index, *BMI* body mass index, *HCC* hepatocellular carcinoma, *SVR* sustained virologic response at 12 weeks, *ULN* upper limit of normal

[a]Clinic visits include ambulatory visit or emergency department visit (excludes urgent care visit)

[b]Diabetes mellitus and alcohol misuse/dependence according to ICD-9 code

[c]ALT upper limit of normal: 30 IU/mL for men, 19 IU/mL for women

[d]Platelet count lower limit of normal cutoff: 150,000 cells/mL

the original dataset (N), we generated $M = 100$ datasets of the same sample size n from the development cohort dataset size N, where $n < N$, using bootstrap samples with replacement. For each one of the new datasets $m = 1,\ldots,M$, the same algorithmic approach as above was used to calculate the bootstrap Harrell's c-index ($c^{(m)}_{\text{boot}}$). For each one of the 100 new models, we calculated its Harrell's c-index, as applied to the original dataset (N) ($c^{(m)}_{\text{orig}}$). For all bootstrap samples, the optimism in the fit was $O^{(m)} = c^{(m)}_{\text{boot}} - c^{(m)}_{\text{orig}}$. The average of these values was the optimism of the original model:

$$\bar{O} = \frac{\sum_{m=1}^{M} O^{(m)}}{M}$$

The optimism-corrected performance of the final model was $c_{\text{adj}} = c_{\text{app}} - \bar{O}$. From that model, the apparent Harrell's c-index was calculated.

Development of Scoring System for HCC Risk

We developed our HCC risk estimates with the same methods used to develop the Framingham Study risk score functions, performing a series of steps as suggested by Sullivan [18]. This process first involved computing the mean Kaplan–Meier (KM) survival function over 1, 3, and 5 years of follow-up and then computing an extended Cox model (i.e., with APRI as a time-dependent variable) with regression coefficients for the predictors selected and validated from the previous sections. An extended Cox model was preferred over a standard model because we hypothesized that the APRI trajectory might differ before (i.e., progressively higher scores) and after (i.e., progressively lower scores) SVR. In order to derive time-dependent regression coefficients for APRI, for each patient we calculated mean annual (calendar year) APRI values from 2005 to 2014. Patients did not necessarily undergo the same frequency and timing of laboratory assessments (to calculate APRI score), although nearly all patients had at least one APRI score available per year of follow-up. Mean annual values derived from a single APRI score were accepted; those without at least one APRI for two consecutive years were right-censored at the date of the last available APRI score. We then built a matrix for all participants vis-à-vis their specific yearly mean APRI scores in a dataset. For those who had HCC, we marked a timestamp in the year of event (HCC diagnosis) and right-censored all yearly mean APRI after that year. Meanwhile, for those who did not have HCC, we used all mean annual APRI scores and right-censored at death, departure from the CHeCS, or the end of the study period follow-up in 2014 (whichever came first). Finally, we applied this dataset to survival analysis using SAS PROC PHREG. The mean annual APRI scores were applied as Heaviside functions in the model, i.e., the value of yearly APRI was assigned to the model one at a time for each specific year.

The regression coefficients for each predictor variable were then weighted, according to their mean or proportion distribution in the model development cohort. Each predictor variable was divided into categories; for age, we used 5-year segments and assigned the lowest age in each segment as the comparative age against the main age referent (e.g., age 20 years was used as the main referent for males and females), for the treatment variable (IFN-based treatment response, as the data were collected before release of the second-generation direct acting antiviral drugs) we used "never treated" as the referent (vs. IFN failure and SVR), and for the remaining variables (all yes/no dummy variables) we used "no" as the referent. We then calculated the degree of increased risk of HCC relative to the referent for each category and assigned point values, accordingly, for all category segments. For each variable, we determined its minimum and maximum point value

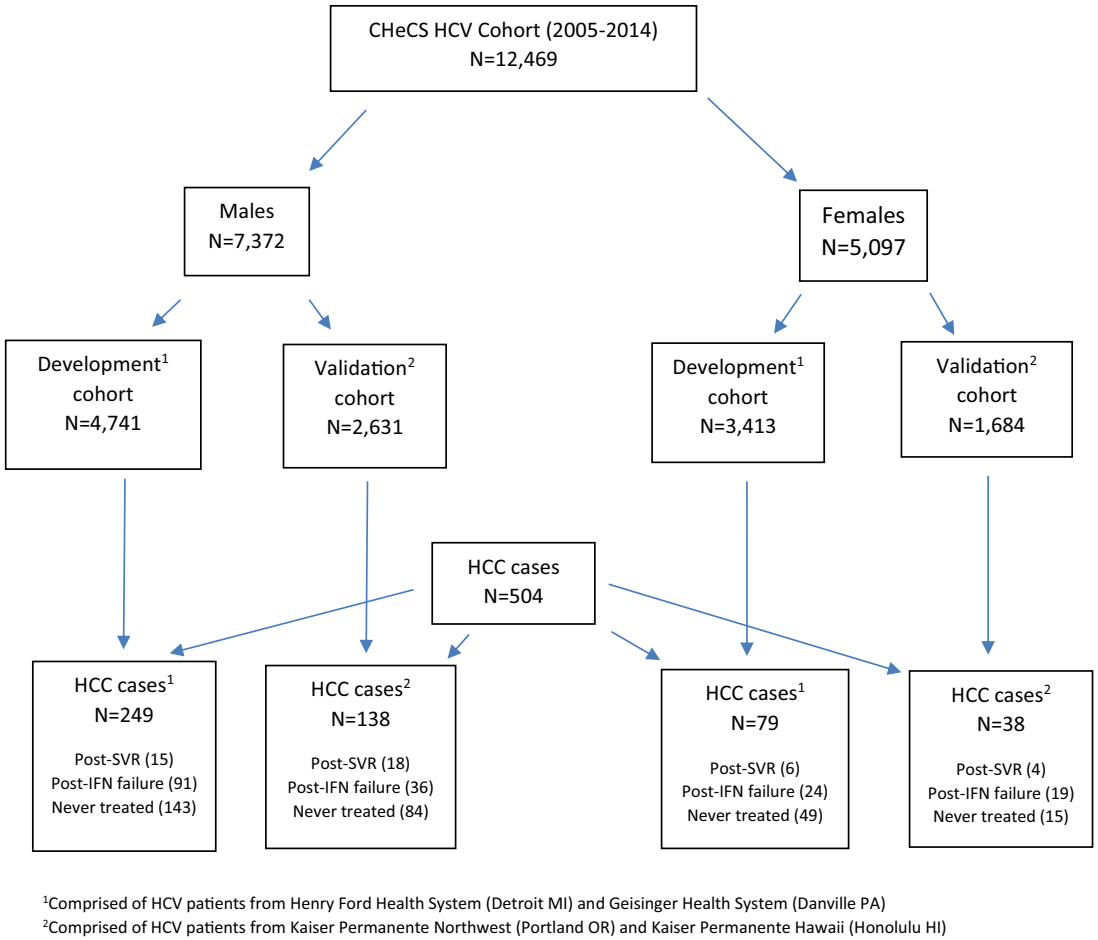

Fig. 1 Derivation of the model development and validation cohorts, Chronic Hepatitis Cohort Study

contributions, the sums of which provided a range of total point values for all predictor variables in the model. For each possible total point value, we calculated 1-, 3-, and 5-year risk estimates of HCC and represented these in condensed tables for clinical use, separately for males and females (see Supporting Information for detailed methods).

Model Calibration

We applied the Greenwood–Nam–D'Agostino goodness-of-fit statistic as a test of calibration. The Greenwood–Nam–D'Agostino statistic is a modification of the Nam–D'Agostino goodness-of-fit test developed for longitudinal survival models and serves a purpose similar to the Hosmer–Lemeshow goodness-of-fit test used for logistic regression models [26, 27]. We calibrated our prediction score model according to the 1, 3-, and 5-year risk estimates of HCC by calculating the prediction score total point value for each patient and obtaining the corresponding predicted probability of HCC. We then categorized total points into groups (four groups for males and three for females, based on the distribution of point scores) and calculated the mean predicted probability of HCC for each group. Lastly, we calculated observed HCC risk with the KM method for each group, determining the variance of each and the corresponding Chi-square and p values for the 1-, 3-, and 5-year risk estimates (see Supporting Information for detailed methods).

All statistical analyses were conducted using SAS 9.3 (Cary, NC).

Results

Characteristics of the Development Cohort and Validation Cohorts

Table 1 shows characteristics of HCV-infected patients with and without HCC during 2005–2014, according to sex and representation in the development and validation cohorts; the derivation of these cohorts from the larger CHeCS hepatitis C cohort is shown in Fig. 1. There were

Table 2 An example of the best subset models selected in first 10 bootstrap samples from total 100 bootstraps used to determine the overall model *c*-index (example using male cohort)

1 Sample	2 Variables[a] in best subset model	3 Number of variables in best subset model	4 Harrell's *c*-index from best subset model fit to bootstrap dataset (80% sampling)	5 Harrell's *c*-index from best subset model fit to bootstrap dataset, applied to original dataset (100% sampling)	6 Difference (i.e., optimism)
1	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap3 Ap2	7	0.8099	0.8155	−0.005559
2	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap3 Ap2	7	0.8145	0.8232	−0.008666
3	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap4 Ap3 Ap2	8	0.8158	0.8366	−0.020778
4	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap3 Ap2	7	0.8102	0.8215	−0.011238
5	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap2	6	0.7864	0.8126	−0.026223
6	Age1 Bmi3 AlcoAbuse Svr2 Ap6 Ap5 Ap3 Ap2	8	0.8044	0.8162	−0.011829
7	Age1 AlcoAbuse Svr2 Ap6 Ap3 Ap2	6	0.8087	0.8172	−0.008450
8	Age1 AlcoAbuse Svr2 Ap6 Ap5 Ap3 Ap2	7	0.7871	0.8275	−0.040415
9	Age1 AlcoAbuse Svr2 Ap6 Ap3 Ap2	6	0.7823	0.8160	−0.033685
10	Age1 AlcoAbuse Ap6 Ap5 Ap4 Ap3 Ap2	7	0.7959	0.8332	−0.037272

Overall *c*-index = Average of Column 4 − Average of Column 6 = 0.7956 − (−0.0210) = 0.8166

[a]For inclusion in the prediction model, we selected the following independent variables known to be or plausibly associated with HCC: age (as a continuous variable); history of alcohol use disorder (yes/no); history of diabetes mellitus (yes/no); body mass index (<25, 25–29, ≥30 kg/m^2—corresponding to Bmi1–3 above); interferon treatment response (yes/no for never treated, IFN failure, and SVR, corresponding to Svr1-3 above); and APRI score (yes/no for six time-dependent dummy variables, corresponding to progressively higher score increments, corresponding to Ap1-6 above)

12,469 patients in the pooled cohort: 7372 were male and 5097 were female, which included 914 (12.4%) men and 714 (14.0%) women who achieved SVR during follow-up. During follow-up, 504 patients developed HCC, of whom 387 were male (5.2% of all males) and 117 were female (2.3% of all females). Among male HCC cases, 227 occurred among patients never treated, 127 occurred after IFN treatment failure, and 33 occurred after SVR; among female cases, 64 occurred among patients never treated, 43 occurred following IFN failure, and 10 occurred after SVR. Both male and female cohorts were comprised predominantly of persons who were white, had private healthcare insurance, and had BMI ≥25 kg/m^2; 9.0% of males and 7.4% of females had diabetes, and 29.2% of males and 17.4% of females had a history of an alcohol use disorder. The median age at initiation of follow-up (i.e., first APRI score) was 53 years for males and 51 years for females; median duration of follow-up was 6.3 years for males and 6.5 years for females. At the initiation of follow-up, 63.3% of males and 71.9% of females had an APRI <1.0 and 18.4% of males and 13.3% of females had an APRI >2.0. The median age at HCC diagnosis was 58 (range 35–65) years for males and 60 (range 46–75) years for females. Among patients who had achieved SVR, the median APRI at SVR was 0.43 for men and 0.36 for women; the median duration of post-SVR follow-up was 5.0 years for men and

Table 3 Extended Cox hazards regression coefficients from the final model, by sex

Risk factor	Male			Female		
	Regression coefficient	Hazard ratio	*p* value	Regression coefficient	Hazard ratio	*p* value
Age	0.05	1.05	<0.001	0.06	1.07	<0.001
Alcohol abuse	0.62	1.87	<0.001	0.60	1.82	0.03
Never treated	0.17	1.18	0.54	−0.01	0.99	0.99
IFN failure	1.01	2.74	<0.001	0.48	1.62	0.30
SVR	0	1	–	0	1	–
APRI 0–1.0	0	1	–	0	1	–
APRI 1.0–2.0	1.15	3.15	<0.001	1.69	5.39	<0.001
APRI 2.0–3.0	1.20	3.31	<0.001	2.12	8.35	<0.001
APRI 3.0–4.0	1.23	3.43	<0.001	1.99	7.28	<0.001
APRI 4.0–5.0	1.56	4.77	<0.001	2.30	9.98	<0.001
APRI >5.0	2.04	7.70	<0.001	2.60	13.44	<0.001

6.1 years for females. Table 1 further demonstrates that there were significant differences between the development and validation cohorts among males and females for nearly all variables (i.e., race, insurance status, BMI, alcohol abuse history, baseline APRI, age at first APRI, IFN response, duration of total follow-up, and frequency and duration of post-SVR follow-up among patients successfully treated).

Model Selection and Validation (Discrimination)

In the final model, we included age, history of alcohol use disorder, IFN treatment response, and APRI score for males and females (additional details are provided with Supporting Information, Supplementary Table 1a [males] and 1b [females], and Supplementary Table 2). The *c*-index, which corresponds to the area under the ROC curve, for males was 0.82 (the derivation of which is shown in Table 2) and for females was 0.84, consistent with good discriminatory capacity.

Development of HCC Risk Scoring System, by Sex

We computed an extended Cox model with regression coefficients for age, alcohol use disorder, IFN treatment response, and APRI (as a time-dependent variable), according to sex, as well as the mean KM survival function over 1, 3, and 5 years of follow-up. Extended Cox hazard regression coefficients for the variables in the final model are shown in Table 3. The minimum and maximum point value contributions for each variable and the range of total point values are shown in 4a (males) and 4b (females). For each possible total point value, we calculated 1-, 3-, and 5-year risk estimates for HCC and represented these in condensed tables for clinical use, one for males and another for females (Table 5a, b). Prediction scores of 6 for males corresponded to a 1.5% and scores of 8 for females corresponded to a 1.8% annual risk of HCC; the commensurate 3- and 5-year risk estimates were, respectively, 2.3 and 2.9% for males, and 2.7 and 3.4% for females. Tables 4 and 5 should be applied for direct clinical use (additional details are provided with Supporting Information, Supplementary Tables 3a, 3b, 4a, and 4b).

Model Calibration

The results of external calibration, demonstrating differences between the score-predicted and KM-observed HCC risk, for 1-, 3-, and 5-year forecasts, are shown in Fig. 2a (males) and 2b (females). There were no significant differences between predicted and observed HCC risk for 1-, 3-, and 5-year forecasts among males ($p = 0.977$, 0.990, and 0.996, respectively) and for 3- and 5-year risk among females ($p = 0.873$ and 0.992, respectively). The Greenwood–Nam–D'Agostino goodness-of-fit statistic could not be applied to calibrate 1-year HCC risk among females because there were an insufficient number of observed cases to provide a comparison prediction score group.

In general, the predicted risk derived from scores exceeded the KM-observed risk, although the absolute difference between observed and predicted values decreased with progressively longer forecast intervals and with higher prediction scores [Figs. 2a, b; additional details available in Supporting Information, Supplementary Table 5a (males) and 5b (females)].

Table 4 Model-derived point values for the clinical scoring system for 1-, 3-, and 5-year estimates of hepatocellular carcinoma risk among patients with chronic hepatitis C, according to risk factors and categories, for males (a) and females (b)

Score components		
Risk factor	Categories	Points
a. Male		
Age (years)	20–24	−3
	25–29	−2
	30–34	−1
	35–39	0
	40–44	1
	45–49	2
	50–54	3
	55–59	4
	60–64	5
	65–69	6
	70–74	7
	75–79	8
	80–84	9
	>85	10
Alcohol use disorder	No	0
	Yes	2
Treatment history	Never treated	1
	IFN failure	4
	SVR	0
APRI	0–1.0	0
	>1.0–2.0	4
	>2.0–4.0	5
	>4.0–5.0	6
	>5.0	8
b. Female		
Age	20–24	−3
	25–29	−2
	30–34	−1
	35–39	0
	40–44	1
	45–49	2
	50–54	3
	55–59	4
	60–64	5
	65–69	6
	70–74	7
	75–79	8
	80–84	9
	≥85	10
Alcohol use disorder	No	0
	Yes	2
Treatment history	Never treated	0
	IFN failure	2
	SVR	0
APRI	0–1.0	0
	>1.0–2.0	5
	>2.0–4.0	6
	>4.0–5.0	7
	>5.0	8

Discussion

Using data collected from patients with chronic hepatitis C during 2005–2014 at two large US healthcare organizations (Henry Ford Health System, Detroit, MI and Geisinger Health System, Danville, PA), we developed a model and HCC prediction scoring system for males and females, incorporating age, treatment response, alcohol use disorder history, and APRI score. We externally validated this model using 2005–2014 data from patients with hepatitis C at two other study sites (Kaiser Permanente, Portland, OR, and Kaiser Permanente, Honolulu, HI), whose demographic and clinical characteristics differed significantly from those of patients from the model development sites. Validation of the model consisted of assessments of discrimination and calibration. Its capacity for discrimination between HCV-infected patients at high versus low risk of HCC was high, with a *c*-index >0.8 for both sexes. The Greenwood–Nam–D'Agostino goodness-of-fit statistic for calibration demonstrated overall good predictive capacity, though the model was superior for predicting 3- and 5-year HCC risk than for predicting 1-year risk. Its performance was also better for males than for females (largely because of higher HCC case numbers for males) and at higher rather than lower total point scores. As we used pre- and post-SVR longitudinal clinical data for model development [28] and incorporated APRI score as a time-dependent variable in an extended Cox model (rather than a baseline score in a standard Cox model), this scoring system may be readily applied and reapplied in any clinical setting during the course of follow-up as a patient's age and APRI score change, before and after achievement of SVR. Thus, the score can be individualized to patients at any point during follow-up and can provide an immediate forecast of HCC risk according to evolving risk determinants over time.

For example, a 52-year-old, treatment-naïve male with no alcohol abuse history and an APRI score of 0.9 would have, using Table 4, a risk score of 4, which corresponds in Table 5 to 1-, 3-, and 5-year risk of HCC of 0.9, 1.4, and 1.7% in Table 5. If the same patient were seen 6 years later, had remained untreated, and had an APRI score of

Table 5 Point totals with corresponding 1-, 3-, and 5-year estimates of hepatocellular carcinoma risk among patients with chronic hepatitis C, for males (a) and females (b)

Total points	HCC risk (%)		
	1 year	3 years	5 years
a. Male			
−3	0.1	0.2	0.3
−2	0.2	0.3	0.4
−1	0.2	0.4	0.5
0	0.3	0.5	0.6
1	0.4	0.6	0.8
2	0.5	0.8	1.0
3	0.7	1.0	1.3
4	0.9	1.4	1.7
5	1.2	1.8	2.2
6	1.5	2.3	2.9
7	2.0	3.0	3.8
8	≥2.5[a]	3.8	4.9
9		5.0	6.3
10		6.4	8.1
11		≥8.3[a]	10.4
12			13.3
13			17.0
14			≥21.5
b. Female			
−3	0.1	0.1	0.1
−2	0.1	0.1	0.1
−1	0.1	0.1	0.2
0	0.1	0.2	0.3
1	0.2	0.3	0.4
2	0.3	0.4	0.5
3	0.4	0.5	0.7
4	0.5	0.7	1.0
5	0.7	1.0	1.3
6	1.0	1.4	1.8
7	1.3	1.9	2.5
8	≥1.8[a]	2.7	3.4
9		≥3.6[a]	4.7
10			6.4
11			8.7
12			≥11.8

[a]Since there were so few individuals in the upper ranges of the distribution, we cut off the risk table to avoid overstating the precision in the risk estimates

1.8, his total score from Table 4 at this point would be 9, corresponding in Table 5 to 1-, 3-, and 5-year HCC risk of >2.5, 5.0, and 6.3%, respectively.

The importance of patient age with regard to HCC risk is evident in our scoring system. In our system, every 5-year increment in advancing age translated to an additional point added to the point total (e.g., a male aged 55–59 years would receive 4 points and one aged 60–64 years would receive 5 points). Patient age has been identified as a major risk factor for HCC among HCV-infected persons, and investigators have found that even when stratified by stage of fibrosis (except among those with cirrhosis), the incidence of HCC was higher among older versus younger patients, and that fibrosis progression accelerated after age 50 years regardless of duration of infection [29–35].

Interferon treatment failure was an important characteristic of the model, a history of which contributed four points for men and two for women to their point totals. That interferon treatment failure was associated with a more rapid progression of liver disease (compared to untreated persons) was consistent with what we reported, using separate methodology, in another publication [36], and is consistent with the findings of Baran et al. [37].

There are a few issues with the application of the scoring system that could be encountered in clinical practice. As we used mean calendar year APRI scores for model development, if serial values are available, we suggest that clinicians use a patient's mean APRI score over a period rather than rely on a single value. Characterization of alcohol use disorder for model development was based on ICD-9 codes, which been shown to underestimate the prevalence of alcohol abuse and dependence when compared with the prevalence ascertained through direct patient questioning. For example, our group found ICD-9 codes to be insensitive but highly specific (i.e., >90%) for detecting alcohol use disorders [38]. With respect to use of the scoring system in a clinical setting, providers should have the discretion of adding points for alcohol use disorder based on knowledge of or a high index of clinical suspicion of such a disorder, independent of whether a patient had an ICD-9 code-based diagnosis in his or her medical record.

There are limitations to our model and scoring system. Application of the system should be limited to HCV monoinfected persons, as persons with known HBV or HIV infection were excluded from the development cohort. As mentioned, data used for development and validation of the model were largely collected during the IFN treatment era, so it is possible the system may not entirely reflect conditions and outcomes related to widespread use of DAAs. Since we had a relatively low number of genotype 3 patients among patients with known genotype, and because genotype data were missing for a large fraction of patients who were never treated, we were unable to assess whether genotype might serve as a predictor variable (e.g., whether patients with genotype 3 infection might accrue more points compared to non-genotype 3 patients). Although we did not include nonalcoholic fatty liver disease specifically

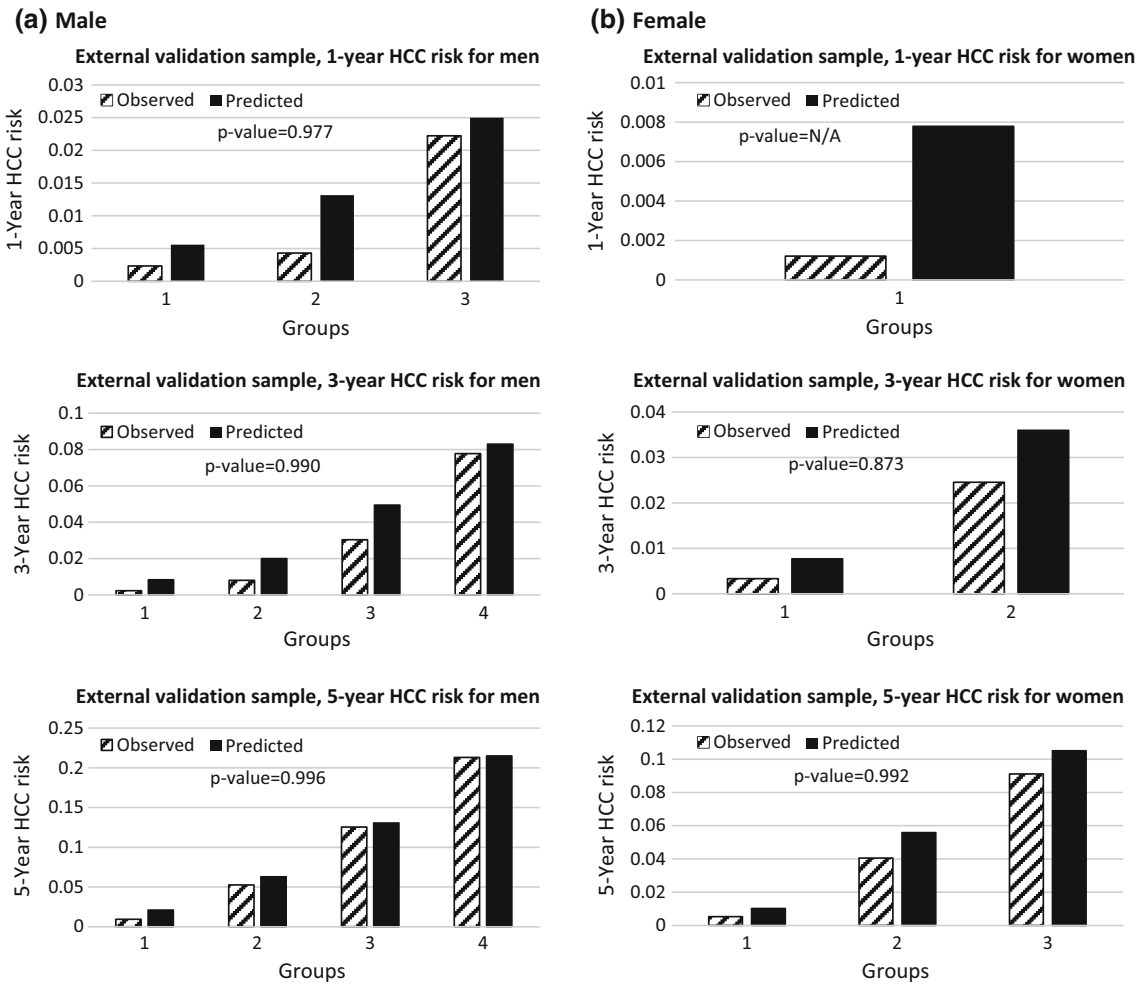

Fig. 2 **a**, **b** Calibration bar grafts comparing predicted estimate of hepatocellular carcinoma (HCC) risk (derived from the development cohort) versus observed HCC cases in the external validation cohort, for males (**a**) and females (**b**)

as a predictor variable, we included BMI and DM as candidate variables in the prediction model, neither of which was determined to belong in the final model. Nonetheless, we believe this simple scoring system for estimating HCC risk can serve as a useful tool to guide decision-making for patients and clinicians. Validation with other pre- and post-SVR cohorts, particularly as more data collected from the DAA era are available, may help further refine and improve its predictive capacity. Given the expanding population of persons in whom HCV infection has been eradicated by the use of DAAs, it will be necessary to determine in the future whether and to what degree these risk estimates should be modified.

Acknowledgments The CHeCS Investigators include the following investigators and sites: Scott D. Holmberg, Eyasu H. Teshale, Philip R. Spradling, Anne C. Moorman, Fujie Xu, Jim Xing, and Yuna Zhong, Division of Viral Hepatitis, National Centers for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), Centers for Disease Control and Prevention (CDC), Atlanta, Georgia; Stuart C. Gordon, David R. Nerenz, Mei Lu, Lois Lamerato, Jia Li, Loralee B. Rupp, Nonna Akkerman, Nancy Oja-Tebbe, Yueren Zhou, and Talan Zhang, Henry Ford Health System, Detroit, Michigan; Joseph A. Boscarino, Zahra S. Daar, Robert E. Smith and Meredith Lewis, Center for Health Research, Geisinger Health System, Danville, Pennsylvania; Connie Mah Trinacty, Yihe G. Daida, and Carmen P. Wong, The Center for Health Research, Kaiser Permanente, Hawaii, Honolulu, Hawaii; Mark A. Schmidt, Judy L. Donald, and Erin M. Keast, the Center for Health Research, Kaiser Permanente-Northwest, Portland, OR.

Funding Henry Ford Health System receives funding for CHeCS from the Centers for Disease Control and Prevention and from Gilead Sciences. CHeCS was previously funded through May 2016 by the CDC Foundation, which received grants from AbbVie; Genentech, a Member of the Roche Group; Gilead Sciences; Janssen Pharmaceuticals, Inc. and Vertex Pharmaceuticals; past partial funders include Bristol-Myers Squibb. Granting corporations do not have access to CHeCS data and do not contribute to data analysis or writing of manuscripts.

Author's contribution JX carried out study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript, and statistical analysis; PS contributed to study concept and design, analysis and interpretation of data, drafting

of the manuscript, critical revision of the manuscript, and study supervision; AM was responsible for study concept and design analysis and interpretation of data, critical revision of the manuscript, and technical support; SH and ET were involved in analysis and interpretation of data, critical revision of the manuscript, technical support, and study supervision; LR, JB, MS, and CT took part in acquisition of data, analysis and interpretation of data, critical revision of the manuscript, and technical support; SG and ML performed analysis and interpretation of data, critical revision of the manuscript, and technical support; FX participated in study concept and design, analysis and interpretation of data, critical revision of the manuscript, statistical analysis, technical support, and study supervision.

Compliance with ethical standards

Conflict of interest Stuart C. Gordon receives grant/research support from AbbVie Pharmaceuticals, Bristol-Myers Squibb, Conatus, CymaBay, Exalenz BioScience, Gilead Pharmaceuticals, Intercept Pharmaceuticals, and Merck. He is also a consultant/advisor for Abbvie, Bristol-Myers Squibb, CVS Caremark, Gilead, Intercept, and Merck, and serves as a speaker/teacher in programs sponsored by Gilead Pharmaceuticals and Intercept Pharmaceuticals. The other authors have no potential conflicts of interest.

Ethical standards The CHeCS investigation follows the guidelines of the US Department of Health and Human Services regarding the protection of human subjects. The study protocol was approved and is renewed annually by the institutional review board at each participating site.

References

1. Kansagara D, Papak J, Pasha AS, et al. Screening for hepatocellular cancer in chronic liver disease: a systematic review. VA-ESP Project #05-225; 2013.
2. El-Serag HB. Hepatocellular carcinoma. *New Engl J Med*. 2011;365:1118–1127.
3. U.S. Cancer Statistics Working Group. United States Cancer Statistics: 1999–2012 Incidence and Mortality Web-based Report. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2015. https://nccd.cdc.gov/uscs/Survival/Relative_Survival_Tables.pdf. Accessed Jan 4, 2016.
4. Bruix J, Sherman M. Management of hepatocellular carcinoma. *Hepatology*. 2005;42:1208–1236.
5. Bruix J, Sherman M. management of hepatocellular carcinoma: an update. *Hepatology*. 2011;53:1020–1022.
6. Bruix J, Sherman M, Llovet JM, et al. for EASL panel of experts on HCC. Clinical management of hepatocellular carcinoma: conclusions of the Barcelona—2000 EASL conference. *J Hepatol*. 2001;35:421–430.
7. Omata M, Lesmana LA, Tateishi R, Chen PJ, Lin SM, Yoshida H. Asian Pacific Association for the Study of the Liver consensus recommendations on hepatocellular carcinoma. *Hepatol Int*. 2010;4:439–474.
8. Holmberg SD, Lu M, Rupp L, et al. Noninvasive serum fibrosis markers for screening and staging chronic hepatitis C virus patients in a large US cohort. *Clin Infect Dis*. 2014;57:240–246.
9. Lin ZH, Xin YN, Dong QJ, et al. Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. *Hepatology*. 2011;53:726–736.
10. Velazquez RF, Rodriguez M, Navascues CA, et al. Prospective analysis of risk factors for hepatocellular carcinoma in patients with liver cirrhosis. *Hepatology*. 2003;37:520–527.
11. Rodriguez-Diaz JL, Rosas-Camargo V, Vega-Vega O, et al. Clinical and pathological factors associated with the development of hepatocellular carcinoma in patients with hepatitis virus-related cirrhosis: a long-term follow-up study. *Clin Oncol*. 2007;19:197–203.
12. Ganne-Carrie N, Chastang C, Chapel F, et al. Predictive score for the development of hepatocellular carcinoma and additional value of liver large cell dysplasia in Western patients with cirrhosis. *Hepatology*. 1996;23:1112–1118.
13. Bonis PAL, Tong MJ, Blatt LM, Conrad A, Griffith JL. A predictive model for the development of hepatocellular carcinoma, liver failure, or liver transplantation for patients presenting with chronic hepatitis C. *Am J Gastroenterol*. 1999;94:1605–1612.
14. Lok AS, Seeff LB, Morgan TR, et al. Incidence of hepatocellular carcinoma and associated risk factors in hepatitis C-related advanced liver disease. *Gastroenterol*. 2009;136:138–148.
15. Hung YC, Lin CL, Liu CJ, et al. Development of risk scoring system for stratifying population for hepatocellular carcinoma screening. *Hepatology*. 2015;61:1934–1944.
16. Wen CP, Lin J, Yang YC, et al. Hepatocellular carcinoma risk prediction model for the general population: the predictive power of transaminases. *J Natl Cancer Inst*. 2012;104:1599–1611.
17. Sherman M. Hepatitis C and hepatocellular carcinoma: Grist for the mill. *Gastroenterol*. 2009;136:39–42.
18. Sullivan LM, Massaro JM, D'Agostino RB. Tutorial in biostatistics: presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Statist Med*. 2004;23:1631–1660.
19. Moorman AC, Gordon S, Rupp L, et al. Baseline characteristics and mortality among people in care for chronic viral hepatitis: the Chronic Hepatitis Cohort Study. *Clin Infect Dis*. 2013;56:40–50.
20. Loaeza-del-Castillo A, Paz-Pineda F, Oviedo-Cárdenas E, Sánchez-Avila F, Vargas-Voráckova F. AST to platelet ratio index (APRI) for the noninvasive evaluation of liver fibrosis. *Ann Hepatol*. 2008;7:350–357.
21. Miao Y. Estimating Harrell's optimism on predictive indices using Bootstrap samples. Paper 504; SAS Global Forum; 2013.
22. Kremers WK. Calculation of the C-statistic (discrimination index) for survival data with time dependent covariates. http://www.mayo.edu/research/documents/survcstdsas/doc-10027180.
23. Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med*. 1984;3:143–152.
24. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Med Decis Mak*. 1993;13:49–58.
25. Harrell FE, Lee KL, Mark DB. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
26. Grønnesby JK, Borgan Ø. A method for checking regression models in survival analysis based on the risk score. *Lifetime Data Anal*. 1996;2:315–320.
27. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. *Stat Med*. 2015;34:1659–1680.
28. Konerman MA, Zhang Y, Zhu J, Higgins PDR, Lok ASF, Waljee AK. Improvement of predicitive models of risk of disease progression in chronic hepatitis C by incorporating longitudinal data. *Hepatology*. 2015;61:1832–1841.
29. Kim H, Kisseleva T, Brenner DA. Aging and liver disease. *Curr Opin Gastroenterol*. 2015;31:184–191.

30. Miki D, Aikata H, Uka K, et al. Clinicopathological features of elderly patients with hepatitis C virus-related hepatocellular carcinoma. *J Gastroenterol*. 2008;43:550–557.
31. Asahina Y, Tsuchiya K, Tamaki N, et al. Effect of aging on risk for hepatocellular carcinoma in chronic hepatitis C virus infection. *Hepatology*. 2010;52:518–527.
32. Hamada H, Yatsuhashi H, Yano K, et al. Impact of aging on the development of hepatocellular carcinoma in patients with post-transfusion hepatitis C. *Cancer*. 2002;95:331–339.
33. Ohishi O, Kitamoto M, Aikata H, et al. Impact of aging on the development of hepatocellular carcinoma in patients with hepatitis C virus infection in Japan. *Scand J Gastroenterol*. 2003;38:894–900.
34. Saneto H, Kobayashi M, Kawamura Y, et al. Clinicopathological features, background liver disease, and survival analysis of HCV-positive patients with hepatocellular carcinoma: differences between young and elderly patients. *J Gastroenterol*. 2008;43: 975–981.
35. Poynard T, Ratziu V, Charlotte F, Goodman Z, McHutchison J, Albrecht J. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C. *J Hepatol*. 2001;34:730–739.
36. Lu M, Li J, Rupp LB, Holmberg SD, et al. Hepatitis C treatment failure is associated with increased risk of hepatocellular carcinoma. *J Viral Hepat*. 2016;23:718–729.
37. Baran B, Gulluoglu M, Soyer OM, et al. Treatment failure may lead to accelerated fibrosis progression in patients with chronic hepatitis C. *J Viral Hepat*. 2014;21:111–120.
38. Boscarino JA, Rupp LB, Moorman AC, et al. Depression and alcohol misuse among patients in the Chronic Hepatitis Cohort Study (CHeCS): comparison of PHQ-8 and Audit-C instruments vs. ICD-9 codes. The American Association for the Study of Liver Diseases 66th Annual Meeting, San Francisco, CA, November 13–17, 2015. Abstract final ID: 525.