HHS Public Access

Author manuscript

J Hepatol. Author manuscript; available in PMC 2022 February 01.

Published in final edited form as:

J Hepatol. 2022 February ; 76(2): 294–301. doi:10.1016/j.jhep.2021.09.009.

Comparative performance of risk prediction models for hepatitis B-related hepatocellular carcinoma in the United States

Hyun-seok Kim[1], **Xian Yu**[2], **Jennifer Kramer**[2], **Aaron P. Thrift**[3,4], **Pete Richardson**[2], **Yao-Chun Hsu**[5], **Avegail Flores**[2,3], **Hashem B. El-Serag**[1,2], **Fasiha Kanwal**[1,2,*]

[1]Section of Gastroenterology and Hepatology, Baylor College of Medicine, Houston, TX, USA

[2]Clinical Epidemiology and Comparative Effectiveness Program, Section of Health Services Research (IQuESt), Michael E. DeBakey Veterans Affairs Medical Center, Houston, TX, USA

[3]Section of Epidemiology and Population Sciences, Baylor College of Medicine, Houston, TX, USA

[4]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX, USA

[5]Center for Liver Diseases and Division of Gastroenterology and Hepatology, E-Da Hospital/I-Shou University, Kaohsiung, Taiwan

Abstract

Background & Aims: Guidelines recommend hepatocellular carcinoma (HCC) surveillance in patients with chronic HBV infection. Several HCC risk prediction models are available to guide surveillance decisions, but their comparative performance remains unclear.

Methods: Using a retrospective cohort of patients with HBV treated with nucleos(t)ide analogues at 130 Veterans Administration facilities between 9/1/2008 and 12/31/2018, we calculated risk scores from 10 HCC risk prediction models (REACH-B, PAGE-B, m-PAGE-B, CU-HCC, HCC-RESCUE, CAMD, APA-B, REAL-B, AASL-HCC, RWS-HCC). We estimated the models' discrimination and calibration. We calculated HCC incidence in risk categories defined by the reported cut-offs for all models.

Results: Of 3,101 patients with HBV (32.2% with cirrhosis), 47.0% were treated with entecavir, 40.6% tenofovir, and 12.4% received both. During a median follow-up of 4.5 years, 113 patients developed HCC at an incidence of 0.75/100 person-years. AUC values for 3-year HCC risk were the highest for RWS-HCC, APA-B, REAL-B, and AASL-HCC (all >0.80). Of these, 3 (APA-B,

*Corresponding author. Address: Section of Gastroenterology and Hepatology, Baylor College of Medicine, 7200 Cambridge Street, Houston, Texas 77030, USA; Phone: 713-798-6847. kanwal@bcm.edu (F. Kanwal).

Authors' contributions

Study and design: all authors. Acquisition, analysis, or interpretation of data: all authors. Drafting of the manuscript: all authors. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: HSK, XY. Administrative, technical, or material support: all authors. Study supervision: all authors. Dr. H.S. Kim had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version of the manuscript.

Conflict of interest

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

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

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2021.09.009.

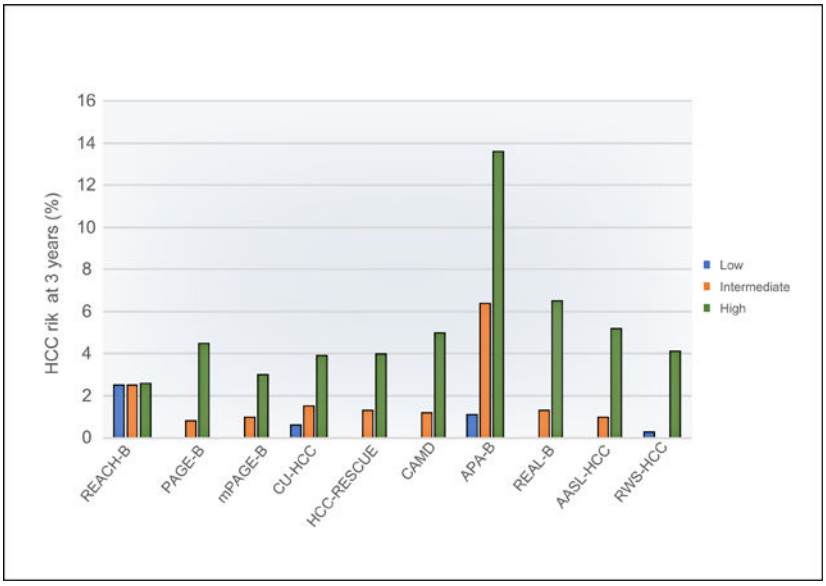
RWS-HCC, REAL-B) incorporated alpha-fetoprotein. AUC values for the other models ranged from 0.73 for PAGE-B to 0.79 for CAMD and HCC-RESCUE. Of the 7 models with AUC >0.75, only APA-B was poorly calibrated. In total, 10–20% of the cohort was deemed low-risk based on the published cut-offs. None of the patients in the low-risk groups defined by PAGE-B, m-PAGE-B, AASL-HCC, and REAL-B developed HCC during the study timeframe.

Conclusion: In this national cohort of US-based patients with HBV on antiviral treatment, most models performed well in predicting HCC risk. A low-risk group, in which no cases of HCC occurred within a 3-year timeframe, was identified by several models (PAGE-B, m-PAGE-B, CAMD, AASL-HCC, REAL-B). Further studies are warranted to examine whether these patients could be excluded from HCC surveillance.

Lay summary:

Risk prediction models for hepatocellular carcinoma (HCC) in patients infected with hepatitis B virus (HBV) could guide HCC surveillance decisions. In this large cohort of US-based patients receiving treatment for HBV, most published models discriminated between those who did or did not develop HCC, although the RWS-HCC, REAL-B, and AASL-HCC performed the best. If confirmed in future studies, these models could help identify a low-risk subset of patients on antiviral treatment who could be excluded from HCC surveillance.

Graphical Abstract



Keywords

Hepatitis B virus; hepatocelluar carcinoma; external validation

Introduction

Chronic HBV infection is the most common chronic viral infection in the world and the leading cause of hepatocellular carcinoma (HCC) globally.[1] HCC risk in HBV is

characterized by considerable variability related to demographic factors (age, sex, race/ethnicity), disease severity and activity (fibrosis stage, HBV DNA level, HBeAg status), metabolic disease (diabetes, obesity), and lifestyle factors (alcohol, smoking).[2] Accurate information regarding future risk of HCC is important for clinicians and patients to make optimal clinical care decisions, including those related to HCC surveillance.

To enable risk stratification, several risk models have been developed to predict future HCC risk among patients with HBV. These include the REACH-B[3] (includes age, sex, alanine aminotransferase [ALT], HBeAg, HBV DNA); PAGE-B (age, sex, platelet)[4]; mPAGE-B[5](age, sex, platelet, albumin); CU-HCC[6] (age, cirrhosis, albumin, bilirubin, HBV DNA); HCC-RESCUE (age, sex, cirrhosis)[7]; CAMD[8] (cirrhosis, age, sex, and diabetes mellitus); APA-B[9] (age, platelet, alpha-fetoprotein [AFP]); REAL-B[10] (age, sex, alcohol use, cirrhosis, diabetes mellitus, platelet, AFP); AASL-HCC[11](age, albumin, sex, liver cirrhosis); and RWS-HCC[12] (age, sex, cirrhosis, AFP) scores – the full names of scores are provided in the abbreviations section. With the exception of PAGE-B, these models were developed and tested in mainly Asian patients infected with HBV, potentially limiting their use for risk stratification of patients seen in clinical practice in the US.[8,13] This is because there are significant differences between Asians and non-Asians in the mode of transmission, HBV genotype, and distribution of other risk factors such as age, race, and obesity which may result in differences in natural history and disease progression to HCC. In addition, these models were developed in cohorts with varying proportions of patients on antiviral treatment (*e.g.*, 0% in the REACH-B cohort, 15.1% in CU-HCC, 100% in the remaining 8 cohorts) and patients with cirrhosis (*e.g.*, 0% in the REACH-B model to nearly 40% in the cohorts for the CU-HCC and AASL-HCC models), making it difficult to apply these models in routine clinical practice. To our knowledge, no study has examined the performance of HCC risk models in US patients with HBV. There are also no data on the comparative effectiveness of different HBV-HCC models in non-Asian patients with HBV. The utility of these models in clinical practice also remains unknown. This information could play a central role in guiding which model(s) to use for clinical decision-making in US patients with HBV.

In this study, we examined and compared the performance of 10 HCC risk prediction models in a large cohort of patients with HBV treated with entecavir or tenofovir in routine clinical practice at 130 Veterans Administration (VA) facilities and their affiliated clinics.

Patients and methods

Data source

The VA healthcare system is the largest integrated healthcare provider in the US. We used data from the national VA Corporate Data Warehouse that includes all laboratory test results, pharmacy, and inpatient and outpatient procedures and diagnosis codes for patients utilizing the VA for healthcare. We also used the VA Purchased Care database of services paid by but rendered outside the VA. The VA Central Cancer Registry (CCR) is a national repository for VA patients with cancer. Local registrars manually abstract data using the North American Association of Central Cancer Registries standards. We obtained the date of death from

the VA Vital Status file that combines death data from Medicare, VA, and Social Security (sensitivity 98.3%; specificity 99.8% relative to the National Death Index).[14]

Study population

The study cohort included patients aged 18 years and older with chronic HBV infection, defined by at least 1 positive HBsAg test between September 1, 2008 and December 31, 2017, who had at least 1 filled prescription for entecavir or tenofovir. We chose September 1, 2008 as a start study time because tenofovir was first approved by the US Food and Drug Administration in September, 2008.[15] We utilized the date of the first dispensed prescription of entecavir or tenofovir as the index date for follow up. We excluded patients with HIV co-infection, defined by presence of ICD-9 or ICD-10 codes, and those with HCV co-infection, defined based on any positive HCV RNA test any time during study duration. Finally, we excluded patients with prevalent HCC defined as HCC diagnosed any time before or during the first year after treatment initiation.

HCC definition

We used a multi-step process to identify incident HCC. HCC was initially identified using the ICD-9 code (155.0: Malignant neoplasm of liver, primary in the absence of 155.1: intrahepatic bile duct carcinoma) and ICD-10 code (C.22.0: Liver cell carcinoma) in the VA Corporate Data Warehouse data. We then examined the VA CCR for patients with HCC diagnosis based on primary site code C220 with histology codes 817XX through 818XX and text searches for liver and hepatocellular carcinoma. For patients who had an ICD-9/10 code but were not identified as having HCC in the CCR, we conducted a manual review of the VA electronic medical record (EMR) to determine their true HCC status. This hierarchical approach ensured high validity of all the captured HCC cases. HCC diagnosis date was defined as the date the patient first met our HCC case definition. The study follow-up ended at the time of diagnosis of HCC, death, last visit recorded in the VA, or December 31, 2019.

Variables for HBV-HCC models

We obtained data on the individual factors included in 10 HBV-HCC models (Table 1 and Table S1). Socio-demographic variables included age at the start of antiviral treatment (index date), sex, and race/ethnicity (White, African American, Hispanic, Asian, and other). For all patients, we extracted data for blood platelet count (10^3/ul), aspartate aminotransferase (AST)/ALT (U/L), albumin (g/dl), bilirubin (mg/dl), international normalized ratio, HBeAg, AFP, and HBV DNA tests that were performed within 1 year prior and closest to the index date. Cirrhosis was determined based on ICD-9 (456.1, 456.21, 456.0, 571.2, 571.5, 571.6, 789.5, 789.60, 789.59, 567.23, 572.2, 572.5, 573.5) or ICD-10 code (E83.110, G93.40, I85.00, I86.40, I85.10, I85.01, I86.41, I85.11, K65.2, K70.30, K70.31, K70.11, K71.91, K71.51, K71.7, K72.11, K74.60, K74.69, K74.3, K74.4, K74.5, K76.7, K76.81, R18.8) any time prior to the index or a fibrosis-4 (FIB-4) ≥3.25 within 1 year of the index date. Our group previously reported that the positive predictive value (probability that cirrhosis is present based on EMR reviews among those with a cirrhosis ICD code) and negative predictive value (probability that cirrhosis is absent based on EMR reviews among those without a cirrhosis ICD code) were 90% and 87%.[16] We identified diabetes and alcohol abuse based on ICD-9/10 diagnosis codes recorded any time before

the index date. We used a combination of ICD codes for alcohol use and results from the annual AUDIT-C scores (≥4 in men and ≥3 in women) any time prior to or during study follow-up to determine history of alcohol abuse. See Table S2 for the ICD codes used to define study variables. We converted HBV DNA reported as picogram/ml or copies/ml to IU/ml (1 picogram/ml = 270,000 copies/ml, 1 IU/ml = 5 copies/ml).

HBV-HCC models

We examined 10 published HBV-HCC prediction models: REACH-B,[3] PAGE-B,[4] mPAGE-B,[5] CU-HCC,[6] HCC-RESCUE,[7] CAMD,[8] APA-B,[9] REAL-B,[10] AASL-HCC, and RWS-HCC.[12] Except for the PAGE-B model which was developed in Europe, all models were developed in Asian patients. REACH-B was developed in patients without cirrhosis. The proportion of patients with cirrhosis ranged from 0% in the cohort used for the development for the REACH-B model to nearly 40% in the cohorts for the CU-HCC and AASL-HCC models. The proportion of patients on antiviral treatment varied from 0% in the REACH-B cohort, to 15.1% in CU-HCC, to 100% in the remaining 8 cohorts. Information on the specific cut-offs and risk parameters in each model is shown in Table S1. Most models relied on information available at the time of antiviral treatment initiation with 1 exception: the APA-B model uses platelet and AFP levels at 12 months following initiation of HBV treatment. HBV-HCC prediction models were developed for different time-periods of risk prediction. AASL-HCC predicted both 3- and 5-year risk; PAGE-B, mPAGE-B, CU-HCC, HCC-RESCUE, CAMD, APA-B dicted 5-year risk; REACH-B and REAL-B models predicted 3-, 5-, and 10-year risk; and RWS-HCC predicted 10-year risk. In this study we examined 3- and 5-year risk since these were the most commonly utilized time periods.

Data analysis

We compared the demographic, clinical, and virus-related risk factors between patients who developed HCC and those who did not using chi-square or Fisher's exact tests for categorical variables and *t* test for continuous variables. We assessed performance of the HBV-HCC risk prediction models using measures of discrimination and calibration.[17] Discrimination describes the ability of the model to distinguish patients who develop an event (HCC) from those who do not. Calibration is the ability of the model to accurately estimate the absolute risk.

We used time to event Cox proportional hazard models to examine the association between each model and risk of incident HCC. We examined each models' discrimination using the area under the curve receiver-operating characteristic (AUC) curve. We assessed the performance of each model for prediction of time to HCC risk within 3 and 5 years to examine the different prediction horizons in the original studies from which the risk scores were derived. For example, for the 3-year risk, we considered individuals as incident cases if they had developed HCC within the first 3 years of follow-up. Patients who developed HCC after 3 years of follow-up were included in the 3-year prediction as non-cases.

We examined calibration via plots to visualize the relationship between predicted risk score and observed HCC risk within the 3-year time window using a logistic regression model.

We plotted the HCC cumulative incidence rate by the risk score for both the predicted and observed risk for HCC within 3 years. We also examined the Hosmer-Lemeshow test for goodness of fit using the logistic regression model.[18,19]

To gain insights into the clinical utility of the models, we estimated the cumulative HCC risk at 3 and 5 years of follow-up according to risk categories defined by the models' clinical cut-offs using cox-regression analysis.

Sensitivity and subgroup analyses—Few laboratory values were missing in more than 2% of patients (Table 1). Specifically, HBeAg tests were missing in 8.8%, HBV DNA in 10.2%, platelet count 5.8%, AST 6.2%, ALT 5.6%, albumin 14.2%, bilirubin 2.0%, and AFP in 27.5% of the cohort. We used complete-case analysis as our primary approach. However, in a sensitivity analysis, we imputed missing data using multiple imputation and repeated all analyses to assess the degree to which our results were sensitive to missing data. The imputation model included outcomes and all covariates. Estimates from regressions performed on 5 imputed data sets were combined using Rubin's rule.[20]

To examine whether each models' performance was different in key racial/ethnic groups, we conducted subgroup analyses in White and African American patients. We used SAS version 9.4 and R studio version 1.2.5019 for all analyses. p values <0.05 were considered significant. This study was approved by the Institutional Review Boards of Baylor College of Medicine and Michael E. DeBakey Veterans Affairs Hospital, Houston, Texas. The data that support the findings of this study are available on request from the corresponding author, [FK]. The data are not publicly available due to privacy/ethical restrictions.

Results

Demographic characteristics

Our study included 3,101 patients with chronic HBV on entecavir (47.0%), tenofovir (40.6%), or both treatments (12.4%) (Table 1). On average, patients were on antiviral treatment for 3.1 years (SD 2.8 year) with 23 (SD, 22) prescriptions per patient during the follow-up period. In total, 28.5% of patients received less than 1 year of treatment.

The mean age at the index date was 56.8 years (SD 13.1) and 94.9% were men. Approximately 39.5% were White, 36.3% African American, and 13.6% were Asians. A total of 32.2% of patients had a diagnosis of cirrhosis, 38.8% were positive for HBeAg, and 54.8% had HBV DNA levels greater than 2,000 IU/ml. The mean AST and ALT values were 86.2 (SD, 141.6) and 101.1 (SD, 163.7) U/L, respectively. Approximately a quarter of the study population had diabetes (26.9%) and one-third had a history of alcohol abuse (30.2%).

During a median follow-up of 4.5 years (IQR 2.15–7.57) and an overall follow-up of 15,159 person-years, 113 (3.6%) patients developed HCC at an incidence rate of 0.75 (95% CI 0.61–0.90) per 100 person-years. Patients with cirrhosis developed HCC at an incidence of 1.74 (95% CI 1.35–2.20) per 100 person-years, whereas HCC incidence was 0.40 (95% CI 0.29–0.54) per 100 person-years in patients without cirrhosis.

In the unadjusted analyses, compared to patients without HCC, those who developed HCC were older, more likely to be African American, and had cirrhosis. Patients who developed HCC also had lower platelet count and albumin level at the time of treatment initiation than those who did not. There were no statistical differences in the baseline values of AFP, HBV DNA, HBeAg status, AST, or ALT between the 2 groups. The type of HBV treatment was not significantly different in those who developed HCC compared to those who did not.

Model discrimination

Table 2 shows the AUC values for the 10 HBV-HCC prediction models. Except for the REACH-B model (AUC value 0.57, 95% CI 0.49–0.65 for 3-year HCC risk and 0.58, 95% CI 0.52–0.64 for 5-year HCC risk), the AUC values of all models were greater than 0.70. The AUC value for 3-year HCC risk was the highest for RWS-HCC (0.89, 95% CI 0.85–0.93), followed by APA-B (0.87, 95% CI 0.83–0.91), AASL-HCC (0.86, 95% CI 0.82–0.90), and REAL-B (0.84, 95% CI 0.78–0.90). The AUC values for 3-year HCC risk were between 0.73 and 0.80 for CU-HCC, HCC-RESCU, and CAMD. There were no significant differences in AUC among these 9 models. The AUC values for 5-year risk showed trends that were similar to those for 3-year risk (Table 2).

These results did not change in the sensitivity analysis after multiple imputation with the highest AUC value of 0.91 (95% CI 0.87–0.95) for the RWS-HCC model (Table S3). In the subgroup analyses by race/ethnicity (White *vs.* African American patients), there was no significant difference in AUC values by race/ethnicity. RWS-HCC, AASL-HCC, REAL-B, and APA-B had AUC values >0.80 in both subgroups (Table 4).

Model calibration

Among the 7 models with useful discrimination (AUC >0.75), most (CU-HCC, HCC-RESCUE, CMAD, REAL-B, AASL-HCC and RWS-HCC) were well calibrated (see Hosmer-Lemeshow *p* values, Table 2). Only APA-B demonstrated poor calibration. Most models underestimated the actual risk in patients with 3-year cumulative incidence of 10% or higher (Fig. 1).

Cumulative HCC risk by each model's risk category

The cumulative HCC risk based on the risk categories defined by clinical cut-offs from each model showed clear discrimination except for the REACH-B model (for HCC risk at 3 years: low risk group: 0%-2.5%, intermediate risk group: 0.8%-6.4%, high risk group: 2.6%-13.6%; for HCC risk at 5 years, low risk: 0%-3.0%, intermediate risk: 1.4%-12.8%, and high risk: 2.6%-15.9%) (Table 3). In total, approximately 10–20% of the cohort was deemed a low-risk group by the reported clinical cut-offs of most models. None of the patients in the low-risk groups defined by PAGE-B (n = 287, 9.7%), m-PAGE-B (n = 344, 13.5%), AASL-HCC (n = 294, 11.1%) and REAL-B (n = 355 16.4%) developed HCC during the study timeframe. Of all models, CAMD identified the largest group of patients deemed at low risk for HCC (n = 677, 21.8%); of these, no patient progressed to HCC within 3 years although 1 had developed HCC at the 5-year follow-up. For the high-risk groups, the 3-year risk of HCC was as high as 13.6% (95% CI 7.2–24.8) defined by APA-B followed by 6.5% (95% CI 4.4–9.6) based on REAL-B.

Discussion

Accurate information regarding patients' risk of HCC is fundamental to optimal surveillance and risk reduction. Our study examined 10 published HBV-HCC prediction models and evaluated their predictive performance in a large multiracial US-based cohort of patients with HBV treated with entecavir or tenofovir as part of routine clinical care. This comparative evaluation is important to guide the selection of the best available models for risk stratification, to inform patient-centered decisions about the risk and benefits of HCC surveillance, and to guide patient selection for future clinical trials.

We found that most models showed good overall discrimination with AUC values of 0.75 or greater for 7 of the 10 models (CAMD, APA-B, HCC-RESCUE, CU-HCC, AASL-HCC, RWS-HCC, REAL-B). AUC values of 3 models that incorporated AFP as a variable (APA-B, RWS-HCC, and REAL-B) were the highest (all >0.80), suggesting that incorporating HCC biomarker data is important for risk prediction in HBV. Indeed, several other risk stratification models for HCC, such as the HES model and the GALAD score rely on AFP testing in their algorithms.[21,22] One model, REACH-B, did not perform well, with an AUC value of 0.57 (95% CI 0.49–0.65). The REACH-B[3] model was developed in a cohort of untreated patients without cirrhosis and did not accurately differentiate patients who developed HCC from those who did not in our cohort of patients (32.2% cirrhosis) who were on antiviral treatment. Wu *et al.* also reported that the AUC value of REACH-B in a Chinese external validation cohort was the lowest among existing HBV-HCC prediction models (AUC 0.68; 95% CI 0.51–0.85).[23]

Discrimination alone is insufficient to assess a model's prediction capability. Calibration allows assessment of model's ability to predict the absolute magnitude of risk. Of the 7 models with good discrimination (AUC >0.75), all except APA-B were well calibrated based on a non-significant Hosmer-Lemeshow goodness of fit test. Of these, CU-HCC, CAMD, AASL-HCC and RWS-HCC models appeared better calibrated based on the visual inspection of plots (Fig. 1). Most models were accurate at estimating 3-year HCC risk for patients in the range from 0% to 10%, but under-estimated risks among patients with 3-year cumulative risk of 10% or higher. Such poor calibration among patients with higher risk may not be a problem since the annual HCC incidence threshold for decision making for HCC surveillance in patients with HBV ranges from 0.2% to 1.0%. Most models accurately predicted HCC risk at the lower end of the risk spectrum.

HBV-HCC models could be clinically valuable if they can be used to identify low-risk patients. None of the patients in the low-risk groups defined by PAGE-B (n = 287, 9.7%), m-PAGE-B (n = 344, 13.5%), AASL-HCC (n = 294, 11.1%) and REAL-B (n = 355 16.4%) developed HCC during the study timeframe (Table 4). The clinical utility of the risk prediction scores is driven not only by their ability to identify the lowest risk group but also by the absolute size of this low-risk group; these patients can be potentially excluded from HCC surveillance, thereby enhancing the overall cost-effectiveness of HCC surveillance. Of all models, CAMD identified the largest group of patients deemed at low risk for HCC (n = 677, 21.8%); of these, no patient progressed to HCC within 3 years although 1 had developed HCC at the 5-year follow-up.

We also found that model performance was not different in subgroups defined based on race/ethnicity. Our study provides the first comprehensive data on the performance of the HBV-HCC prediction models in African American patients with HBV and fills in an important evidence gap.

Our study has several limitations. Our analysis was restricted to patients diagnosed and treated in the VA healthcare system, thus generalizability to other US patients, especially female patients, requires further evaluation. We did not have complete information on HBV DNA and AFP on all patients given variations in testing practices, although we used multiple imputation to overcome this missingness.[24] In addition, we could not conduct meaningful subgroup analyses by race/ethnic group, especially for Asians, due to insufficient sample sizes (Only 6 HCC cases among 428 Asians). Our study is limited by the observational retrospective nature of its design and missing some variables such as family history of HCC. We are also limited by the accuracy of the ICD code-based diagnosis for key variables. Yet, we confirmed each HCC case in the patients' EMR and calculated FIB-4 to complement our definition of cirrhosis.

In conclusion, in this national multi-racial/ethnic cohort of US-based patients with treated HBV, most published models identified patients at high risk of progressing to HCC. These included CAMD, APA-B, HCC-RESCUE, CU-HCC, AASL-HCC, RWS-HCC, REAL-B model. Of these, RWS-HCC, REAL-B, and AASL-HCC had the highest AUC values and also predicted the actual risk of HCC closely. All low-risk patients (~10–20% of the cohort) defined by PAGE-B, m-PAGE-B, AASL-HCC, CMAD and REAL-B models did not develop HCC during the first 3-years of follow-up. Further studies are warranted to examine whether this low-risk group may be excluded from HCC surveillance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support

NIH U01 CA230997, P30 DK056338, T32 DK083266 to Dr. Fasiha Kanwal. Dr. Kim is supported by T32 DK083266.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author, [FK]. The data are not publicly available due to privacy/ethical restrictions.

Abbreviations

AASL-HCC	age, albumin, sex, liver cirrhosis-HCC
AFP	alpha-fetoprotein

ALT	alanine aminotransferase
APA-B	age, platelet, AFP
AST	aspartate aminotransferase
CAMD	cirrhosis, age, male sex, and diabetes mellitus
CU-HCC	Chinese University HCC
EMR	electronic medical record
FIB-4	fibrosis-4
HCC	hepatocellular carcinoma
HCC-RESCUE	HCC-Risk Estimating Score in CHB patients Under Entecavir
mPAGE-B	modified PAGE-B
REACH-B	Risk estimation for hepatocellular carcinoma in chronic hepatitis B
REAL-B	Real-World Effectiveness From the Asia Pacific Rim Liver Consortium for HBV Risk Score
RWS-HCC	Real-world risk score for hepatocellular carcinoma
VA	Veterans Affairs.

References

[1]. Polaris Observatory C. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. Lancet Gastroenterol Hepatol 2018;3:383–403. [PubMed: 29599078]

[2]. Kim WR. Risk of incident hepatocellular carcinoma in hepatitis B-infected patients treated with tenofovir disoproxil fumarate versus entecavir: a US administrative claims analysis. Hepatology 2019;70:302A–303A.

[3]. Yang HI, Yuen MF, Chan HL, Han KH, Chen PJ, Kim DY, et al. Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score. Lancet Oncol 2011;12:568–574. [PubMed: 21497551]

[4]. Papatheodoridis G, Dalekos G, Sypsa V, Yurdaydin C, Buti M, Goulis J, et al. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. J Hepatol 2016;64:800–806. [PubMed: 26678008]

[5]. Kim JH, Kim YD, Lee M, Jun BG, Kim TS, Suk KT, et al. Modified PAGE-B score predicts the risk of hepatocellular carcinoma in Asians with chronic hepatitis B on antiviral therapy. J Hepatol 2018;69:1066–1073. [PubMed: 30075230]

[6]. Wong VW, Chan SL, Mo F, Chan TC, Loong HH, Wong GL, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. J Clin Oncol 2010;28:1660–1665. [PubMed: 20194845]

[7]. Sohn W, Cho JY, Kim JH, Lee JI, Kim HJ, Woo MA, et al. Risk score model for the development of hepatocellular carcinoma in treatment-naive patients receiving oral antiviral treatment for chronic hepatitis B. Clin Mol Hepatol 2017;23:170–178. [PubMed: 28506056]

[8]. Hsu YC, Yip TC, Ho HJ, Wong VW, Huang YT, El-Serag HB, et al. Development of a scoring system to predict hepatocellular carcinoma in Asians on antivirals for chronic hepatitis B. J Hepatol 2018;69:278–285. [PubMed: 29551708]

[9]. Chen CH, Lee CM, Lai HC, Hu TH, Su WP, Lu SN, et al. Prediction model of hepatocellular carcinoma risk in Asian patients with chronic hepatitis B treated with entecavir. Oncotarget 2017;8:92431–92441. [PubMed: 29190928]

[10]. Yang HI, Yeh ML, Wong GL, Peng CY, Chen CH, Trinh HN, et al. Real-world effectiveness from the Asia pacific Rim liver Consortium for HBV risk score for the prediction of hepatocellular carcinoma in chronic hepatitis B patients treated with oral antiviral therapy. J Infect Dis 2020;221:389–399. [PubMed: 31550363]

[11]. Yu JH, Suh YJ, Jin YJ, Heo NY, Jang JW, You CR, et al. Prediction model for hepatocellular carcinoma risk in treatment-naive chronic hepatitis B patients receiving entecavir/tenofovir. Eur J Gastroenterol Hepatol 2019;31:865–872. [PubMed: 30694912]

[12]. Poh Z, Shen L, Yang HI, Seto WK, Wong VW, Lin CY, et al. Real-world risk score for hepatocellular carcinoma (RWS-HCC): a clinically practical risk predictor for HCC in chronic hepatitis B. Gut 2016;65:887–888. [PubMed: 26786688]

[13]. Lee HW, Ahn SH. Prediction models of hepatocellular carcinoma development in chronic hepatitis B patients. World J Gastroenterol 2016;22:8314–8321. [PubMed: 27729738]

[14]. Sohn MW, Arnold N, Maynard C, Hynes DM. Accuracy and completeness of mortality data in the department of veterans Affairs. Popul Health Metr 2006;4:2. [PubMed: 16606453]

[15]. Gilead. U.S. Food and Drug administration approves viread(R) for chronic hepatitis B in adults. 2008. https://www.gilead.com/news-and-press/press-room/press-releases/2008/8/us-food-and-drug-administration-approves-vireadr-for-chronic-hepatitis-b-in-adults.

[16]. Kramer JR, Davila JA, Miller ED, Richardson P, Giordano TP, El-Serag HB. The validity of viral hepatitis and chronic liver disease diagnoses in Veterans Affairs administrative databases. Aliment Pharmacol Ther 2008;27:274–282. [PubMed: 17996017]

[17]. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605. [PubMed: 19477892]

[18]. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of predictio nmodels: aframewor kfo rtraditional and novel measures. Epidemiology 2010;21:128–138. [PubMed: 20010215]

[19]. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–387. [PubMed: 8668867]

[20]. Marshall A,Altm an DG,Hol der RL,Roys to n P.Combin ingestima te sofinterest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol 2009;9:57. [PubMed: 19638200]

[21]. Tayob N, Christie I, Richardson P, Feng Z, White DL, Davila J, et al. Validation of the hepatocellular carcinoma early detection screening (HES) algorithm in a cohort of veterans with cirrhosis. Clin Gastroenterol Hepatol 2019;17:1886–1893 e1885. [PubMed: 30557738]

[22]. Yang JD, Addissie BD, Mara KC, Harmsen WS, Dai J, Zhang N, et al. GALAD score for hepatocellular carcinoma detection in comparison with liver ultrasound and proposal of GALADUS score. Cancer Epidemiol Biomarkers Prev 2019;28:531–538. [PubMed: 30464023]

[23]. Wu S, Zeng N, Sun F, Zhou J, Wu X, Sun Y, et al. HCC prediction models in chronic hepatitis B: a systematic review of 14 models and external validation. Clin Gastroenterol Hepatol 2021.

[24]. Serper M, Choi G, Forde KA, Kaplan DE. Care delivery and outcomes among US veterans with hepatitis B: a national cohort study. Hepatology 2016;63:1774–1782. [PubMed: 26561023]

Highlights

- CAMD, APA-B, HCC-RESCUE, CU-HCC, AASL-HCC, RWS-HCC, REAL-B models identified patients at high risk of developing HCC.
- RWS-HCC, REAL-B, and AASL-HCC had the highest AUCs and also predicted the actual risk of HCC closely.
- No low-risk patients (~10–20% of cohort) defined by PAGE-B, m-PAGE-B, AASL-HCC, CMAD and REAL-B models developed HCC within 3 years.
- Further studies are warranted to examine whether this low-risk group may be excluded from HCC surveillance.

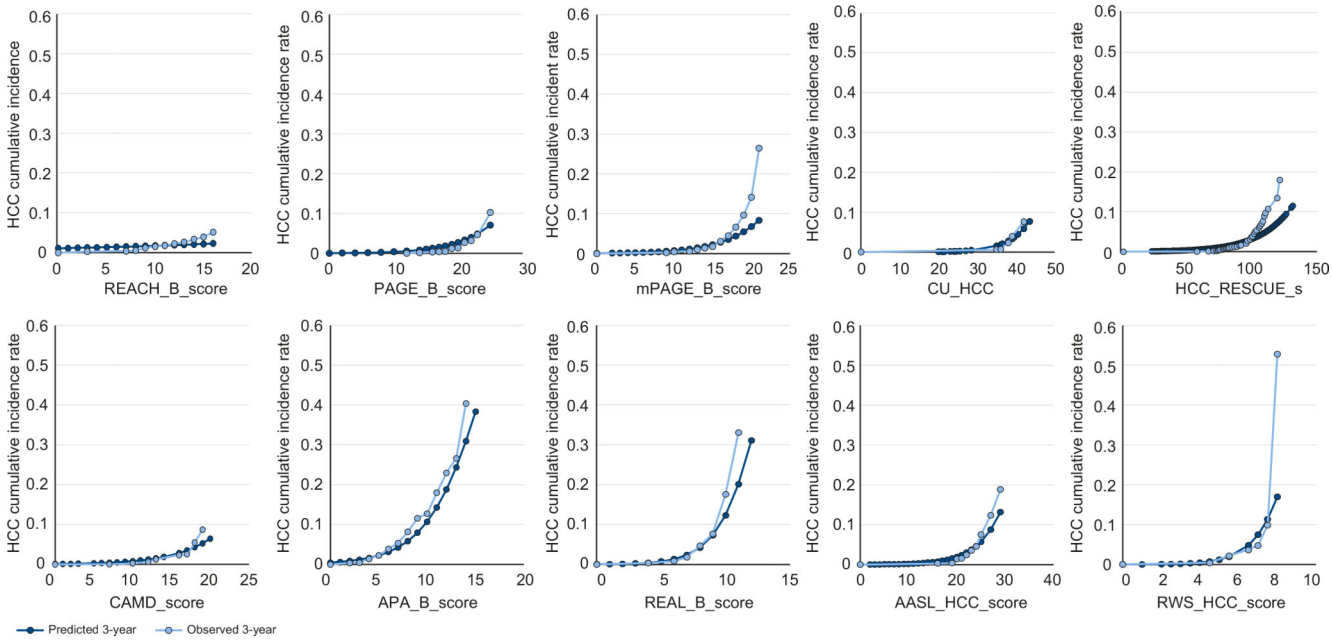


Fig. 1. Calibration plot of HBV-HCC prediction models for HCC risk at 3 years.

X axis denotes the scores of HBV-HCC prediction model and Y axis denotes HCC cumulative incidence rate, ranging from 0 to 1. Calibration plots were made to visualize the relationship between predicted risk score and observed HCC risk within the 3-year time window using a logistic regression model. AASL-HCC, age, albumin, sex, liver cirrhosis-HCC; AFP, alpha-fetoprotein; APA-B, age, platelet, AFP; CAMD, cirrhosis, age, male sex, and diabetes mellitus; CU-HCC, Chinese University HCC; HCC, hepatocellular carcinoma; HCC-RESCUE, HCC-Risk Estimating Score in CHB patients Under Entecavir; mPAGE-B, modified PAGE-B; REACH-B, Risk estimation for hepatocellular carcinoma in chronic hepatitis B; REAL-B, Real-World Effectiveness From the Asia Pacific Rim Liver Consortium for HBV Risk Score; RWS-HCC, Real-world risk score for hepatocellular carcinoma.

Table 1.

Baseline characteristics.

N(%)	All, N = 3,101 (%)	No HCC, n = 2,988 (%)	HCC, n = 113 (%)	*p* value
Age (in years), mean (SD)	56.8 (13.1)	56.7 (13.2)	59.4 (10.8)	<0.01
Male	2,942 (94.9)	2,830(94.71)	112 (99.1)	0.02
Race/ethnicity				
Non-Hispanic White	1,226 (39.5)	1,180 (41.3)	46 (41.8)	<0.01
Non-Hispanic Black	1,124 (36.3)	1,071 (35.8)	53 (46.9)	
Hispanic	99 (3.2)	95 (3.2)	4(3.5)	
Asian	423 (13.6)	417 (14.0)	6(5.3)	
Other	96 (3.1)	95 (3.2)	1 (0.9)	
Diabetes	834 (26.9)	803 (26.9)	31 (27.3)	0.90
Alcohol abuse	936 (30.2)	894 (29.9)	42 (37.2)	0.10
Cirrhosis	919 (32.2)	851 (28.8)	68 (60.2)	<0.01
HBV treatment				
Entecavir	1,457 (47.0)	1,396 (46.7)	61 (54.0)	0.31
TDF	1,259 (40.6)	1,220 (40.8)	39 (34.5)	
Both	385 (12.4)	372 (12.5)	13 (11.5)	
HBeAg				
Positive	1,202 (38.8)	1,157 (38.7)	45 (39.8)	0.97
Negative	1,625 (52.4)	1,567 (52.4)	58 (51.3)	
Missing	274 (8.8)	264 (8.8)	10 (8.9)	
HBV DNA				
<2,000 IU/ml	1,088 (35.1)	1,051 (35.2)	37 (32.7)	0.81
≤2,000 IU/ml	1,398 (54.8)	1,635 (54.7)	63 (55.8)	
Missing	315 (10.1)	302 (10.1)	13 (11.5)	
Platelet (10^9/L), mean (SD)	190.5 (74.3)	192.3 (74.2)	142.9 (61.1)	<0.01
AST (IU/L), mean (SD)	86.2 (141.6)	85.7 (142.0)	98.9 (130.6)	0.35
ALT (IU/L), mean (SD)	101.1 (163.7)	101.5 (164.9)	91.2 (131.1)	0.43
Albumin (g/dl), mean (SD)	3.8 (0.6)	3.8 (0.6)	3.5 (0.7)	<0.01
Total bilirubin, mean (SD)	1.2 (2.7)	1.2 (2.7)	1.4 (2.0)	0.22

N(%)	All, N = 3,101 (%)	No HCC, n = 2,988 (%)	HCC, n = 113 (%)	*p* value
AFP (μg/L), mean (SD)	9.3 (48.0)	9.2 (48.8)	12.5 (19.0)	0.14
Follow-up (years), mean (SD)	4.9 (3.1)	4.9 (3.1)	4.1 (2.5)	<0.01

N (%) missing: HBeAg (n = 274, 8.8%); HBV DNA (n = 315, 10.2%), platelet (n = 179, 5.8%), AST (192, 6.2%), ALT (173, 5.6%), albumin (442, 14.2%), total bilirubin (64, 2.0%), AFP (853, 27.5%). Chi-square or Fisher's exact tests for categorical variables and t test for continuous variables were used to compared people who developed HCC and people who did not.

AFP, alpha-fetoprotein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; HCC, hepatocellular carcinoma; TDF, tenofovir disoproxil fumarate.

Table 2.

Discrimination and calibration of 10 HBV-HCC models for prediction of HCC risk.

	Total number of patients in complete case analysis	Number of HCC in 3 years	Number of HCC in 5 years	AUC		Hosmer-Lemeshow χ^2	
				3 years (95% CI)	5 years (95% CI)	3 years (*p* value)	5 years (*p* value)
REACH-B	2,496	45	71	0.57 (0.49–0.65)	0.58 (0.52–0.64)	8.68 (0.37)	16.11 (0.04)
PAGE-B	2,922	52	80	0.73 (0.67–0.79)	0.73 (0.67–0.79)	4.15 (0.66)	4.85 (0.56)
mPAGE-B	2,586	47	73	0.73 (0.65–0.81)	0.74 (0.68–0.80)	3.33 (0.85)	6.12 (0.53)
CU-HCC	2,414	43	69	0.79 (0.75–0.83)	0.80 (0.76–0.84)	12.64 (0.08)	17.68 (0.01)
HCC-RESCUE	3,101	53	83	0.79 (0.73–0.85)	0.77 (0.69–0.83)	7.54 (0.48)	7.54 (0.48)
CAMD	3,101	53	83	0.79 (0.73–0.85)	0.77 (0.73–0.81)	4.74 (0.69)	3.88 (0.80)
APA-B*	1,749	44	65	0.87 (0.83–0.91)	0.78 (0.74–0.82)	12.40 (0.03)	14.67 (0.01)
REAL-B*	1,858	46	68	0.84 (0.78–0.90)	0.82 (0.78–0.86)	6.27 (0.39)	12.85 (0.05)
AASL-HCC	2,659	47	74	0.86 (0.82–0.90)	0.86 (0.84–0.88)	13.07 (0.11)	28.03 (<0.01)
RWS-HCC*	1,947	46	70	0.89 (0.85–0.93)	0.88 (0.86–0.90)	11.14 (0.19)	19.45 (0.01)

Of note, AASL-HCC predicted both 3- and 5-year risk; PAGE-B, mPAGE-B, CU-HCC, HCC-RESCUE, CAMD, APA-B predicted 5-year risk; REACH-B and REAL-B models predicted 3-, 5-, and 10-year risk; and RWS-HCC predicted 10-year risk. Area under the curve (AUC) and Hosmer-lemeshow chi-square tests were conducted for each model.

AASL-HCC, age, albumin, sex, liver cirrhosis-HCC; AFP, alpha-fetoprotein; APA-B, age, platelet, AFP; CAMD, cirrhosis, age, male sex, and diabetes mellitus; CU-HCC, Chinese University HCC; HCC, hepatocellular carcinoma; HCC-RESCUE, HCC-Risk Estimating Score in CHB patients Under Entecavir; mPAGE-B, modified PAGE-B; REACH-B, Risk estimation for hepatocellular carcinoma in chronic hepatitis B; REAL-B, Real-World Effectiveness From the Asia Pacific Rim Liver Consortium for HBV Risk Score; RWS-HCC, Real-world risk score for hepatocellular carcinoma.

*Highlights the models incorporating AFP variable.

Table 3.

Cumulative HCC risk by risk category of HBV-HCC models.

Models	HCC risk at 3 years (%, 95% CI)	HCC risk at 5 years (%, 95% CI)
REACH-B		
Low	2.5% (1.0–2.4)	3.0% (2.1–4.2)
Intermediate	2.5% (1.6–3.8)	5.2% (3.7–7.3)
High	2.6% (0.4–17.2)	2.6% (0.4–17.2)
PAGE-B		
Low	0% (n.a.)*	0% (n.a.)*
Intermediate	0.8% (0.4–1.5)	1.9% (1.2–3.1)
High	4.5% (3.5–5.9)	6.5% (5.0–8.3)
mPAGE-B		
Low	0% (n.a.)*	0% (n.a.)*
Intermediate	1.0% (0.4–2.2)	1.4% (0.7–2.9)
High	3.0% (2.2–4.3)	6.6% (5.1–8.5)
CU-HCC		
Low	0.6% (0.2–1.6)	1.1% (0.5–2.4)
Intermediate	1.5% (0.7–3.1)	4.2% (2.5–6.8)
High	3.9% (2.7–5.7)	7.3% (5.4–10.0)
HCC-RESCUE		
Low	0% (n.a.)*	0.2% (0–1.4)
Intermediate	1.3% (0.8–2.3)	3.4% (2.4–4.8)
High	4.0% (2.9–5.7)	6.9% (5.1–9.2)
CMAD		
Low	0% (n.a.)*	0.2% (0.0–1.6)
Intermediate	1.2% (0.7–2.0)	3.1% (2.2–4.3)
High	5.0% (3.5–7.2)	8.3% (6.1–11.2)
APA-B		
Low	1.0% (0.6–1.8)	2.2% (1.5–3.4)
Intermediate	6.4% (0.4–11.1)	12.8% (8.0–18.1)
High	13.6% (7.2–24.8)	15.9% (8.7–28.1)
REAL-B		
Low	0% (n.a.)*	0% (n.a.)*
Intermediate	1.3% (0.8–2.2)	3.1% (2.2–4.5)
High	6.5% (4.4–9.6)	10.7% (7.6–14.8)
AASL-HCC		
Low	0% (n.a.)*	0% (n.a.)*
Intermediate	1.0% (0.6–1.7)	2.5% (1.8–3.7)
High	5.2% (3.5–7.6)	9.5% (7.0–12.9)
RWS-HCC		

Models	HCC risk at 3 years (%, 95% CI)	HCC risk at 5 years (%, 95% CI)
Low	0.3% (0.1–1.0)	1.2% (0.6–2.2)
High	4.1% (3.0–5.7)	7.4% (5.6–9.6)

Cox-proportional hazard models were used to calculate HCC risk at 3 years and 5 years. AASL-HCC, age, albumin, sex, liver cirrhosis-HCC; AFP, alpha-fetoprotein; APA-B, age, platelet, AFP; CAMD, cirrhosis, age, male sex, and diabetes mellitus; CU-HCC, Chinese University HCC; HCC, hepatocellular carcinoma; HCC-RESCUE, HCC-Risk Estimating Score in CHB patients Under Entecavir; mPAGE-B, modified PAGE-B; REACH-B, Risk estimation for hepatocellular carcinoma in chronic hepatitis B; REAL-B, Real-World Effectiveness From the Asia Pacific Rim Liver Consortium for HBV Risk Score; RWS-HCC, Real-world risk score for hepatocellular carcinoma.

* n.a.: not applicable due to no case in the category.

Table 4.

Discrimination and calibration of 10 HBV-HCC models for prediction of HCC risk bv race/ethnicity.

	Total number	Number of HCC in 3 years	Number of HCC in 5 years	AUC		Hosmer-Lemeshow χ^2	
				3 years (95% Cl)	5 years (95% Cl)	3 years (*p* value)	5 years (*p* value)
White							
REACH-B	956	18	29	0.57 (0.47–0.67)	0.53 (0.45–0.61)	8.17 (0.23)	11.94 (0.06)
PAGE-B	1,162	22	33	0.71 (0.61–0.81)	0.72 (0.64–0.80)	11.56 (0.12)	9.50 (0.22)
mPAGE-B	1,010	20	30	0.75 (0.65–0.85)	0.75 (0.67–0.83)	3.58 (0.73)	4.00 (0.68)
CU-HCC	914	17	28	0.69 (0.63–0.75)	0.70 (0.64–0.76)	17.32 (0.02)	19.50 (<0.01)
HCC-RESCUE	1,226	22	35	0.75 (0.65–0.85)	0.74 (0.66–0.82)	10.02 (0.26)	11.83 (0.16)
CAMD	1,226	22	35	0.76 (0.66–0.86)	0.75 (0.67–0.83)	2.27 (0.89)	4.35 (0.63)
APA-B	673	17	23	0.82 (0.76–0.88)	0.83 (0.77–0.89)	8.23 (0.22)	6.31 (0.39)
REAL-B	714	19	25	0.83 (0.75–0.91)	0.82 (0.74–0.90)	2.69 (0.75)	5.45 (0.36)
AASL-HCC	1,036	20	31	0.86 (0.80–0.92)	0.85 (0.81–0.89)	4.89 (0.77)	10.57 (0.23)
RWS-HCC	748	19	27	0.87 (0.81–0.93)	0.85 (0.79–0.91)	3.68 (0.60)	7.05 (0.22)
African American							
REACH-B	942	20	33	0.63 (0.51–0.75)	0.62 (0.52–0.72)	8.88 (0.35)	15.16 (0.06)
PAGE-B	1,078	22	36	0.69 (0.59–0.79)	0.69 (0.61–0.77)	3.79 (0.71)	5.14 (0.53)
mPAGE-B	994	21	34	0.68 (0.56–0.80)	0.69 (0.59–0.79)	8.35 (0.40)	9.42 (0.31)
CU-HCC	955	20	33	0.83 (0.79–0.87)	0.84 (0.80–0.88)	1.45 (0.98)	2.52 (0.93)
HCC-RESCUE	1,124	23	37	0.79 (0.71–0.87)	0.75 (0.67–0.83)	5.19 (0.74)	8.19 (0.42)
CAMD	1,124	23	37	0.79 (0.71–0.88)	0.75 (0.67–0.83)	5.60 (0.69)	6.84 (0.55)
APA-B*	678	19	32	0.87 (0.79–0.95)	0.83 (0.75–0.91)	3.33 (0.65)	6.57 (0.25)
REAL-B*	733	19	33	0.85 (0.77–0.92)	0.81 (0.73–0.89)	4.24 (0.52)	6.84 (0.23)
AASL-HCC	1,024	21	34	0.85 (0.79–0.91)	0.85 (0.81–0.89)	13.86 (0.05)	21.81 (<0.01)
RWS-HCC*	759	19	33	0.91 (0.85–0.97)	0.89 (0.85–0.93)	6.57 (0.58)	11.97 (0.15)

Area under the curve (AUC) and Hosmer-lemeshow chi-square tests were conducted for each model. AASL-HCC, age, albumin, sex, liver cirrhosis-HCC; AFP, alpha-fetoprotein; APA-B, age, platelet, AFP; CAMD, cirrhosis, age, male sex, and diabetes mellitus; CU-HCC, Chinese University HCC; HCC, hepatocellular carcinoma; HCC-RESCUE, HCC-Risk Estimating Score in CHB patients Under Entecavir; mPAGE-B, modified PAGE-B; REACH-B, Risk estimation for hepatocellular carcinoma in chronic hepatitis B; REAL-B, Real- World Effectiveness From the Asia Pacific Rim Liver Consortium for HBV Risk Score; RWS-HCC, Real-world risk score for hepatocellular carcinoma.

Author Manuscript

Author Manuscript

Author Manuscript

* Highlights the models incorporating AFP variable.