ORIGINAL RESEARCH—CLINICAL

Modeling 5-Year Hepatocellular Carcinoma Risk in Alaska Native Peoples With Hepatitis B Virus Infection

Xiaohong Li,[1] Janet M. Johnston,[2] Chriss Homan,[2] Tracey L. Marsh,[1] Nicole J. Kim,[3] Yongjun Liu,[4] Trang VoPham,[1] William M. Grady,[5] Jorge Mera,[6] Brian J. McMahon,[2] George N. Ioannou,[3] Ziding Feng,[1] and Qianchuan He[1]

[1]Public Health Sciences Division & Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, Washington; [2]Liver Disease and Hepatitis Program, Alaska Native Tribal Health Consortium, Anchorage, Alaska; [3]Division of Gastroenterology, Department of Medicine, Veterans Affairs Puget Sound Health Care System and University of Washington, Seattle, Washington; [4]Department of Laboratory Medicine and Pathology, University of Washington School of Medicine, Seattle, Washington; [5]Translational Science and Therapeutics Division, Public Health Sciences and Clinical Research Divisions, Fred Hutchinson Cancer Center, Division of Gastroenterology, University of Washington School of Medicine, Seattle, Washington; and [6]Infectious Disease Department, Cherokee Nation Health Services, Tahlequah, Oklahoma

BACKGROUND AND AIMS: Modeling hepatocellular carcinoma (HCC) risk in Alaska Native (AN) peoples with chronic hepatitis B virus (HBV) infection is important for risk stratification and surveillance. Existing HCC risk prediction models use baseline characteristics ascertained at the time of HBV diagnosis, rather than predicting HCC risk within 5 years of a relevant time point (such as a clinic visit), and do not include the HBV genotype (GT). We aimed to develop an HCC risk prediction model that addresses these limitations. **METHODS:** We used longitudinal data from a cohort of 1163 AN peoples with HBV. We considered age, sex, GT, serum alpha fetoprotein (AFP), along with serum alanine transaminase, albumin, aspartate aminotransferase, bilirubin, hepatitis B-e-antigen, platelet count, and fibrosis 4 score. To build a 5-year risk model, we structured the longitudinal data into multiple 5-year segments, using AFP as the landmark biomarker. We used the generalized estimation equation approach as well as the Random Forest approach to build risk prediction models. **RESULTS:** Among the 11 predictors included in our final models, AFP was the most important followed by platelet count and GT. Based on cross-validation, the generalized estimation equation model had an area under the receiver operating characteristic curve of 0.81, with 46.5% sensitivity at 90% specificity for 5-year HCC risk prediction. The Random Forest model was superior with an area under the receiver operating characteristic curve of 0.88 and 70% sensitivity at 90% specificity, outperforming the PAGE-B, mPAGE-B, REACH-B and REAL-B models. **CONCLUSION:** We developed an HCC risk prediction model using rich information from different time points in a patient's disease trajectory. Our model can accurately estimate HCC risk at different time points during follow-up for risk stratification and risk-based surveillance.

Keywords: Alaska Native; HBV Genotype; Hepatocellular Carcinoma; Longitudinal Data; Risk Modeling; 5-Year Risk

Introduction

Hepatocellular carcinoma (HCC) is the one of the fastest-rising major malignancies in the United States[1], and hepatitis B virus (HBV) is one of the most important risk factors for HCC.[2–4] Currently, HCC screening is recommended by the American Association for the Study of Liver Diseases in high-risk patients with HBV including men from an endemic country age ≥40 years, women from an endemic country age ≥50 years, persons from Africa, those with a family history of HCC and persons with cirrhosis at any age.[5] Alaska Native (AN) peoples living in Alaska are eligible to receive care through the Alaska Native Tribal Health Consortium (ANTHC) and all known AN peoples with HBV are linked to ANTHC's liver specialty clinic. The ANTHC sends reminder letters every 6 months to AN patients with chronic HBV to have liver function tests, serum alpha fetoprotein (AFP) and liver ultrasound for HCC screening if indicated. Less than 1% of AN persons with HBV have asked to be removed from the every 6-month reminder letters.[6] Modeling HCC risk in AN peoples with HBV is needed to better optimize risk stratification and offer more targeted risk-based surveillance.

A number of published HCC risk prediction models have been developed and are based on different populations; however, none of these models included a significant proportion of AN peoples. The PAGE-B model[7] which includes age, sex and platelet count (PLT) and has an area under the receiver operating characteristic curve (AUC) of 0.73 was developed in a cohort of Caucasian patients with HBV who received entecavir or tenofovir. The mPAGE-B model[8]

Abbreviations used in this paper: **AFP, alpha fetoprotein; ALB, albumin; ALT, alanine transaminase; AN peoples, Alaska native peoples; ANTHC, Alaska native tribal health consortium; AST, aspartate aminotransferase; AUC, area under the receiver operating characteristic curve; FIB-4, fibrosis 4 score; FPF, false positive fraction; GEE, generalized estimating equations; GT, hepatitis B virus genotype; HBeAg, Hepatitis B-e-antigen; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; PLT, platelet count; RF, random forest; TPF, true positive fraction.**

Most current article

2772-5723
https://doi.org/10.1016/j.gastha.2025.100661

includes age, sex, PLT, and serum albumin (ALB) level, and achieved an AUC of 0.82 in Asian patients who received antiviral therapy. Models REACH-B[9] and REAL-B[10,11] have also reported with AUC of 0.796 and 0.81, respectively. Recently, a gradient-boosting machine model was also developed for predicting HCC risk in Korean and Caucasian patients with HBV.[12] This gradient-boosting machine model consists of 10 baseline variables, and the corresponding c-index was 0.79 in the Korean cohort and 0.81 in the Caucasian cohort. In parallel to these studies, some recently developed biomarkers, e.g., AFP-3 and des-gamma-carboxy prothrombin, have been combined with AFP, age, and sex to evaluate their performance as an HCC screening modality in prospective or phase 3 biomarker studies.[13,14]

While AN peoples are the only US-born population with endemic HBV infection, currently, there is no HCC risk model that is specifically developed for AN peoples with HBV and none of the existing models includes the HBV genotype (GT). AN peoples in Alaska have a unique GT prevalence composition. Specifically, 5 different GTs (A2 (13%), B6 (4%, this is pure genotype B without recombination of genotype C), C2 (7%), D2/3 (58%), and F1b (18%)) have been identified in an AN population with 90% racial identity (Yupik and Inupiaq), with significantly different HCC incidence rates associated with different genotypes, especially high in those with genotype F1b[15], which is also more common in AN peoples than other ethnic groups or geographic regions. HBV-infected AN peoples have unique HCC risk factors such as GT (eg genotype F1b), HBV epidemiology, and an extremely high risk of HCC. This suggests a need for an HCC risk model that is specifically developed for AN peoples with HBV.

Additionally, existing models are mostly constructed using baseline characteristics ascertained at the time of HBV diagnosis or entry into the cohort and thus the predicted HCC risk is essentially constant for each person and does not reflect risk factor changes over time. In clinical practice, patients with HBV have regular clinical visits and laboratory tests performed to monitor disease progression. Thus, it is desirable to account for up-to-date laboratory results and predict HCC risk in the medium term, such as within the next 5 years following a clinical visit to plan for appropriate HCC screening strategies accordingly. Furthermore, existing HCC risk models, in general, do not include GT, which is an important factor for HCC risk especially among AN peoples.[15]

Based on a cohort of AN peoples with HBV,[15] this study aimed to develop and internally validate HCC risk prediction models using both Random Forest (RF) and regression methods and compare their performance with that of existing models such as PAGE-B, mPAGE-B, REACH-B and REAL-B. Our models represent comprehensive HCC risk modeling in AN peoples with HBV and provide a powerful tool to predict the HCC risk within 5 years from a given time point. We believe that risk scores from our model can be used for risk stratification and risk-based HCC surveillance, and our methodology could also be emulated for HCC risk model development in other settings.

Methods

Study Population and Data Collection

To develop and validate HCC risk prediction models, we used data from an existing well-characterized population-based cohort of AN peoples with HBV (n = 1,163, 6.7% coinfected with hepatitis C virus), who consented to be in an institutional review board and tribally approved longitudinal study, including 42 patients with incident HCC. The data were collected prospectively between 1983 and 2018. This long-term cohort with over 33,000 person-years of follow-up provides unique opportunities for risk modeling of HCC. This cohort includes incident HCC cases that occurred during prospective follow-up in patients who were undergoing regular HCC surveillance. We considered for inclusion in prediction models the following 11 variables that are known or suspected to be associated with HCC in AN peoples with HBV: age, sex, serum AFP, serum alanine transaminase (ALT), serum ALB, serum aspartate aminotransferase (AST), bilirubin, HBeAg, platelet, fibrosis 4 score (FIB-4) and GT. These characteristics were chosen because (1) they are readily available and objectively determined in our study; and (2) they are important and robust predictors of HCC among AN peoples or they were significant predictors in current models of HCC in patients with HBV that we and others have developed.[16–18] Details of data preprocessing are shown in Supplementary Materials. All research was conducted in accordance with both the Declarations of Helsinki and Istanbul, all research was approved by the appropriate ethics and institutional review committee(s), including the full name of the committees and approval number where possible, and written consent was given in writing by all subjects.

Segmentation of the Longitudinal Data into 5-Year Segments

To build a model for predicting HCC within 5 years following a clinical visit, we segmented the longitudinal data of each patient into multiple 5-year segments. Because AFP was the most frequently measured biomarker, we used the AFP as the landmark biomarker (ie, time 0) for each segment. Around time 0, we searched its forward time-window and backward time-window to identify the values of other risk factors. Based on patients' check-up frequencies, the forward time-window and backward time-window are defined as follows: 1 month forward and 12 months backward for platelets, ALT, AST, bilirubin, and FIB-4 score; 1 year forward and 3 years backward for HBeAg. If multiple values for the same biomarker were available in the forward or backward windows, the one that is closest to time 0 was chosen as the biomarker's value. We used the AST, ALT, and platelet results closest to time 0 to calculate FIB-4. We used a nonparametric imputation method[19] to impute the missing data. For each AFP measurement (ie, time 0 of each segment), if a patient developed HCC after year 1 and within year 5 from time 0, then the HCC outcome is 'yes', and 'no' otherwise. We did not consider segments in which HCC developed within 1 year from time 0, because HCC occurrence would otherwise be too close to biomarker measurements. An illustration of this segmentation process is shown in Figure A1.

Statistical Analysis

Analysis of Risk Factors Associated with HCC

We first conducted univariate analysis to evaluate the association between each risk factor and the 5-year HCC

risk. To accommodate the correlated longitudinal data within the same subject, we fit a generalized estimation equation (GEE) model for each individual risk factor against the HCC outcome (1 or 0) based on the segmented data. These models assess the marginal association for each risk factor. Subsequently, we fit a multivariate GEE model that includes all the risk factors against the HCC outcome based on the segmented data. This model assesses the association of each risk factor while accounting for other risk factors in the model.

Risk Prediction Modeling

We used the GEE approach and the RF approach[20] to conduct 5-year HCC risk prediction. Since there is no external validation dataset for AN people, we used a multiple-data-splitting and cross-validation strategy to internally validate the constructed models. For the GEE approach, the segmented data were randomly split into 8 portions, with 7 portions being the training dataset and 1 portion being the validation dataset. For the training dataset, we fit a multivariate GEE model with the 11 risk factors plus an interaction term between age (young and old) and GT (F or non-F).[15] The yielded regression model was then evaluated on the validation dataset, which is completely independent of the training dataset. Such a data-splitting process was repeated 50 times to obtain the average performance metrics. For the RF approach, we adopted the same data-splitting and cross-validation strategy to build the risk prediction model and internally validate the model. We did not include the age*genotype interaction in the RF model because the RF approach naturally accommodates interactions among predictors. The discrimination performances of the constructed prediction models (GEE or RF) were compared using the area under the receiver operating characteristic curve.

We compared with 4 risk prediction models in the literature: PAGE-B, mPAGE-B, REACH-B, REAL-B .[7–11] For each of these models, we used its marker panel to fit a GEE model in our AN data (which are based on 5-year segments) to test the prediction performance of the fitted model. For the REAL-B model, we did not include 'alcohol use' due to lack of standardized variable. It is worth to point out that the 4 existing models are based on baseline data, while our model is based on time-dependent segments from longitudinal data. Thus, the predicted risks for previous models are with respect to the baseline time, while our predicted risk is with respect to a recent clinical visit during follow-up.

To verify that the HCC risk changes over time, we examined the risk scores predicted from our model for the HCC patients in the following 3 time periods: >10 years before HCC diagnosis, 5–10 years before HCC diagnosis, and 1–4 years before HCC diagnosis. For non-HCC, we examined the predicted risk scores in the following time periods: >10 years before the last clinical visit, 5–10 years before the last clinical visit, and 1–4 years before the last clinical visit. Then, we compared the trend of risk score changes for the HCC patients versus the non-HCC patients.

Table 1. Study Cohort and Patients Characteristics at Baseline of Enrollment

Participant characteristics and HBV genotype at baseline time (n = 1163)	
Age 18–40 (y)	84%(n = 977)
40–60	12.6%(n = 146)
>=60	3.4%(n = 40)
GT (% of F)	19.2%
GT (% of A, B, C & D)	80.8%
Sex (% of female)	44.6%
HCC lab markers at baseline time (mean, standard deviation)	
AFP (ng/mL)	2.0 (1.53)
ALT (U/L)	33.8 (16.2)
ALB (g/dL)	4.1 (0.68)
AST (U/L)	28.9 (11.0)
$HBeAg^+$ (%)	18%
PLT (uL)	286.0 (85.9)
Total bilirubin (mg/dL)	0.5 (0.28)
FIB-4 (score)	0.6 (0.48)

Importance of Individual Predictors

To evaluate the relative contribution of each variable to the prediction model, we used a Permutation Feature Importance method.[20] To assess the importance of the GT, we removed the GT from the RF model and GEE model to evaluate the prediction performance of the reduced model. Since HCC risk prediction can be used for risk-based HCC surveillance, it is useful to evaluate the predictiveness of the constructed models in addition to their classification accuracy. We used the approach by Pepe et al.[21] to plot the predictiveness curve, which provides a comprehensive assessment of a risk model.

Results

Study Cohort and Baseline Patient Characteristics

The study cohort consisted of 1163 AN peoples with HBV, among whom 42 developed HCC during a mean follow-up of 28.1 years. We examined demographics, GT, and baseline values of the 8 laboratory biomarkers, and the results are summarized in Table 1. There were more men (55.4%) than women (44.6%), and the majority were ≤40 years of age (84%) at the time of entry into the cohort. The frequency of GTs A, B, C, D, and F were 15.4%, 3.2%, 7%, 55.2%, and 19.2%, respectively.

Characteristics of the 5-Year Segment Data Based on Longitudinal Follow-Up Data

The longitudinal data of each patient were restructured into multiple 5-year segments, as described in Methods. Applying this segmentation procedure to all the 1163 patients yielded 186 HCC segments and 36,392 non–HCC segments. The values of the laboratory measurements, age, sex, and GT in HCC segments versus non–HCC segments are presented in Table 2. For example, among the 186 HCC segments, the average AFP, ALT, platelet and FIB-4 values were 5.6, 31.4, 228.7, and 1.7, respectively, while among the

Table 2. Summary Statistics of the 5-Y Segments for 11 Risk Factors

Markers	HCC segment (N1 = 186)	Non-HCC segment (N2 = 36,392)
Age	46.4 (14.3)	42.9 (15.4)
Sex(M/F)	40.3%	48.0%
Genotype(F%)	53.3%	19.9%
AFP (ng/mL)	5.6 (1.8)	2.8 (0.6)
ALT (u/L)	31.4 (6.3)	26.3 (6.3)
ALB (g/dL)	4.1 (0.2)	4.3 (0.2)
AST (u/L)	33.0 (5.2)	24.3 (4.2)
$HBeAg^+$ (%)	11.0%	5.6%
PLT (uL)	228.7 (21.5)	283.3 (26.3)
Total bilirubin (mg/dL)	0.5 (0.1)	0.4 (0.1)
FIB-4 (score)	1.7 (0.4)	1.0 (0.2)

36,392 non–HCC segments, these measurements were 2.8, 26.3, 283.3, and 1.0 respectively. Overall, Table 2 shows that, on average, the ALB and PLT are lower in HCC segments than non–HCC segments, whereas AFP, ALT, AST, HBeAg, total bilirubin, and FIB-4 are higher in HCC segments than non–HCC segments.

GEE Analysis of Risk Factors Based on the Segmented Data

Since our goal is to model the 5-year HCC risk at the time of a clinical visit, we used the segmented data for the following risk modeling. We first analyzed the risk factors individually using univariate GEE model. Specifically, we fitted a GEE model for each risk factor individually to evaluate their marginal effects on the HCC outcome. The estimated coefficients of these risk factors are shown in Table 3. Among these risk factors, GT, AFP, ALB, AST, FIB-4 were highly significant ($P < .0001$), and ALT and age were significant ($P < .05$) in univariate analyses. To further examine the joint effects of the risk factors on the HCC outcome, we used a multivariate GEE model to analyze the 11 risk factors. Table 3 shows that the overall trend of the association direction in the univariate model is consistent with that in the multivariate model. Table 3 also shows that the GT and AFP remain to be highly significant in the multivariate GEE analysis.

HCC Risk Prediction Modeling

Prediction Performance of GEE and RF for 5-Year HCC Risk

We first used the GEE approach to build the 5-year HCC risk prediction model. We observed that the mean AUC of the GEE model was 0.81 (see Figure A2). Then, we used the RF approach to build the risk prediction model for HCC, and the RF model had a superior performance, with a mean AUC of 0.88. The receiver operating characteristic curve and the empirical confidence interval of the RF model is shown in Figure 1. The RF model achieved 70% sensitivity at 90% specificity and 80.4% sensitivity at 80% specificity.

Using this dataset, we also compared our model with the PAGE-B model,[7] which includes age, sex and PLT at baseline, as well as the mPAGE-B model,[8] which includes age, sex, PLT, and ALB at baseline. We fit GEE models with our AN data using variables from the PAGE-B, mPAGE-B, REACH-B, or REAL-B models to test their prediction performance, and the AUCs of these models reached 0.62, 0.64, 0.77, and 0.82, respectively (Figure A3 and A4). These AUCs are lower than that of our RF model (0.88).

The risk scores predicted from our model in different time periods for the HCC patients versus the non-HCC patients are presented in Figure A5. As shown, the risk scores

Table 3. Univariate and Multivariate Analysis of Risk Factors Using GEE Model Based on Segmented Data

	Univariate model		Multivariate model	
Biomarkers	Odds ratio (95% confidence interval)	*P* value	Odds ratio (95% confidence interval)	*P* value
Sex	1.37 (0.67–2.81)	.3928	1.16	.72141
GT	4.82 (2.40–9.66)	<.0001	17.46	.00025
Log10 (AFP) (ng/mL)	7.48 (4.15–13.47)	<.0001	6.46	<.0001
ALT (u/L)	1.0 (1.00–1.01)	.0017	1.00	.15816
ALBUMIN (g/dL)	0.29 (0.16–0.54)	.0001	0.42	.03970
AST (u/L)	1.0 (1.00–1.01)	<.0001	0.99	.11023
$HBeAg^+$	2.52 (1.08–5.90)	.0328	2.10	.08090
PLT (uL)	0.99 (0.98–1.00)	.0062	1.00	.29240
Total bilirubin (mg/dL)	1.86 (1.32–2.62)	.0004	1.18	.42504
FIB-4 (score)	1.17 (1.09–1.24)	<.0001	1.07	.20043
Age	2.59 (1.24–5.38)	.0109	8.11	.00478
Age*GT	–	–	0.18	.06081

GT was coded as F or non-F, in the analysis. Age was coded as young or old (see Supplementary Material for details). Age*GT is an interaction term between age and genotype.

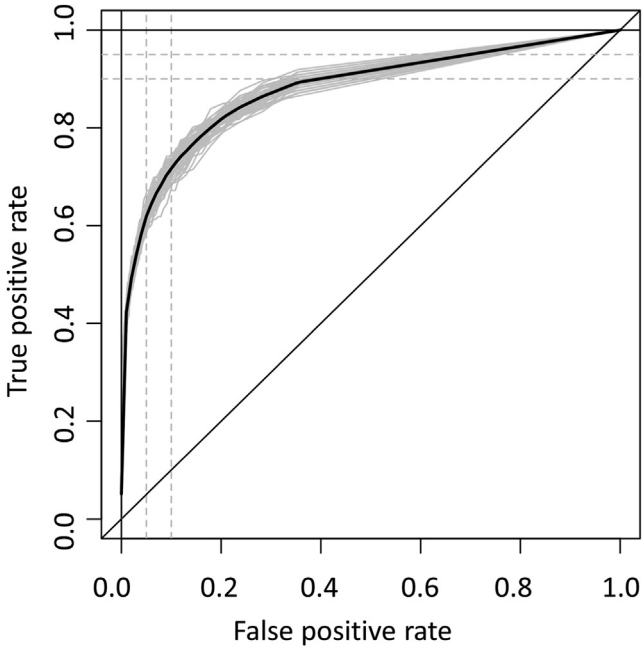

Figure 1. The receiver operating characteristic curve of the RF model for HCC 5-year risk prediction.

of the HCC patients increased substantially over time before the HCC diagnosis, while the risk scores of the non-HCC patients did not show such a pattern of change before their last clinical visit.

GTs Significantly Improved HCC Risk Prediction

Next, we evaluated the importance of each predictor in its contribution to the HCC prediction. Since the RF model showed better performance than the GEE model, we first focused on the RF model. Based on the Permutation Feature Importance approach, we found that AFP was the most important risk factor, followed by PLT and GT (Figure A6). To verify its importance, we removed AFP from our model, and found that the AUC dropped to 0.7. We also observed that if GT was removed, the AUC dropped from 0.88 to 0.83 (see Figure A7). For instance, at a specificity of 90%, the sensitivities with and without the GT were 70% and 63.4%, respectively (see Table 4 and Figure A7 for more details). The GEE model also had decreased AUC when GT was removed (Figure A2 and Table 4). These results demonstrated that GT is an important factor for predicting 5-year risk of HCC in this population.

Predictiveness of the Constructed Risk Prediction Models

We examined the predictiveness of the RF model using the approach by Pepe et al.[21] (Figure 2). For example, when the predicted risk score is 0.02, the risk percentile is 0.94 (94% of the considered population have a risk score lower than 0.02 and 6% have a score higher than 0.02); the corresponding true positive fraction (TPF, a fraction of predicted positive out of total positive cases) is 61%; and the false positive fraction (FPF, a fraction of negative samples predicted as positive out of total negative) is 6%. Similarly, when the predicted risk score is 0.01, 88% of the considered population have their risk score lower than 0.01; the corresponding TPF and FPF are 75% and 12%, respectively. The predictiveness curve for the GEE model is shown in Figure A8, for which the specificities and sensitivities at various thresholds are shown in Table 4. Overall, the predictiveness curve integrates information on specificity, sensitivity, and risk distribution in population, and informs both predictiveness and classification performance.

Discussion

Using a large prospective cohort of AN peoples with HBV infection, we developed an HCC risk prediction model using the RF approach with vigorous cross-validation. The constructed model provides a powerful tool to predict the HCC risk within 5 years from any given time point. This model harnesses rich information from longitudinal data, incorporates GT and other risk factors, and yields highly valuable information for HCC risk stratification and cancer surveillance in AN people.

A major strength of our model is that, unlike many other HCC risk studies, where risk was assessed only using baseline predictors, our model is adaptive and able to reflect dynamic changes in patients' health conditions over time. We constructed 5-year data segments based on each participant's clinical and laboratory test history and then utilized these segments to estimate the HCC risk within 5 years of a clinic visit. Five-year risk prediction is also more clinically relevant for making surveillance decisions. This is different from most existing models, which are focused on baseline measurements and do not update the HCC risk in a longitudinal manner.

Another unique aspect of our study is that we considered GT, which is not commonly assessed in many population-based HCC risk studies. Our results showed that incorporating GT into prediction models can substantially improve the accuracy of the prediction (increasing AUC from 0.83 to 0.88), and we recommend that this variable, if available, be used in HBV-related HCC risk prediction in AN peoples. Prior studies have reported that genotype F can significantly increase the risk of HCC.[22] This might partly explain why the GTs made a significant contribution in the modeling of HCC risk. Our model did not include cirrhosis and HBV-DNA load, as they are correlated with some of the risk factors in our model. For example, in our data, HBV DNA load is correlated with ALT and HBeAg; cirrhosis is correlated with AST, total bilirubin, and FIB-4 score. In fact, when we included cirrhosis or HBV-DNA load into our model, we did not observe any improvement of the AUC. Thus, these 2 factors were not included in our final prediction model for the sake of model simplicity. Additionally, antiviral treatment was not included in our modeling

Table 4. Classification Performance of the GEE and RF Models With and Without GT

	GEE with GT	GEE without GT	RF with GT	RF without GT
Specificity	Sensitivity	Sensitivity	Sensitivity	Sensitivity
90.0%	46.5%	39.5%	70.0%	63.4%
85.0%	56.5%	45.9%	76.1%	68.5%
80.0%	67.6%	52.4%	80.4%	72.5%
75.0%	74.4%	60.4%	84.1%	76.1%
70.0%	80.8%	64.2%	86.9%	79.0%
60.0%	89.3%	72.1%	89.9%	83.7%

because it was confounded by disease severity in our data, ie, patients with more advanced or active disease were more likely to receive treatment, thus making it appear paradoxically as if treatment increased the risk of HCC.

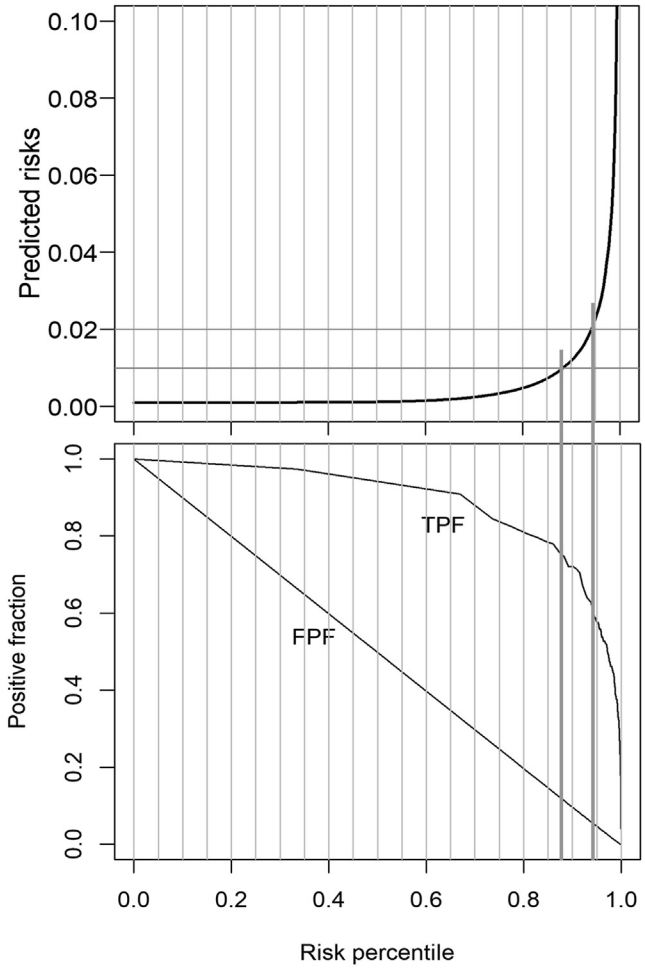

Figure 2. Predictiveness curve of the RF model for predicting 5-year HCC risk. For example, at a predicted risk score of 0.01 (lower horizontal line in upper panel), 88% of samples have risk scores lower than this level, with TPF being 75% and FPF being 12%. At risk score of 0.02 (higher horizontal line in upper panel), 94% of samples have risk scores lower than this level, with TPF being 61% and FPF being 6%.

We also considered 2 approaches: GEE and RF modeling. The GEE approach is easy to implement and can give estimates of risk effects (such as odds ratios) for the predictors. We used GEE in both univariate analysis and multivariate analysis to assess the marginal and joint effects of predictors. These effect estimates for HCC risk are easy to interpret and can be readily used to calculate the HCC risk. RF is a machine learning approach and can be more difficult to properly implement. It does not have simple equations to calculate the risks but can potentially capture nonlinear effects of predictors. Existing literature suggests that the model diversity introduced by the RF approach tends to yield excellent prediction accuracy,[23] and our analysis appears to resonate with those findings. Additionally, we used data-splitting and cross-validation to validate the constructed prediction models. Ideally, one would obtain another independent population-based dataset to validate the findings, but it is practically difficult to find another cohort of AN peoples with HBV and similar longitudinal data as well as clinical, laboratory and genotypic features. Nevertheless, by data-splitting, our training data and validation data are independent of each other, and thus essentially resolve the overfitting issue in risk modeling analysis. Our prediction model uses 8 laboratory markers in addition to age, sex, and GT. Age and sex have been shown to be significant HCC risk factors in population studies.[24] HCC surveillance is currently recommended in all patients with chronic HBV who have cirrhosis, whereas for those without cirrhosis, age, sex, race, and family history are the key factors to determine when surveillance should begin.[5] However, adherence to screening recommendations is low. The risk scores calculated by our model can be used in clinical practice to identify patients at greater risk of developing HCC to improve screening outreach efforts and adherence in these high-risk groups. For example, this model can be potentially used for identifying patients with high-risk scores for intensive outreach to improve adherence to surveillance. As shown in Figure 2, if outreach for improved surveillance adherence is focused on the top 12% of patients with highest risk scores (88% have risk scores lower than this level), the TPF (ie sensitivity) will be 75% and the FPF will be 12%. In other words, by focusing surveillance outreach efforts on only 12% of the population, we can ensure that HCC surveillance takes place in this subset

where 75% are HCC cases; this will significantly reduce the costs and burden of monitoring the patients who are predisposed to HCC. The exact threshold to be chosen for surveillance will require expertise from multiple disciplines, including public health experts and medical providers. We tested the marker panels of PAGE-B, mPAGE-B, Reach-B, and Real-B models in our AN data, and observed that their AUCs were 0.62, 0.64, 0.77, and 0.82, respectively. Built upon the 5-year segments, our 11-marker HCC model achieved substantially better performance with an AUC of 0.88. Overall, our results suggest that the proposed model can significantly improve the performance for HCC prediction in AN peoples with HBV over existing models.

Conclusion

Based on data derived from a longitudinal cohort of AN peoples with HBV, we have developed an HCC risk prediction model that incorporates multiple laboratory markers, GT, and demographic variables. This adaptive model can be utilized to predict 5-year HCC risk for patients with HBV in 5-year intervals and can be used to facilitate risk stratification in HCC surveillance in a population already at high-risk for developing HCC.

Supplementary Material

Material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.gastha.2025.100661.

References

1. Annual report to the nation 2022: overall cancer statistics. https://seer.cancer.gov/report_to_nation/statistics.html#new.2015-2019. Accessed May 1, 2024.
2. Levrero M, Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. J Hepatol 2016;64(1 Suppl):S84–S101.
3. Chaturvedi VK, Singh A, Dubey SK, et al. Molecular mechanistic insight of hepatitis B virus mediated hepatocellular carcinoma. Microb Pathog 2019;128:184–194.
4. de Martel C, Maucort-Boulch D, Plummer M, et al. World-wide relative contribution of hepatitis B and C viruses in hepatocellular carcinoma. Hepatology 2015;62:1190–1200.
5. Singal AG, Llovet JM, Yarchoan M, et al. AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. Hepatology 2023;78(6):1922–1965.
6. McMahon BJ, Bulkow L, Harpster A, et al. Screening for hepatocellular carcinoma in Alaska natives infected with chronic hepatitis B: a 16-year population-based study. Hepatology 2000;32:842–846.
7. Papatheodoridis G, Dalekos G, Sypsa V, et al. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. J Hepatology 2016;64:800–806.
8. Kim JH, Kim YD, Lee M, et al. Modified PAGE-B score predicts the risk of hepatocellular carcinoma in Asians with chronic hepatitis B on antiviral therapy. J Hepatology 2018;69:1066–1073.
9. Yang HI, Yuen MF, Chan HL, et al. , REACH-B Working Group. Risk estimation for hepatocellular carcinoma in chronic hepatitis B (REACH-B): development and validation of a predictive score. Lancet Oncol 2011;12(6):568–574.
10. Yang HI, Yeh ML, Wong GL, 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(3):389–399.
11. Wu S, Zeng N, Sun F, et al. Hepatocellular carcinoma prediction models in chronic hepatitis B: a systematic review of 14 models and external validation. Clin Gastroenterol Hepatol 2021;19(12):2499–2513.
12. Kim HY, Lampertico P, Nam JY, et al. An artificial intelligence model to predict hepatocellular carcinoma risk in Korean and Caucasian patients with chronic hepatitis B. J Hepatol 2022;76:311–318.
13. Singal AG, Tayob N, Mehta A, et al. GALAD demonstrates high sensitivity for HCC surveillance in a cohort of patients with cirrhosis. Hepatology 2022;75:541–549.
14. Tayob N, Kanwal F, Alsarraj A, et al. The performance of AFP, AFP-3, DCP as biomarkers for detection of hepatocellular carcinoma (HCC): a phase 3 biomarker study in the United States. Clin Gastroenterol Hepatol 2023;21:415–423.
15. McMahon BJ, Nolen LD, Snowball M, et al. HBV genotype: a significant risk factor in determining which patients with chronic HBV infection should undergo surveillance for HCC: the hepatitis B Alaska study. Hepatology 2021;74:2965–2973.
16. Lok AS. Hepatitis B: liver fibrosis and hepatocellular carcinoma. Gastroenterol Clin Biol 2009;33:911–915.
17. Su F, Weiss NS, Beste LA, et al. Screening is associated with a lower risk of hepatocellular carcinoma-related mortality in patients with chronic hepatitis B. J Hepatol 2021;74:850–859.
18. McMahon B, Block J, Block T, et al. Hepatitis-associated liver cancer: gaps and opportunities to improve care. J Natl Cancer Inst 2015;108:1–6.
19. Stekhoven DJ, Buehlmann P. MissForest - nonparametric missing value imputation for mixed-type data. Bioinformatics 2012;28:112–118.
20. Breiman L. Random forests. Mach Learn 2001;45:5–32.
21. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. Am J Epidemiol 2008;167:362–368.
22. Sánchez-Tapias JM, Costa J, Mas A, et al. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. Gastroenterology 2002;123:1848–1856.
23. Fawagreh K, Gaber MM, Elyan E. Random forests: from early developments to recent advancements. Syst Sci Control Eng An Open Access J 2014;2:602–609.
24. Lee CM, Lu SN, Changchien CS, et al. 1999. Age, sex, and local geographic variations of viral etiology of hepatocellular carcinoma in a hyperendemic area for hepatitis B virus infection. Cancer 1999;86(7):1143–1150.

Received June 3, 2024. Accepted March 18, 2025.

Correspondence:
Address correspondence to: Qianchuan He, PhD, Public Health Sciences Division, Fred Hutchinson Cancer Center, 1100 Fairview Ave N, Seattle, WA 98109. e-mail: qhe@fredhutch.org.

Acknowledgments:
The authors wish to thank the Alaska Native peoples who participated in this study.

Authors' Contributions:
Xiaohong Li: Data curation; Formal analysis; Methodology; Writing – original draft; Writing – review & editing. Janet M. Johnston: Data curation; Writing – review & editing. Chriss Homan: Data curation; Writing – review & editing. Tracey L Marsh: Writing – review & editing. Nicole Kim: Writing – review & editing. Yongjun Liu: Writing – review & editing. Trang VoPham: Writing – review & editing. William M Grady: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing – review & editing. Jorge Mera: Writing – review & editing. Brian J McMahon: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing – review & editing. George N. Ioannou: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing – review & editing. Ziding Feng: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing – review & editing. Qianchuan He: Conceptualization; Formal analysis; Funding acquisition; Investigation; Methodology; Visualization; Writing – review & editing; Corresponding Author

Conflicts of Interest:
These authors disclose the following: William M Grady is a consultant for Guardant Health, Diacarta, Karius and Freenome and receives research support from Lucid Diagnostics. Nicole J. Kim receives Clinical trial funding (to the institution) from Madrigal Pharmaceuticals. The remaining authors disclose no conflicts.

Funding:
This study is partly supported by the National Institutes of Health Fund P20CA252732, R01CA223498, and R01GM105785.

Ethical Statement:
The study was approved by the Institutional Review Board of the University of Washington and the Alaska Area Native Health Service Research and Publications Committee, the Anchorage Native Health Board, the Alaska Area Institutional Review Board, the Alaska Native Tribal Health Consortium, and Southcentral Foundation.

Data Transparency Statement:
The analytic methods can be made available to other researchers. The data and materials are owned by the tribes and cannot be made available for others.

Reporting Guidelines:
Declaration of Helsinki.