OncoTargets and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Development and Validation of a Machine Learning-Based Model Used for Predicting Hepatocellular Carcinoma Risk in Patients with Hepatitis B-Related Cirrhosis: A Retrospective Study

Yixin Hou[1,*], Jianguo Yan[2,*], Ke Shi[1,3,*], Xiaoli Liu[1], Fangyuan Gao[1], Tong Wu[1], Peipei Meng[1], Min Zhang[2], Yuyong Jiang[1], Xianbo Wang[1]

[1]Center of Integrative Medicine, Beijing Ditan Hospital, Capital Medical University, Beijing, 100015, People's Republic of China; [2]People's Liberation Army Fifth Medical Center, Beijing, 100039, People's Republic of China; [3]Department of Gastroenterology, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, 100700, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xianbo Wang, Center of Integrative Medicine, Beijing Ditan Hospital, Capital Medical University, No. 8 Jing Shun East Street, Beijing, 100015, People's Republic of China, Email wangxb@ccmu.edu.cn; Min Zhang, People's Liberation Army Fifth Medical Center, Beijing, 100039, People's Republic of China, Email gcmw2001@163.com

Object: Our objective was to estimate the 5-year cumulative risk of HCC in patients with HBC by utilizing an artificial neural network (ANN).

Methods: We conducted this study with 1589 patients hospitalized at Beijing Ditan Hospital of Capital Medical University and People's Liberation Army Fifth Medical Center. The training cohort consisted of 913 subjects from Beijing Ditan Hospital of Capital Medical University, while the validation cohort comprised 676 subjects from People's Liberation Army Fifth Medical Center. Through univariate analysis, we identified factors that independently influenced the occurrence of HCC, which were then used to develop the ANN model. To evaluate the ANN model, we assessed its predictive accuracy, discriminative ability, and clinical net benefit using metrics such as the area under the receiver operating characteristic curve (AUC), concordance index (C-index), calibration curves.

Results: In total, we included nine independent risk factors in the development of the ANN model. Remarkably, the AUC of the ANN model was 0.880, significantly outperforming the AUC values of other existing models including mPAGE-B (0.719) (95% CI 0.670–0.768), PAGE-B (0. 710) (95% CI 0.660–0.759), FIB-4 (0.693) (95% CI 0.640–0.745), and Toronto hepatoma risk index (THRI) (0.705) (95% CI 0.654–0.756) ($p<0.001$ for all). The ANN model effectively stratified patients into low, medium, and high-risk groups based on their 5-year In the training cohort, the positive predictive value (PPV) for low-risk patients was 26.2% (95% CI 25.0–27.4), and the negative predictive value (NPV) was 98.7% (95% CI 95.2–99.7). For high-risk patients, the PPV was 54.7% (95% CI 48.6–60.7), and the NPV was 91.6% (95% CI 89.4–93.4). These findings were validated in the independent validation cohort.

Conclusion: The ANNs model has good individualized prediction performance and may be helpful to evaluate the probability of the 5-year risk of HCC in patients with HBC.

Keywords: machine learning-based model, hepatocellular carcinoma, risk, hepatitis B-related cirrhosis

Introduction

Hepatocellular carcinoma (HCC) was the third leading cause of cancer mortality.[1] Hepatitis B-related cirrhosis (HBC) is a well-known risk factor for HCC and accounts for 50%-80% of HCC cases worldwide.[2,3] Given the increasing incidence of HCC related to HBC, managing patients with HBC in an ambulatory setting is crucial. Although the current first-line oral nucleos(t)ide analogues (NAs), such as entecavir (ETV) or tenofovir disoproxil fumarate (TDF), can effectively inhibit hepatitis B virus (HBV) replication, HCC can still develop in chronic hepatitis B (CHB) patients undergoing NA treatment.[2,4,5] A 10-year follow-up study

Received: 13 October 2023
Accepted: 8 March 2024
Published: 23 March 2024

conducted in Europe revealed the occurrence of new HCC cases even after more than 5 years of NA treatment.[6] Thus, it is essential to stratify the risk of HCC in patients with HBC.

Evidence suggests that large-scale community-wide ultrasonography (USG) screening for HCC can significantly reduce HCC-related mortality[2,7] According to a updated meta-analysis about 24% of the patients with cirrhosis adhered to regular HCC screening prior to their HCC diagnoses.[8] Current expert guidelines recommend HCC screening for all patients with HBC, including chronic carriers of HBV, particularly Asian men aged 40 years and women aged 50 years.[9] However, even with these guidelines, a considerable number of individuals at risk below these age limits still develop HCC.[10] Consequently, there is a need for an stratify HCC risks prediction model specifically designed for HBC patients to classify the population into different risk categories. The implications of the HCC risk stratification model are twofold. Firstly, the current prediction model lacks a hierarchical structure, which hinders effective clinical stratification management. Secondly, incorporating hierarchical models can help identify low-risk populations, leading to reduced need for surveillance and consequently lowering economic costs and preventing wastage of medical resources.

Several HCC risk models have been developed for patients diagnosed with chronic hepatitis B (CHB). Although these predictive models exhibit excellent performance, they lack the ability to stratify HCC risks in patients with HBC. Early scoring systems primarily originated from untreated patients in Asia and include factors such as age, sex, HBV DNA levels, core promoter mutations, and cirrhosis-HCC (GAG-HCC)[11] In previous studies, it has been observed that these scores show moderate ability in predicting HCC among Caucasian patients treated with NA, but the performance of HCC risk scores was low in Asian patients., with sensitivity ranging from 18% to 73%, and c-statistics from 0.71 to 0.85[12,13] In order to predict the risk of HCC for up to 5 years in Caucasian patients on entecavir (ETV) and tenofovir disoproxil fumarate (TDF), a simple risk score called PAGE-B was developed based on age, sex, and platelet count.[14] Although it can estimate the occurrence of liver cancer within 5 years, it is not suitable for categorizing the risk levels. Korean researchers made slight modifications to the PAGE-B score by incorporating serum albumin (ALB) and adjusting the weighting of age, gender, and platelet count, resulting in mPAGE-B.[15] However, the accuracy of mPAGE-B is still limited. A recent meta-analysis demonstrated that a fibrosis index known as FIB-4 showed a stronger correlation with the severity of fibrosis compared to other available tests.[16] Nonetheless, additional longitudinal cohort studies are necessary to assess whether the FIB-4 index can be utilized to stratify HCC risks and complement current criteria for identifying patients with the lowest HCC risk. A score known as THRI (age, sex, etiology, and platelets) might provide guidance for HCC surveillance recommendations in patients with cirrhosis.[17] However, as with most other studies focusing on the development of risk prediction models, various HCC risk models lack the ability to stratify predictions.

Artificial neural networks (ANNs), a form of machine learning, are mathematical models inspired by the biological nervous system.[18] ANNs have been extensively employed in medical decision-making due to their capability to handle statistical analyses of linear, logistic, or nonlinear complex relationships.[19] The objective of this research was to employ an ANN in constructing an early-stage warning model to predict HCC in patients with HBC, specifically targeting the identification of high-risk groups that may experience disease progression

Materials and Methods

Patients

We recruited a total of 2466 individuals who were diagnosed with first-diagnosed HBC at the Beijing Ditan Hospital of Capital Medical University (Beijing, China) and the People's Liberation Army Fifth Medical Center between January 2011 and January 2016. To be included in the study, participants had to meet the following criteria: (1) They were diagnosed with HBC for the first time. (2) Their age ranged from 18 to 75 years. (3) They had been receiving antiviral treatment with ETV or TDF for at least one year. On the other hand, individuals were excluded if they met any of the following criteria: (1) Their age was below 18 or above 75. (2) They were infected with other forms of hepatitis (including A, C, D, and E) and the human immunodeficiency virus. (3) They were diagnosed with HCC prior to the index date. (4) They had liver failure or had undergone liver transplantation. (5) They had active alcoholism or severe fatty liver. (6) They either died within the five-year follow-up period or were lost to follow-up. In total, the study included 1589 patients who had compensated and decompensated cirrhosis with HBC. The training cohort consisted of 913 subjects from Beijing Ditan Hospital of Capital Medical University, while the validation cohort comprised 676 subjects from the People's Liberation Army Fifth Medical Center (Figure 1). The Ethics

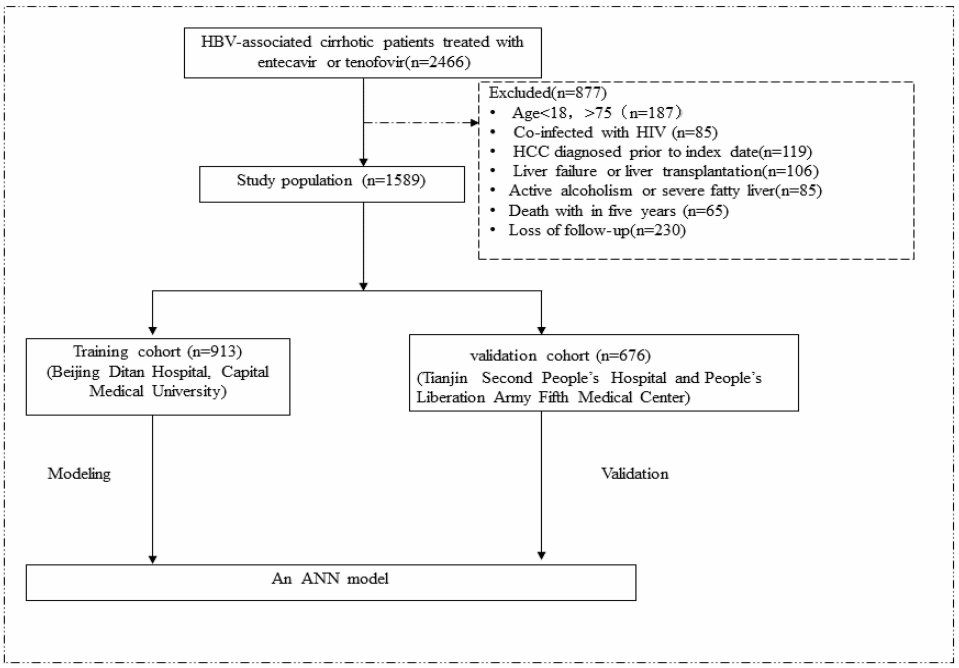

Figure 1 Outline of the recruitment and grouping of Hepatitis B-related cirrhosis cohort study.

Committee of Beijing Ditan Hospital approved this study, and since it was retrospective in nature, written informed consent was not obtained from all patients. To ensure patient privacy, their records and information were anonymized and de-identified before the analysis.

Clinical Definition and Follow-Up

Compensatory cirrhosis resulting from chronic hepatitis B was diagnosed by fulfilling the following criteria: (1) HBsAg positivity for a duration exceeding half a year; (2) confirmation through liver biopsy which showed a Metavir fibrosis score of F4; (2) Detection of Esophageal varices during endoscopy, ruling out non-cirrhotic portal hypertension; (3) In cases lacking histology or endoscopy, two of the four criteria below must be met: ① ultrasonography, computed tomography (CT), or magnetic resonance imaging (MRI) indicating morphological changes in the liver, such as nodules in the hepatic parenchyma and irregularities on the liver surface; ② platelet count below 100×109 cells/L with no other plausible causes; ③ ALB level below 35.0 g/L, international normalized ratio surpassing 1.3, or prothrombin time extension of more than 3 seconds; ④LSM >12.4 kPa when alanine amino transferase (ALT) levels are considered. The diagnosis of decompensated cirrhosis was premised on the presence of portal and venous hypertension complications, and/or impaired liver function using the following guidelines: (1) Confirmation of cirrhosis; (2) Presence of complications associated with portal hypertension, such as ascites, bleeding in the esophagus and stomach due to varices, and hepatic encephalopathy (HE)[20] Detection of hepatocellular carcinoma (HCC) relied on the use of imaging techniques encompassing multiple detector CT scans or dynamic contrast-enhanced MRI. Diagnosis could be conclusively made if the typical vascular indicators of HCC (marked increase in blood vessels during the arterial phase followed by rapid decline during the portal venous or delayed phase) were observed in a nodule exceeding 1 cm in diameter using either of these two methods.[21] Virological response (VR) was defined as the absence of detectable HBV DNA load either at the time of HCC development or at the conclusion of the follow-up period. The reference date for this study was defined as the date of the initial cirrhosis diagnosis at the hospital.

All patients underwent CT, MRI, ultrasonography, or serum AFP tests every 3 months. When the serum AFP level of the patient increased or new intrahepatic nodules were found by ultrasonography, dynamic CT or MRI was used to determine whether the HCC had happened The endpoint of the study was either the occurrence of the first HCC diagnosis or the conclusion of the five-year follow-up period. Clinical data encompassed demographic information (age, sex, and family history of HCC), as well as complications (bleeding from gastrointestinal varices, ascites, and HE) The biochemical

parameters analyzed in this study included alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TBIL), albumin (ALB), γ-glutamyl transpeptidase (GGT), white blood cell count (WBC), neutrophil count (NC), lymphocyte count (LC), platelet count (PLT), creatinine (CREA), prothrombin time (PT), international normalized ratio (INR), alpha-fetoprotein (AFP), and hepatitis B virus (HBV) DNA. To monitor any changes over time, routine laboratory tests including AFP levels as well as radiological examinations were conducted every 3–6 months.

Statistical Analysis

Data analysis was conducted using SPSS 22.0 and R version 3.3.2 (R core development team, 2010). The presentation of data was done as median (range) or n (%), depending on the nature of the data. Statistical significance of differences among continuous and categorical variables was assessed using Student's *t*-test (or Mann–Whitney test, if appropriate) or a chi-squared test (or Fisher's exact test, if appropriate). We presented hazard ratios (HRs) and their 95% confidence intervals (CIs), along with corresponding p values. To construct the ANNs model, we utilized Mathematica 11.1.1 for Microsoft windows (64-bit) and considered factors related to the 5-year development of HCC in patients with HBC. To assess discrimination performance, we utilized receiver operating characteristic (ROC) curves and computed the area under the ROC (AUROC) curve, which generated the Harrell's c-index. In addition, we compared the performance of the ANN model with well-established scores such as mPAGE-B, PAGE-B, FIB-4, and THRI scores in the ROC curves.[15,22–24] The scores were calculated using a previously published scoring formula. For graphical evaluation of agreement between the predicted 5-year probability of HCC-free status and observed probability, we utilized a calibration plot. In all statistical tests, a p value of less than 0.05 was deemed to indicate a statistically significant difference.

Results

Baseline Characteristics

To conduct our research, a total of 913 patients were recruited for the training set, while the validation set consisted of 676 patients. The baseline data of the enrolled patients can be found in Table 1. Within the training set, a majority of the patients, 630 individuals (69.0%), were male, with a median age of 49 years (interquartile range, 40–58 years). In terms of achieving VR, the numbers were 91% and 93.5% for the first year, and 99% and 98% for the fifth year, in the training and validation sets respectively. It's worth noting that no significant differences were observed in the baseline characteristics between the two sets. Throughout the 5-year follow-up period, a diagnosis of HCC was made in 182 patients (19.9% of the training cohort) and 108

Table 1 Basic Clinical Characteristics of Patients with HBV-Related Cirrhosis

Variables	All Patients (n=1589)	Training Cohort (n = 913)	Validation Cohort (n =676)	*P* value
Age, years	49.0 (41.0–58.0)	49.0 (40.0–58.0)	49.0 (43.0–57.0)	0.466
Male sex, n (%)	1076 (67.7)	630(69.0)	446(66.0)	0.120
Family history of HCC, n (%)	169 (10.6)	71 (7.8)	98 (14.5)	0.277
Smoking, n (%)	365(23.0)	214(23.4)	151 (22.3)	0.856
Alcohol consumption, n (%)	338(21.3)	173 (18.9)	165 (24.4)	0.599
Diabetes, n (%)	258 (16.2)	130(14.2)	128 (18.9)	0.074
Hypertension, n (%)	229 (14.4)	142 (15.6)	87 (12.9)	0.935
Ascites, n (%)	337(21.2)	190 (20.8)	147 (21.7)	0.385
Encephalopathy, n (%)	71(4.5)	43 (4.7)	28 (4.1)	0.358
Gastrointestinal varices with bleeding, n (%)	132(14.8)	88(13.8)	44(16.0)	0.338
HBeAg positivity, n (%)	827 (52.0)	481 (52.7)	346 (51.2)	0.631
CTP score	7.0 (5.0–10.0)	7.0 (6.0–9.0)	7.0 (5.0–10.0)	0.726
MELD score	10.1 (7.8–12.9)	10.1 (8.0–12.6)	10.1 (7.7–13.6)	0.951
Alanine aminotransferase, U/L	41.3 (26.5–105.8)	43.6 (27.1–112.4)	37.5 (25.1–95.5)	0.064
Aspartate aminotransferase, U/L	47.9 (30.8–106.1)	48.6 (32.1–107.9)	46.7 (29.7–98.1)	0.182

(*Continued*)

Table 1 (Continued).

Variables	All Patients (n=1589)	Training Cohort (n = 913)	Validation Cohort (n =676)	*P* value
Total bilirubin, μmol/L	22.4 (14.2–39.2)	22.7 (14.1–38.5)	22.3 (14.5–40.7)	0.905
Albumin, g/L	33.6 (30.3–39.2)	33.8 (30.4–40.1)	33.2 (30.1–38.4)	0.248
Gamma-glutamyl transpeptidase, U/L	56.3 (37.2–97.8)	55.3 (36.1–99.7)	57.3 (40.4–95.6)	0.487
White blood cell count, $\times10^9$/L	3.9 (2.8–5.3)	3.9 (2.8–5.3)	3.8 (2.7–5.3)	0.441
Neutrophil count, $\times10^9$/L	2.2 (1.5–3.1)	2.2 (1.5–3.2)	2.2 (1.5–3.0)	0.334
Lymphocyte count, $\times10^9$/L	1.1 (0.8–1.6)	1.1 (0.8–1.6)	1.1 (0.7–1.6)	0.821
Neutrophil-lymphocyte ratio	2.0 (1.4–2.7)	2.0 (1.4–2.8)	1.9 (1.4–2.6)	0.498
Platelets, $\times10^9$/L	87.0 (65.8–118.6)	87.0 (65.0–116.0)	89.0 (67.0–121.0)	0.099
Creatinine, μmol/L	66.0 (56.0–76.0)	66.1 (56.1–76.2)	65.0 (56.0–73.9)	0.502
Blood urea nitrogen, mmol/L	5.1(4.0–6.7)	5.1(4.0–6.7)	5.2 (4.1–6.7)	0.617
Prothrombin time, s	14.3 (12.7–16.3)	14.3 (12.8–16.1)	14.1 (12.5–16.7)	0.901
Prothrombin activity, %	67.0(54.0–80.0)	67.0(54.0–80.0)	67.0(53.0–81.0)	0.844
International normalized ratio	1.2 (1.1–1.3)	1.2 (1.1–1.3)	1.2 (1.1–1.4)	0.965
Alpha-fetoprotein, ng/mL	7.1 (3.4–30.1)	7.2 (3.4–30.5)	6.7 (3.3–26.0)	0.275
HBV DNA, $\log_{10}$IU/mL	3.5 (2.7–5.6)	3.5 (2.7–5.7)	3.5 (2.7–5.4)	0.993
5-year hepatocellular carcinoma, n (%)	290(18.3)	182(19.9)	108(16.0)	0.162

Abbreviations: HBV, hepatitis B virus; CTP, Child-Turcotte-Pugh; MELD, model of end-stage liver disease; HBeAg, hepatitis B e antigen.

patients (16.0% of the validation cohort). The cumulative incidences of HCC at 1, 3, and 5 years were 3.9%, 14.7%, and 19.9% respectively in the training cohort, and 3.3%, 13.1%, and 16.0% respectively in the validation cohort (p = 0.184, Table 1)

Construction of ANN Model

The results obtained from the Cox regression analysis are presented in Table 2. The analysis revealed significant associations between various factors and the appearance of HCC. Age (HR = 1.045, 95% CI 1.029–1.061, p < 0.001), ascites (HR = 1.935, 95% CI 1.354–2.766, p < 0.001), HE (HR = 2.513, 95% CI 1.274–4.956, p = 0.008), gastrointestinal varices with bleeding (HR = 2.153, 95% CI 1.420–3.265, p < 0.001), Child-Pugh classification (HR = 1.838, 95% CI 01.173–2.879, p = 0.008), ALT (HR = 0.998, 95% CI 0.996–0.999, p = 0.004), ALB (HR = 0.972, 95% CI 0.947–0.997, p = 0.028), NLR (HR = 1.331, 95% CI 1.262–1.403, p < 0.001), and PLT (HR = 0.979, 95% CI 0.972–0.985, p < 0.001)

Table 2 Factors Associated with Prediction Incidence of Liver Cancer

Variables	Univariate Analysis		*P* value
	β	HR (95% CI)	
Age (yr)	0.044	1.045 (1.029,1.061)	< 0.001
Sex(male)	0.267	1.306 (0.873,1.954)	0.194
Family history of HCC	0.392	1.480 (0.842,2.603)	0.173
Smoking	−0.081	0.922 (0.605,1.405)	0.706
Alcohol consumption	0.161	1.175 (0.767,1.800)	0.458
Diabetes	0.255	1.291 (0.807,2.064)	0.287
Ascites	0.660	1.935 (1.354,2.766)	< 0.001
Hepatic encephalopathy			
0	Reference		
I–II	0.921	2.513 (1.274,4.956)	0.008
III–IV	1.831	6.242 (1.539,25.318)	< 0.001
Gastrointestinal varices with bleeding	0.767	2.153 (1.420,3.265)	< 0.001

(*Continued*)

Table 2 (Continued).

Variables	Univariate Analysis		*P* value
	β	HR (95% CI)	
Child-Pugh classification			
A (5–6)	Reference		
B (7–9)	0.362	1.435 (0.932,2.212)	0.101
C (10–13)	0.618	1.838 (1.173,2.879)	0.008
MELD score	0.002	1.002 (0.961,1.044)	0.924
Alanine aminotransferase (U/L)	−0.002	0.998 (0.996,0.999)	0.004
Aspartate aminotransferase (U/L)	−0.002	0.998 (0.996,0.999)	0.023
Total bilirubin (mg/dl)	−0.003	0.997 (0.992,1.001)	0.114
Albumin (g/L)	−0.029	0.972 (0.947,0.997)	0.028
gamma-glutamyl transpeptidase (U/L)	0.001	1.001 (0.998,1.003)	0.637
White blood cell count ($\times 10^9$/L)	−0.007	0.993 (0.913,1.079)	0.861
Neutrophil count ($\times 10^9$/L)	0.090	1.095 (1.016,1.179)	0.017
Lymphocyte count ($\times 10^9$/L)	−1.400	0.247 (0.162,0.374)	< 0.001
NLR	0.286	1.331 (1.262,1.403)	< 0.001
Platelets ($\times 10^9$/L)	−0.022	0.979 (0.972,0.985)	< 0.001
Creatinine (μmol/L)	0.001	1.001 (0.995,1.007)	0.758
International normalized ratio	0.559	1.749 (0.936,3.268)	0.080
HBeAg positivity	0.224	1.251 (0.878,1.785)	0.216
Alpha-fetoprotein (ng/mL)	−0.002	0.998 (0.996,1.000)	0.063
HBV DNA ($\log_{10}$IU/mL)	0.013	1.013 (0.927,1.107)	0.778

Abbreviations: OR, odds ratio; 95% CI, 95% confidence interval; HBV, hepatitis B virus; MELD model of end-stage liver disease; HBeAg, hepatitis B e antigen.

were all found to have significant associations with HCC appearance in the training group. These factors were also incorporated in constructing the artificial neural network (ANN) model (https://lixuan.me/annmodel/myg-v4/). The most commonly used structure for ANN is the multilayer perceptron (MLP), which consists of input, hidden, and output layers. The input layer comprises clinical and biochemical parameters, while the output layer includes corresponding prognosis outcomes.[10] The ANN model for predicting the 5-year risk of HCC development in patients with HBC is provided (https://lixuan.me/annmodel/myg-v4/). Neurons in the model are interconnected via weighted links, resulting in a total of 14 input neurons and two output neurons. To enhance the MLP's performance, we implemented four hidden layers after numerous rounds of debugging and testing

Application of the ANN Model for Risk Stratification

According to the ANN model scores, we classified all patients into three strata: Strata 1, which represents low-risk patients; Strata 2, which represents medium-risk patients; and Strata 3, which represents high-risk patients. In the training cohort, we compared Strata 2 and Strata 3 to Strata 1 as the reference group. The hazard ratios (HRs) for Strata 2 and Strata 3 were 0.48 (95% CI 0.43–0.82) and 23.5 (95% CI 19.11–26.82) ($p < 0.0001$) respectively. We observed similar survival differences for all stratifications in the validation cohort. Therefore, the ANN model effectively differentiated patients with different risk levels in both the training and validation cohorts. In the training cohort, the positive predictive value (PPV) for low-risk patients was 26.2% (95% CI 25.0–27.4), and the negative predictive value (NPV) was 98.7% (95% CI 95.2–99.7). For high-risk patients, the PPV was 54.7% (95% CI 48.6–60.7), and the NPV was 91.6% (95% CI 89.4–93.4). Similar results were obtained in the validation cohort, with a low-risk PPV of 20.9% (95% CI 19.6–22.2) and an NPV of 100% (95% CI 99.6–100). The high-risk PPV was 41.5% (95% CI 32.8–50.8), and the NPV was 91.9% (95% CI 88.6–94.3) (Table 3). We then utilized the patient information in the ANN model to assess the 5-year risk of developing liver cancer and classify patients accordingly (Figure 2). In the training set, 228 (25.0%) patients were classified as low-risk, 456 (49.9%) as intermediate-risk, and 229 (25.1%) as high-risk. The cumulative probabilities of hepatocellular carcinoma (HCC) development over 5 years were 1.3%, 9.4%, and

Table 3 Positive Predictive and Negative Predictive Values

Cohort	Models	5-Year Risk of HCC	
		Positive(%)(95% CI)	Negative(%) (95% CI)
Training	ANN (low)	26.2(25.0–27.4)	98.7(95.2–99.7)
	ANN(high)	54.7(48.6–60.7)	91.6(89.4–93.4)
Validation	ANN (low)	20.9 (19.6–22.2)	100(99.6–100)
	ANN(high)	41.5(32.8–50.8)	91.9 (88.6–94.3)

51.5% in the low-, intermediate-, and high-ANN model groups, respectively ($p < 0.001$; Figure 3). Similarly, in the validation set, the cumulative probabilities of HCC development over 5 years were 0%, 11.9%, and 41.8% in the low-, intermediate-, and high-ANN model groups, respectively ($p < 0.001$; Figure 3)

Discrimination and Calibration of the ANN Model

The ANN model showed superior performance in predicting HCC occurrence compared to other models in both the training and validation cohorts. In the training cohort, the AUROC curve value of the ANN model was 0.880 (95% CI 0.847–0.914), and the C-index value was 0.843 (95% CI 0.812–0.874) (Table 4). These values were significantly higher than those of the mPAGE-B, PAGE-B, THRI, and FIB-4 models ($p < 0.001$). Similarly, in the validation cohort, the ANN model demonstrated excellent predictability with an AUROC curve value of 0.824 (95% CI 0.757–0.884) and a C-index value of 0.809 (95% CI 0.761–0.887), which were significantly higher than the values obtained from the mPAGE-B, PAGE-B, THRI, and FIB-4 models ($p < 0.001$). Moreover, when compared to the mPAGE-B, PAGE-B, THRI, and FIB-4 models, the ANN model exhibited a significant net benefit in both the training and validation cohorts (Figure 4). The calibration curves for the ANN model demonstrated good agreement between the predicted HCC-free probability and the observed probability in 5 years in both the training (Figure 5A) and validation (Figure 5B) cohorts. This finding highlights the superior clinical practicability of the ANN model over other models

Discussion

For the first time in this study, we constructed an artificial neural network (ANN) prediction model using machine learning. This model is specifically designed to calculate the likelihood of hepatocellular carcinoma (HCC) within a 5-year period for individual patients. It combines patient demographics and laboratory markers, including age, hepatocellular encephalopathy (HE), gastrointestinal varices with bleeding, Child-Pugh classification, alanine aminotransferase (ALT) level, albumin (ALB) level, neutrophil-to-lymphocyte ratio (NLR), and platelet count (PLT). Our ANN model categorizes patients into low-, medium-, and high-risk groups based on estimated HCC risk. Implementing HCC surveillance strategies guided by our model's risk estimates has shown greater overall benefit compared to previous models.

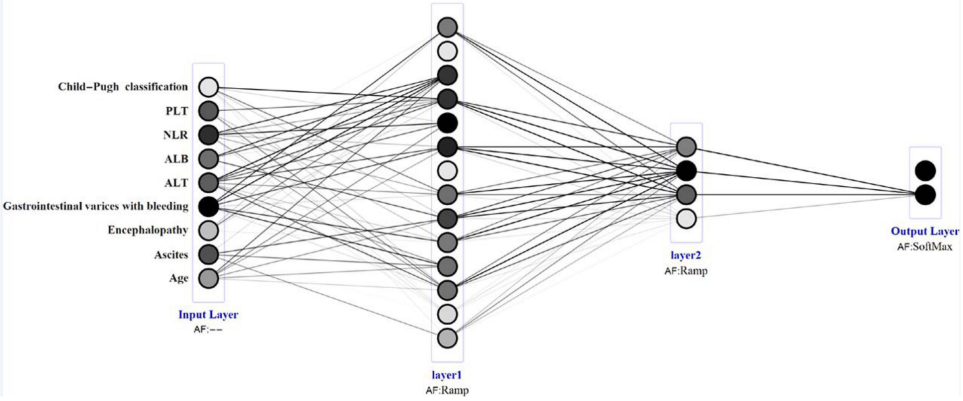

Figure 2 The ANNs model predicts probability of 5-year risk of HCC development. visit.

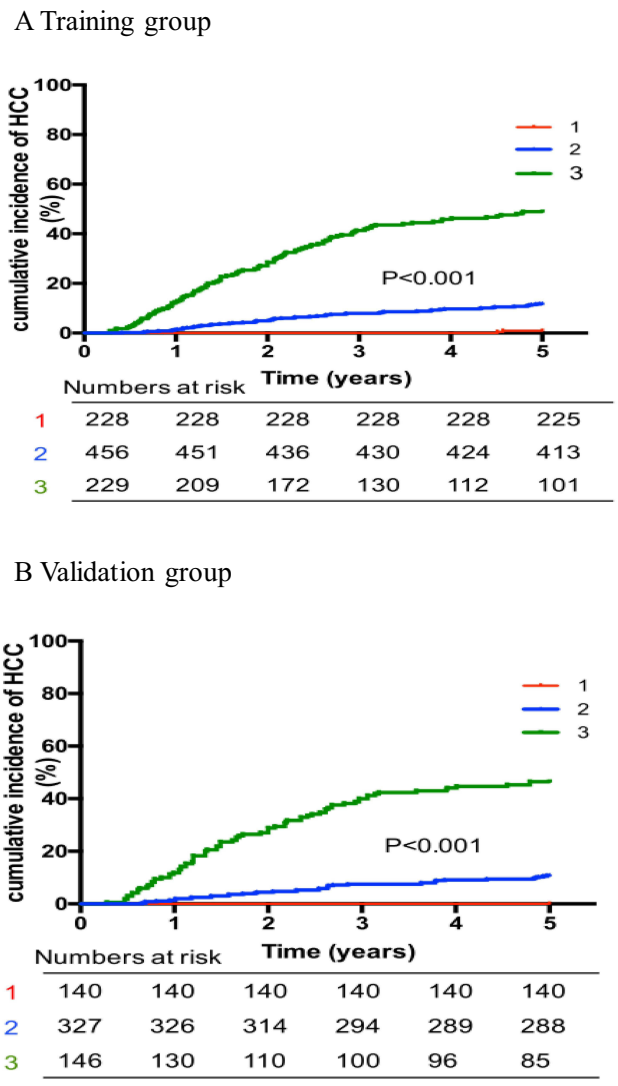

Figure 3 The Kaplan-Meier incidence of HCC in training cohort (**A**), and validation cohort (**B**). ANNs model. According to the ANN model, the patient training and validation sets are divided into three risk layers as follows: low (Stratum1), medium (Stratum2), and high (Stratum3).

The performance of the ANN model in predicting HCC development at 5 years was excellent, as evidenced by the high area under the curve (AUC) value of 0.880 obtained from training and calibration curves. In comparison, other models such as mPAGE-B (age, sex, ALB, and PLT), PAGE (age, sex, and PLT), FIB-4 (age, aspartate aminotransferase (AST), ALT, and PLT), and THRI (age, sex, etiology, and PLT) had lower AUC values. The superior predictive performance of our ANN model was particularly evident in patients with hepatitis B virus infection ($p < 0.001$ for all). One advantage of ANN models is their ability to learn from data and adjust the connections among variables to optimize prediction accuracy. Unlike traditional logistic regression or Cox regression models, ANN models are nonlinear and repeatedly train the factors relevant to the outcome. This iterative process allows ANN models to achieve higher prediction accuracy. Although the PAGE score, which is based on age, gender, and PLT, can predict the occurrence of liver cancer within a 5-year period, but its prediction performance needs to be improved. The modified mPAGE-B score, incorporating ALB and adjusted weighting for age, sex, and PLT, showed limited accuracy with an AUROC of only 0.719, lower than that of our ANN model (0.880) The THRI score (platelets, age, sex, and etiology) can serve as a useful tool for guiding recommendations pertaining to HCC surveillance among cirrhosis patients. However, it's important to note that these models have not been proven to be applicable to patients with HBC. Nevertheless, the ANN model

Table 4 Comparison of Performance and Discriminative Ability Among the Current Model and Other Models

Cohort	Models	5-Year Risk of HCC		*P* value
		AUROC (95% CI)	C-Index (95% CI)	
Training	ANN	0.880(0.847–0.914)	0.843(0.812–0.874)	
	mPAGE-B	0.719 (0.670–0.768)	0.691 (0.646–0.735)	<0.001
	PAGE-B	0.710 (0.660–0.759)	0.681 (0.636–0.726)	< 0.001
	FIB-4	0.693 (0.640–0.745)	0.659 (0.612–0.707)	< 0.001
	THRI	0.705 (0.654–0.756)	0.676 (0.63–0.723)	< 0.001
Validation	ANN	0.824 (0.761–0.887)	0.809 (0.753–0.865)	
	mPAGE-B	0.718(0.642–0.793)	0.707 (0.643–0.772)	< 0.001
	PAGE-B	0.716 (0.640–0.793)	0.701 (0.631–0.771)	< 0.001
	FIB-4	0.711 (0.625–0.797)	0.710 (0.632–0.788)	< 0.001
	THRI	0.720 (0.644–0.797)	0.713 (0.642–0.783)	< 0.001

Abbreviations: mPAGE-B, modified platelets, age, gender-hepatitis B scores; PAGE-B, platelets, age, gender-hepatitis B scores; FIB-4, Fibrosis-4 index; THRI Toronto HCC risk index; HCC, hepatocellular carcinoma.

presents several advantages over previous risk models. These include its ability to inform HCC screening strategies for patients at different clinical stages of HBC, its capacity to calculate the annual incidence of HCC, its large sample size, and its excellent performance in both training and validation cohorts.

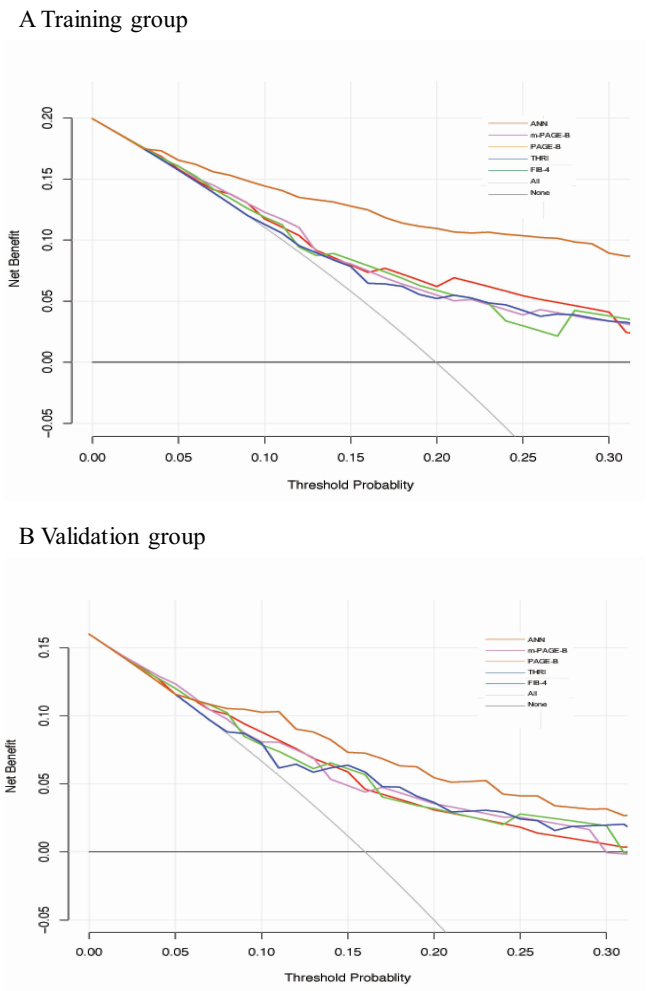

Figure 4 The calibration curve for predicting 5-year incidence of HCC in training (**A**) and validation cohort (**B**).

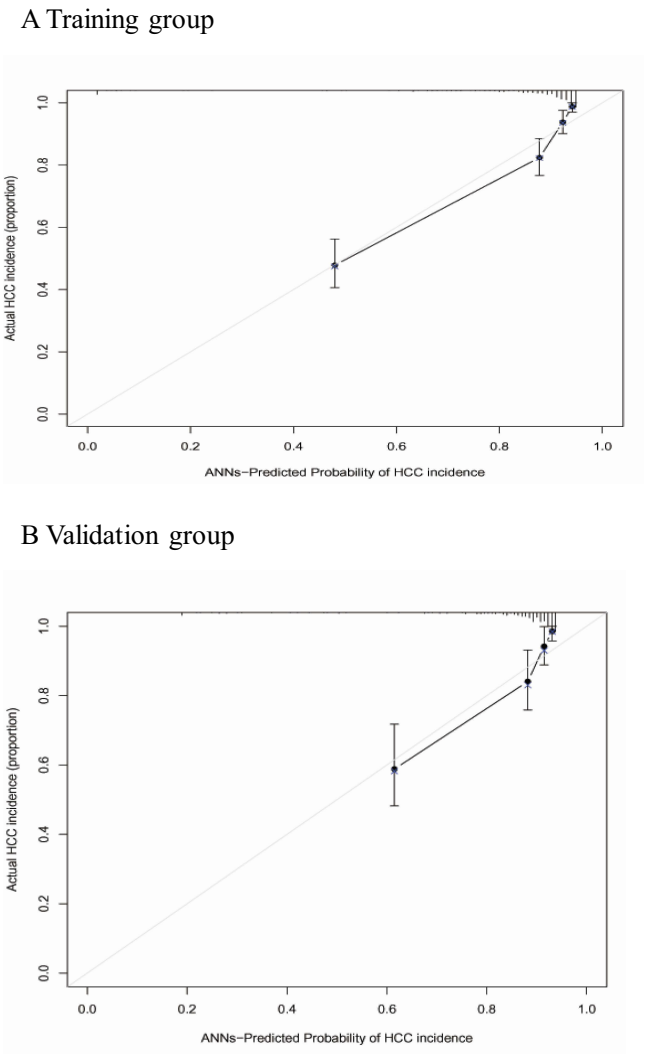

Figure 5 The cumulative probabilities of hepatocellular carcinoma of our model at 5 years in the training (**A**) and validation (**B**) data sets.

The current major international guidelines propose biannual abdominal ultrasound for HCC surveillance in cirrhosis patients, regardless of their individual risk.[25,26] However, studies have indicated that the risk of developing HCC is not the same for everyone. As a result, the current "one-size-fits-all" approach often leads to either an overestimation or underestimation of HCC risk for each individual.[27,28] Therefore, the hierarchical management strategy can be applied to patients with liver cirrhosis. The ANN model is able to classify patients into three distinct risk categories based on their 5-year cumulative incidence of HCC. In the training cohort, the low-risk group displayed a negative predictive value of 98.7%, while in the validation cohort, it was 100%. This indicates its ability to accurately identify patients who are less likely to develop liver cancer. By implementing a screening strategy based on our prediction model, we observed a greater net benefit compared to the "All" approach and other models. According to the risk stratification provided by the ANN model, surveillance for HCC may be reduced among patients from the low-risk group. This may lead to a decrease in potentially harmful and costly diagnostic workups, given the estimated annual incidence of approximately 0.48%[29,30]

Additionally, patients belonging to the high-risk category demonstrated a notably elevated incidence of HCC development, reaching a cumulative rate of 56.8% after 5 years of antiviral therapy. This finding emphasizes the criticality of intensive surveillance in promptly identifying HCC at its early stages to facilitate the implementation of suitable treatments, thereby maximizing the chances of better survival outcomes.[31] Furthermore, patients can enhance

their adherence to treatment protocols, optimize the cost-effectiveness of healthcare interventions, and ensure the optimal allocation of limited medical resources.

It is imperative to acknowledge the limitations of our study. Primarily, this investigation was carried out retrospectively, leading to an inherent selection bias. Nevertheless, it possesses the advantages of comprehensive clinical data and a robust sample size. Coupled with the utilization of deep learning techniques and extensive training of the artificial intelligence neural network, the ANN model exhibited exceptional predictive capabilities in evaluating the risk of HCC. Nonetheless, the replication of these results on a larger scale and prospects for prospective studies are necessary. Secondly, differential interpretations by various radiologists or hepatologists, as well as variances in compliance across different institutions, might have impacted the diagnosis of HCC, as observed in our study. Thirdly, the absence of hepatitis B surface antigen levels and genetic studies for the majority of our patients prevented the evaluation of the potential role of such markers in our study.

Summary

This study employed an ANN model to construct a predictive tool for estimating the 5-year risk of HCC development in patients with HBC undergoing antiviral therapy. The ANN model demonstrated promising individualized prediction performance, thereby offering valuable assessment of HCC risk in clinical settings for patients with HBC.

Ethical Approval

The study was approved by the Ethics Committee of Beijing Ditan Hospital, Capital Medical University. Written informed consent was obtained from each patient. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008.

Funding

This work was supported by Beijing Hospitals Authority Youth Programme (QMl220201802), the Beijing Traditional Chinese medicine science and Technology Development Fund Project (No. Qn-2020-25), Application of Clinical Features of Capital City of Science and Technology commission (z181100001718052),High-level public health technical personnel construction project (The backbone of the discipline-03-07), National Key R&D Program of China (20232023YFC230880).

Disclosure

The authors declare that they have no conflicts of interest with regard to the publication of this research report.

References

1. Wild C. *World Cancer Report 2020[M]. Imprimerie Faurite*. France: World Health Organization; 2020.
2. Lee SW, Choi J, Kim SU, Lim Y-S. Entecavir versus tenofovir in patients with chronic hepatitis B: enemies or partners in the prevention of hepatocellular carcinoma. *Clin Mol Hepatol*. 2021;27(3):402–412. doi:10.3350/cmh.2021.0179
3. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *Ca a Cancer J Clinicians*. 2018;68(6):394–424. doi:10.3322/caac.21492
4. Korean Association for the Study of the Liver (KASL). KASL clinical practice guidelines for management of chronic hepatitis B. *Clin Mol Hepatol*. 2022;28(2):276–331. doi:10.3350/cmh.2022.0084
5. Chan HLY. Okuda lecture: challenges of hepatitis B in the era of antiviral therapy. *J Gastroenterol Hepatol*. 2019;34:501–506. doi:10.1111/jgh.14534
6. Yu JH, Cho SG, Y- J J, Lee J-W. The best predictive model for hepatocellular carcinoma in patients with chronic hepatitis B infection. *Clin Mol Hepatol*. 2022;28:351–362. doi:10.3350/cmh.2021.0281
7. El-Serag HB, Kanwal F, Feng Z, Marrero J, Khaderi S, Singal AG. Risk factors for cirrhosis in contemporary hepatology practices-findings from the Texas hepatocellular carcinoma consortium cohort. *Gastroenterology*. 2020;159:376–377. doi:10.1053/j.gastro.2020.03.049
8. Amit GS, Emily Z, Manasa N, et al. HCC surveillance improves early detection, curative treatment receipt, and survival in patients with cirrhosis: a meta-analysis. *J Hepatol*. 2022;77(77):128–139. doi:10.1016/j.jhep.2022.01.023
9. Marasco G, Colecchia A, Colli A, et al. Role of liver and spleen stiffness in predicting the recurrence of hepatocellular carcinoma after resection[J]. *J Hepatol*. 2019;70(3):440–448. doi:10.1016/j.jhep.2018.10.022
10. McGlynn KA, Petrick JL, El- Serag HB. Epidemiology of hepatocellular carcinoma. *Hepatology*. 2021;73(4):4–13. doi:10.1002/hep.31288

11. YVan der Pol CB, Lim CS, Sirlin CB, et al. Accuracy of the liver imaging reporting and data system in computed tomography and magnetic resonance image analysis of hepatocellular carcinoma or overall malignancy— a systematic review. *Gastroenterology*. 2019;156(4):976–986. doi:10.1053/j.gastro.2018.11.020
12. Liang LY, Lee HW, Wong VW, et al. Serum fibrosis index- based risk score predicts hepatocellular carcinoma in untreated patients with chronic hepatitis B. *Clin Mol Hepatol*. 2021;27(3):499–509. doi:10.3350/cmh.2020.0333
13. Arends P, Sonneveld MJ, Zoutendijk R, et al. Entecavir treatment does not eliminate the risk of hepatocellular carcinoma in chronic hepatitis B: limited role for risk scores in Caucasians. *Gut*. 2015;64(8):1289–1295. doi:10.1136/gutjnl-2014-307023
14. Papatheodoridis G, Dalekos G, Sypsa V, et al. PAGE-Bpredicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitisB on 5-year antiviral therapy. *J Hepatol*. 2016;64:800–806. doi:10.1016/j.jhep.2015.11.035
15. 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 Hepatol*. 2018;69(5):1066–1073. doi:10.1016/j.jhep.2018.07.018
16. Xiao G, Yang J, Yan L. Comparison of diagnostic accuracy of aspartate aminotransferase to platelet ratio index and fibrosis-4 index for detecting liver fibrosis in adult patients with chronic hepatitis B virus infection: a systemic review and meta-analysis. *Hepatology*. 2015;61:292–302. doi:10.1002/hep.27382
17. Sharma SA, Kowgier M, Hansen BE, et al. Toronto HCC risk index: a validated scoring system to predict 10-year risk of HCC in patients with cirrhosis[J]. *J Hepatol*. 2018;68:92–99. doi:10.1016/j.jhep.2017.07.033
18. Nam JY, Sinn DH, Bae J, Jang ES, Kim JW, Jeong SH. Deep learning model for prediction of hepatocellular carcinoma in patients with HBV-related cirrhosis on antiviral therapy. *JHEP Rep*. 2020;2:100175. doi:10.1016/j.jhepr.2020.100175
19. Makridakis S, Spiliotis E, Assimakopoulos V. Statistical and machine learning forecasting methods: concerns and ways forward. *PloSOne*. 2018;13: e0194889. doi:10.1371/journal.pone.0194889
20. Wu S, Kong Y, Piao H, et al. On-treatment changes of liver stiffness at week 26 could predict 2-year clinical outcomes in HBV-related compensated cirrhosis. *Liver Int*. 2018;38:1045–1054. doi:10.1111/liv.13623
21. Omata M, Cheng AL, Kokudo N, et al. Asia–Pacific clinical practice guidelines on the management of hepatocellular carcinoma, a 2017 update. *Hepatol Int*. 2017;11:317–370. doi:10.1007/s12072-017-9799-9
22. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet*. 1995;346(8982):1075–1079. doi:10.1016/S0140-6736(95)91746-2
23. Baxt WG. Application of artificial neural networks to clinical medicine. *Lancet*. 1995;346(8983):1135–1138. doi:10.1016/S0140-6736(95)91804-3
24. Sharma SA, Kowgier M, Hansen BE, et al. Toronto HCC risk index, A validated scoring system to predict 10-year risk of HCC in patients with cirrhosis. *J Hepatol*. 2017;S0168-8278(17):32248.
25. Kim WR, Berg T, Asselah T, et al. Evaluation of APRI and FIB-4 scoring systems for non-invasive assessment of hepatic fibrosis in chronic hepatitis B patients. *J Hepatol*. 2016;64:773–780. doi:10.1016/j.jhep.2015.11.012
26. Research EO, European Association for the Study of the Liver. EASL–EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol*. 2012;56:908–943. doi:10.1016/j.jhep.2011.12.001
27. Terrault NA, Bzowej NH, Chang KM, et al. AASLD guidelines for treatment of chronic hepatitis B. *Hepatology*. 2016;63(1):261–283. doi:10.1002/hep.28156
28. Atiq O, Tiro J, Yopp AC, et al. An assessment of benefits and harms of hepatocellular carcinoma surveillance in patients with cirrhosis. *Hepatology*. 2017;65(4):1196–1205. doi:10.1002/hep.28895
29. Heleno B, Thomsen MF, Rodrigues DS, Jørgensen KJ, Brodersen J. Quantification of harms in cancer screening trials: literature review. *BMJ*. 2013;347(sep16 1):f5334. doi:10.1136/bmj.f5334
30. Goossens N, Singal AG, King LY, et al. Cost-Effectiveness of Risk Score-Stratified Hepatocellular Carcinoma Screening in Patients with Cirrhosis. *Clin Transl Gastroenterol*. 2017;8:e101. doi:10.1038/ctg.2017.26
31. Wang JH, Chang KC, Kee KM, et al. Hepatocellular carcinoma surveillance at 4- vs. 12-month intervals for patients with chronic viral hepatitis: a randomized study in community. *Am J Gastroenterol*. 2013;108(3):416–424. doi:10.1038/ajg.2012.445