Development and Validation of a Diagnostic Model for AFP- Negative Hepatocellular Carcinoma

Zhou Yu
The First Affiliated Hospital of Fujian Medical University
Dongmei Chen
The First Affiliated Hospital of Fujian Medical University
Yansong Zheng
The First Affiliated Hospital of Fujian Medical University
Xuedan Wang
The First Affiliated Hospital of Fujian Medical University
Shuna Huang
The First Affiliated Hospital of Fujian Medical University
Tiansheng Lin
Fujian Provincial Hospital
Yihan Lin
Fujian Provincial Hospital
Yanfang Zhang
The First Affiliated Hospital of Fujian Medical University
Yingna Huang
Fujian Medical University
Qishui Ou
The First Affiliated Hospital of Fujian Medical University
Jinlan Huang (hjlan@fjmu.edu.cn)
The First Affiliated Hospital of Fujian Medical University

Research Article

Keywords: AFP-negative HCC, diagnostic model, nomogram, PIVKA-II

Posted Date: April 11th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2790615/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Journal of Cancer Research and Clinical Oncology on June 27th, 2023. See the published version at https://doi.org/10.1007/s00432-023-04997-4.

Abstract

Purpose AFP appears to be negative about 30% of overall hepatocellular carcinoma (HCC). Our study aimed to develop a nomogram model to diagnose AFP negative HCC (AFPN-HCC).

Patients and methods: The training set and the external validation set consisted of 516 and 456 objects. LASSO, univariate and multivariable logistic regression were performed to construct the model and then transformed into a visualized nomogram. We further used the receiver operating characteristic (ROC) curves, the calibration curve, decision curve analysis (DCA) and clinical impact curve (CIC) for validation.

Results: Four variables included age, PIVKA-II, platelet (PLT) counts and prothrombin time(PT) were selected to establish the nomogram. The area under the curve (AUC) of the ROC to distinguish AFPN-HCC patients was 0.937(95%CI, 0.892-0.938) in training set and 0.942(95%CI, 0.921-0.963) using the validation set and indicated satisfactory discriminative ability of the model. The calibration plots showed favorable consistency between the prediction of the nomogram and actual observations. DCA and CIC showed that the nomogram was clinically useful.

Conclusions: Our model was effective for discrimination of AFPN-HCC from control subsets, and might be helpful for the diagnosis for AFPN-HCC.

1 Introduction

Hepatocellular carcinoma (HCC) is the dominant histological type of liver cancer, and accounts for 75–85% of all cases. According the global cancer statistics 2022, HCC is the sixth most common malignancy and the third leading cause of cancer-related death in the world[1].Surgical resection is still the preferred method for the treatment of HCC but is not suitable for HCC patients with advanced stage[2].Unfortunately, due to the occult onset of HCC and the lack of specifically early markers, most patients often diagnosed at an advanced stage, which associated with a high recurrence rate, metastasis rate and a poor prognosis[3]. Accurate surveillance and differential diagnosis of HCC can significantly improve patient survival.

As the prognosis of HCC depends largely on the stage at which the tumor is detected, early detection of HCC is critical to improve the survival of affected patients. Professional society guidelines from the American Association for the Study of Liver Diseases (AASLD) and European Association for the Study of the Liver (EASL) both suggest a liver ultrasonography (US) for patients with high-risk to develop into HCC as the first level of surveillance[4, 5]. However, the effectiveness of US mainly depends on the experience of operators and it is difficult to distinguish tumors from liver cirrhosis nodules which closely correlated with HCC[6]. In a meta-analysis detect that the sensitivity of ultrasound was only 47% (95% CI, 33–61%) for detection of early-stage HCC[7].Magnetic resonance imaging (MRI), computed Tomography (CT), and other cross-sectional imaging techniques have a higher accuracy, but they are too expensive for widespread screening[8].

Biomarker assessments are more objective, easily accessible, and noninvasive tools for HCC diagnosis. Of all biomarkers, alpha-fetoprotein (AFP) is the most widely used serological indicator for HCC worldwide[9]. However, about 30% of overall HCC patients cannot be observed with elevated serum AFP and AFP can also elevate in other benign liver diseases such as chronic hepatitis B(CHB) and liver cirrhosis(LC)[10].These facts spark controversy about the use of AFP and the AASLD no longer recommend the use of AFP during HCC surveillance[11]. A delayed diagnosis of AFP-negative HCC (AFPN-HCC) is frequently lead to delays in treatment and subsequently to a serious consequence.

The detection of circulating biomarkers associated with AFPN-HCC can improve diagnostic accuracy and overcome the disadvantages of current diagnostic strategies[12]. Prothrombin induced by vitamin K absence II (PIVKA-II), also known as Des-γ-carboxyprothrombin (DCP), is an abnormal prothrombin molecule and has been used as a biological marker which is increased in HCC[13]. Evidence has been presented that PIVKA-II can improve the positive rate of diagnosis for AFPN-HCC patients[14].However, serum PIVKA-II levels are also increased without HCC because of a shortage of vitamin K or usage of vitamin K antagonists[15]. This makes its clinical applications as a marker of HCC limited. Lens culinary agglutinin-reactive fraction of fetoprotein (AFP-L3), is the glycosylated subfraction of AFP and is a more specific indicator than total AFP for HCC[16]. However, previous studies have indicated that AFP-L3 presented low diagnostic sensitivity in cases where AFP is not markedly elevated[17, 18]. Moreover, the strategy of combined multiple biomarkers has been shown to significantly enhance diagnostic performance[19, 20]. GALAD is the most extensively studied which includes gender, age, AFP-L3%, AFP, and DCP. In an international cohort of 6834 patients (2430 with HCC and 4404 with chronic liver disease), GALAD achieved sensitivities ranging from 60–80% for early HCC detection[21]. But numerous studies tended to omit or ignore the application on AFPN-HCC[22, 23]. Liu. et al developed a serum-based GAAP which was based on gender, age, AFP and PIVKA-II for surveillance of HCC and the model had an AUC value of 0.888 for discriminating AFPN-HCC from the entire control[24]. Although data are promising, confirmation of the clinical effectiveness in larger studies is needed.

The purpose of this study was to develop a diagnostic model with combined multiple biomarkers or serological examinations for AFPN-HCC via LASSO regression analysis, univariate logistic regression analysis and multivariable logistic regression. Then, the diagnostic model was transformed into a visualized nomogram and its discrimination, calibration and the net benefits were further validated. External validation was made on validation sets. We hope that the diagnostic model could be applied for the diagnosis of AFPN-HCC.

2 Methods

2.1 Study Population

All HCC patients confirmed by postoperative pathology from January 2019 to May 2022 were included in this study. After screening according to inclusion and exclusion criteria, 294 AFPN-HCC patients from the

First Affiliated Hospital of Fujian Medical University were selected as training set and 227 patients from Fujian Provincial Hospital were enrolled as validation set for this retrospective study. Training set also included a cohort of 159 healthy objects, 63 patients with CHB and 64 patients with LC as controls. The validation set from Fujian Provincial Hospital consisted of 137 healthy controls objects ,47 CHB patients and 45 patients with LC. A written informed consent was obtained from all subjects. All procedures followed were in accordance with Helsinki declaration. The development of the study followed the criteria of the TRIPOD statement[25]. The Ethics Committee of the First Affiliated Hospital of Fujian Medical University reviewed and approved this study (2018[048]).

2.2 Inclusion Criteria and Exclusion Criteria

The HCC inclusion criteria were listed as follows: (1) postoperative pathological diagnosis of HCC; (2) preoperative AFP negative (< 20 ng/mL); (3) first onset without any anticancer treatment before hepatectomy; (4) complete preoperative clinical data.

The HCC exclusion criteria was listed as follows: (1) preoperative AFP positive (≥ 20 ng/mL); (2) received anticancer treatment before hepatectomy; (3) recrudescent HCC; (4) miss data.

The healthy control group was defined as: (1) no family history of cancer and liver-related disease; (2) routine tests including full blood count, electrolytes, liver and kidney function tests, coagulation function, AFP, CEA, PIVKA and other laboratory results within normal range.

The CHB patients were selected according to the guidelines of prevention and treatment for chronic hepatitis B[26, 27]. The LC patients were enrolled according to guidelines on the management of liver cirrhosis[28, 29].

2.3 Laboratory Examination and Data Collection

Demographic data (age and gender) was obtained from the electronic medical records. Laboratory examination data including complete blood routine, electrolytes, liver and kidney function, coagulation function, AFP, CEA, PIVKA-II were obtained within one week before surgery from the Laboratory Information System (LIS). The laboratory methods were detailed briefly below. The serum AFP and CEA were measured by the electroche-miluminescence detection system (Roche, Basel, Switzerland); PIVKA-II was measured by the chemiluminescent microparticle immunoassay (Abbott Laboratories, IL, USA). The blood routine was performed using an ADVIA 2120 automatic blood analyzer (Siemens, Erlangen, Germany). The biochemical indexes were detected via the Cobas-8000 automatic biochemical analyzer (Roche Diagnostics, Basel, Switzerland). Automated coagulation tests were performed using CS5100 coagulometric auto analyzers (Sysmex, Kobe, Japan).

2.4 Statistical Analysis

All statistical analysis were performed by SPSS 26.0 (IBM Corporation, 2020, USA) and R (version 4.2.0, R Foundation for Statistical Computing, Vienna, Austria. URL https://www. R-project.org/). Continuously distributed variables were reported as mean ± standard deviation (SD) for data with normal distribution,

or median and interquartile range for nonnormally distributed data. Categorical covariates were expressed as percentage. LASSO regression analysis and univariate logistic regression were used to identify individual factors of AFPN-HCC. Then the significant variables with significance $P < 0.05$ were selected for multivariate analyses. A visualized nomogram was conducted based on the results of the multivariate model. The predictive performance of the algorithm models was measured by the receiver operating characteristic (ROC) curves analysis. DeLong's test was used to compared the validity between the AUCs of each model. A calibration curve was generated for evaluating the calibration. Decision curve analysis (DCA) and clinical impact curves (CIC) was conducted to determine the clinical benefit of the model. Summary of the study design was presented in Fig. 1.

3 Results

3.1 Characteristics of patients

A total of 521 patients(mean age, 59.2 ± 11 years; male to female ratio, 4.1:1) were enrolled in the present study. Demographics, baseline patient characteristics, main laboratory data of AFPN-HCC patients for both training (294 patients) and validation sets (227 patients) were shown in Table 1. There was no significant difference between training set and validation set for demographic and clinical laboratory characteristics.

Table 1
Demographic and Clinical Laboratory Characteristics of the Training and Validation set

		Training set	Validation set	*P* value
Number of cases		294	227	
Gender, Male/female	n (%)	250/44	196/31	0.568
Age, years	Mean (SD)	59.3 (11.2)	59.1 (11.0)	0.819
AFP, ng/mL	Median [Q1; Q3]	5.14 (4.68)	4.84 (4.16)	0.449
CA199, ng/mL	Median [Q1; Q3]	14.8 [8.36;25.8]	14.5 [8.46;23.0]	0.087
CEA, ng/mL	Median [Q1; Q3]	2.60 [1.73;3.72]	2.57 [1.77;3.66]	0.381
PIVKA-II, mAU/mL	Median [Q1; Q3]	100 [38.0;634]	146 [44.5;1137]	0.142
Basophil **counts, 10^9/L**	Median [Q1; Q3]	0.03 [0.02;0.04]	0.03 [0.02;0.04]	0.933
Eosinophils counts, 10^9/L	Median [Q1; Q3]	0.11 [0.07;0.17]	0.11 [0.06;0.18]	0.61
Lymphocyte counts, 10^9/L	Mean (SD)	1.72 (0.98)	1.69 (0.62)	0.675
Monocyte counts, 10^9/L	Mean (SD)	0.42 (0.21)	0.41 (0.20)	0.612
Neutrophil counts, 10^9/L	Mean (SD)	3.99 (3.66)	3.84 (2.58)	0.59
WBC counts, 10^9/L	Mean (SD)	6.06 (2.25)	6.17 (2.53)	0.599
PLT counts, 10^9/L	Mean (SD)	180 (70.4)	181 (71.9)	0.81
RBC counts, 10^{12}/L	Median [Q1; Q3]	4.58 [4.22;4.90]	4.56 [4.18;4.89]	0.232
HCT, L/L	Median [Q1; Q3]	0.42 [0.39;0.46]	0.42 [0.40;0.45]	0.759
HGB, g/L	Median [Q1; Q3]	144 [131;153]	142 [133;151]	0.369
MCH, pg	Median [Q1; Q3]	31.3 [30.3;32.4]	31.2 [30.0;32.3]	0.286

Abbreviations:

AFP, -fetoprotein; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen199; PIVKA-II, protein induced by Vitamin K absence or antagonist-II; PLT, platelet; WBC, white blood cell ; RBC, red blood cell; HCT, red blood cell specific volume; HGB, hemoglobin; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1;CA, calcium; AST, aspartate aminotransferase; CK, creatine kinase; CL, chlorine; CREA, creatinine,; DBIL, direct bilirubin; GGT, -glutamyl transpeptidase; GLO, globin; GLU, glucose; HDLC, high-density lipoprotein cholesterol; IBIL, indirect bilirubin; K kalium; LDH, lactate dehydrogenase; LDLC, low-density lipoprotein cholesterol; MG, magnesium; NA, natrium; P, phosphorus; TBIL, total bilirubin; TCHO, total cholesterol; TG, total triglycerides; TP, total protein; UA, uric acid; UREA, urea nitrogen; APTT, activated partial thromboplastin time; Fg, fibrinogen; PT, prothrombin time; INR, international normalized ratio; TT, thrombin time

		Training set	Validation set	*P* value
MCHC, g/L	Median [Q1; Q3]	337 [330;345]	335 [328;342]	0.215
MCV, fL	Median [Q1; Q3]	92.9 [90.4;95.8]	92.8 [89.6;96.1]	0.688
RDW, %	Median [Q1; Q3]	13.4 [12.6;14.2]	13.2 [12.5;14.1]	0.226
ALB, g/L	Median [Q1; Q3]	42.8 [38.5;46.0]	41.8 [38.0;44.8]	0.153
ALP, U/L	Mean (SD)	92.7 (54.0)	101 (76.5)	0.135
ALT, U/L	Median [Q1; Q3]	27.0 [19.0;40.5]	27.0 [19.2;41.8]	0.184
AST, U/L	Median [Q1; Q3]	28.0 [21.5;36.0]	28.0 [21.0;40.5]	0.247
CA, mmol/L	Median [Q1; Q3]	2.32 [2.25;2.40]	2.25 [2.17;2.33]	0.334
CK, U/L	Median [Q1; Q3]	89.0 [61.5;112]	83.0 [66.0;114]	0.658
CKMB, U/L	Median [Q1; Q3]	14.0 [11.0;18.0]	14.0 [11.0;19.0]	0.689
CL, mmol/L	Mean (SD)	103 (8.88)	101 (14.2)	0.151
CREA, umol/L	Median [Q1; Q3]	70.1 [59.0;80.0]	69.0 [60.8;79.0]	0.779
DBIL, umol/L	Median [Q1; Q3]	4.90 [3.80;6.65]	4.80 [3.70;6.55]	0.611
GGT, U/L	Median [Q1; Q3]	44.0 [26.5;85.0]	41.0 [27.0;71.5]	0.518
GLO, g/L	Median [Q1; Q3]	27.8 [25.0;30.5]	27.4 [24.8;30.6]	0.593
GLU, mmol/L	Median [Q1; Q3]	5.12 [4.56;6.13]	5.13 [4.58;6.37]	0.642
IBIL, umol/L	Median [Q1; Q3]	7.70 [5.30;10.8]	7.70 [5.35;11.1]	0.881
K, mmol/L	Median [Q1; Q3]	4.10 [3.90;4.40]	4.09 [3.90;4.30]	0.279
LDH, U/L	Median [Q1; Q3]	190 [166;216]	189 [166;216]	0.882
MG, mmol/L	Median [Q1; Q3]	0.88 [0.83;0.94]	0.87 [0.82;0.92]	0.175

Abbreviations:

AFP, -fetoprotein; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen199; PIVKA-II, protein induced by Vitamin K absence or antagonist-II; PLT, platelet; WBC, white blood cell ; RBC, red blood cell; HCT, red blood cell specific volume; HGB, hemoglobin; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1;CA, calcium; AST, aspartate aminotransferase; CK, creatine kinase; CL, chlorine; CREA, creatinine,; DBIL, direct bilirubin; GGT, -glutamyl transpeptidase; GLO, globin; GLU, glucose; HDLC, high-density lipoprotein cholesterol; IBIL, indirect bilirubin; K kalium; LDH, lactate dehydrogenase; LDLC, low-density lipoprotein cholesterol; MG, magnesium; NA, natrium; P, phosphorus; TBIL, total bilirubin; TCHO, total cholesterol; TG, total triglycerides; TP, total protein; UA, uric acid; UREA, urea nitrogen; APTT, activated partial thromboplastin time; Fg, fibrinogen; PT, prothrombin time; INR, international normalized ratio; TT, thrombin time

		Training set	Validation set	*P* value
NA, mmol/L	Median [Q1; Q3]	141 [140;142]	141 [139;142]	0.722
P, mmol/L	Median [Q1; Q3]	1.04 [0.94;1.16]	1.0·7 [0.97;1.19]	0.103
TBIL, umol/L	Median [Q1; Q3]	12.7 [9.40;17.6]	12.2 [9.40;17.1]	0.788
TP, g/L	Median [Q1; Q3]	70.8 [64.7;74.4]	70.0 [64.8;74.4]	0.554
UA, umol/L	Mean (SD)	342 (97.8)	336 (97.1)	0.491
UREA, mmol/L	Median [Q1; Q3]	5.10 [4.40;6.11]	5.20 [4.50;6.40]	0.273
APTT, s	Median [Q1; Q3]	28.6 [25.8;35.0]	27.4 [25.5;33.6]	0.052
Fg, g/L	Median [Q1; Q3]	2.79 [2.30;3.37]	2.70 [2.26;3.33]	0.373
PT, s	Median [Q1; Q3]	12.0 [11.2;13.2]	12.0 [11.1;13.0]	0.242
INR	Median [Q1; Q3]	1.03 [0.97;1.10]	1.03 [0.97;1.08]	0.478
TT, s	Median [Q1; Q3]	17.3 [16.4;18.0]	17.2 [16.3;18.1]	0.616

Abbreviations:

AFP, -fetoprotein; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen199; PIVKA-II, protein induced by Vitamin K absence or antagonist-II; PLT, platelet; WBC, white blood cell ; RBC, red blood cell; HCT, red blood cell specific volume; HGB, hemoglobin; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1;CA, calcium; AST, aspartate aminotransferase; CK, creatine kinase; CL, chlorine; CREA, creatinine,; DBIL, direct bilirubin; GGT, -glutamyl transpeptidase; GLO, globin; GLU, glucose; HDLC, high-density lipoprotein cholesterol; IBIL, indirect bilirubin; K kalium; LDH, lactate dehydrogenase; LDLC, low-density lipoprotein cholesterol; MG, magnesium; NA, natrium; P, phosphorus; TBIL, total bilirubin; TCHO, total cholesterol; TG, total triglycerides; TP, total protein; UA, uric acid; UREA, urea nitrogen; APTT, activated partial thromboplastin time; Fg, fibrinogen; PT, prothrombin time; INR, international normalized ratio; TT, thrombin time

3.2 Identify Independent Variables Significantly Associated with AFPN-HCC

To identify the impact of individual factors associated with AFPN-HCC, univariate analysis was performed in training set. The results showed that 17 variables including age, gender, CEA, PIVKA-II, monocyte counts, neutrophil counts, PLT counts, WBC counts, MCHC, RDW, ALP, CK, CREA, GLU, APTT, Fg, PT with statistical significance ($P < 0.05$) were brought into the next analysis (Table 2). Summary statistics for all independent variables were presented in Supplementary Table 1. Given the large number of variables, LASSO regression was used for dimensionality reduction analysis to further screen AFPN-HCC-related factors from the univariate analysis result. Figure 2A showed the path of all candidate variable coefficients included in the model according to the level of logarithmic transformation λ, and as the optimal penalization coefficient λ increased, the number of independent coefficients tended toward zero. Identification of the λ in the LASSO model used tenfold cross-validation and minimum criterion. The

confidence interval (CI) under each λ was shown in Fig. 3D. The non-zero coefficients were considered to have strong prognostic potential in the LASSO penalized regression model. As a result, a total of 17 significant factors from the result of the univariate regression were used for the LASSO regression, and 8 key variables including age, gender, PIVKA-II, monocyte counts, PLT counts, ALP, PT, MCHC were left (Table 3). All 12 variables were compared between the HCC, LC, CHB and HC groups. (Fig. 2).

Table 2
Significant Variables of Univariate Logistic Regression Analysis(n = 17)

Variables	β	OR	95%CI		*P*-value
			Lower	Upper	
Age	0.085	1.088	1.07	1.107	< 0.001
Gender	-0.843	0.430	0.286	0.648	< 0.001
CEA	0.147	1.158	1.059	1.267	0.001
PIVKA-II	0.002	1.002	1.001	1.003	< 0.001
Monocyte counts	2.576	13.147	4.318	40.028	< 0.001
Neutrophil counts	0.189	1.209	1.084	1.347	0.001
PLT counts	-0.003	0.978	0.964	0.991	0.001
WBC counts	0.114	0.997	0.994	0.999	0.002
MCHC	-0.015	1.121	1.032	1.217	0.007
RDW	0.022	0.888	0.795	0.991	0.034
ALP	0.008	1.008	1.004	1.013	< 0.001
CK	-0.003	0.997	0.995	1.000	0.030
CREA	0.013	1.013	1.004	1.023	0.006
GLU	0.223	1.249	1.122	1.391	< 0.001
APTT	-0.053	0.948	0.921	0.976	< 0.001
Fg	0.320	1.377	1.135	1.671	0.001
PT	-0.369	0.691	0.608	0.786	< 0.001

Note: Data were presented as the odds ratio with the confidence interval

Abbreviations: OR, odds ratio; CI, confidence interval. CEA, carcinoembryonic antigen; PIVKA-II, protein induced by Vitamin K absence or antagonist-II; PLT, platelet; WBC, white blood cell ; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; ALP, alkaline phosphatase; CK, creatine kinase; CREA, creatinine; GLU, glucose; APTT, activated partial thromboplastin time; Fg, fibrinogen; PT, prothrombin time

Table 3
The Result of LASSO Regression Analysis(n = 8)

Variables	LASSO regression coefficient
Gender	4.08E-02
Age	1.29E-02
PIVKA-II	3.23E-06
Monocyte counts	1.82E-01
PLT counts	-3.68E-04
ALP	2.59E-04
PT	-5.03E-02
MCHC	-6.59E-11

2.2 Developing and Visualized a Multivariate Logistic Regression Model

A total of 8 variables were included in the multivariate logistic regression analysis to evaluate combination effects of multiple factors. To facilitate clinical usefulness and practicality, the PIVKA-II was transformed into four levels(1, 2, 3, 4)according to quartile. At last, four significant predictors (age, PLT counts, PT, PIVKA-II) were included in the final model. The OR, 95% CI, and its statistical significance for each variable were presented in Table 4. Univariate logistic regression analysis showed that age (OR = 1.082, $P < 0.001$), PIVKA-II (OR = 6.318, $P = < 0.001$), PLT (OR=-0.006, $P = 0.01$), PT (OR=-0.588, $P = 0.003$) could construct a diagnosis model of AFPN-HCC.

Table 4
Significant Variables of Multivariate Logistic Regression Analysis(n = 4)

Variables	β	OR	95%CI		*P*-value
			Lower	Upper	
Age	0.079	1.082	1.057	1.108	< 0.001
PIVKA-II	1.843	6.318	4.644	8.595	< 0.001
PT	-0.588	0.555	0.458	0.673	< 0.001
PLT counts	-0.006	0.994	0.991	0.998	0.001

We then established a nomogram for AFPN-HCC diagnosis including these four independent factors based on the multivariate logistic regression analysis (Figure.4D). Use of the nomogram is simple. First

the points corresponding to each variable were marked, and then the sum of the points was calculated as the total points, at last we can get the predicted probability value corresponded the total point.

2.3 Validation prediction models

The AUC of ROC curve to distinguish HCC patients from controls in training set was 0.937(95% CI, 0.917–0.956) and in the validation set was 0.942(95% CI, 0.9096–0.964) (Fig. 4A,4E). In addition, with a cutoff point maximizing the sum of sensitivity and specificity, the model achieved sensitivity values of 0.902 and 0.921, and corresponding specificity values of 0.854, 0.882 in training and validation set respectively. The control population comprised of three parts: CHB patients, LC patients, and healthy controls. In order to further confirm the discriminatory power of the model, ROC curve was performed to compare AFPN-HCC with each of the three other groups in training (Fig. 4B-D) and validation sets (Fig. 4F-H). In training set, the model had an AUC of 0.958(95% CI, 0.941–0.975), 0.885(95% CI, 0.845–0.927), and 0.936(95% CI, 0.909–0.964) for distinguish HCC from the HC subsets, the CHB subset, and the LC subset. In validation set, the AUC was 0.956(95% CI, 0.937–0.975), 0.941(95% CI, 0.918–0.963), and 0.911(95% CI, 0.863–0.957).

To evaluate the agreement between predicted probability and the fraction of true observed outcome, the calibration curve was plotted. The calibration curve (Fig. 5A,5B) demonstrated that there was a good agreement between the actual observations and predicted probabilities of AFPN-HCC, and the nomogram model appeared to be well-calibrated in training set (mean absolute error = 0.016) and validation set (mean absolute error = 0.010).

To assess the clinical practicality and usefulness of our model, the DCA and CIC was conducted. DCA was conducted to determine the clinical utility of the model by quantifying the net benefits at different threshold probabilities. With the extension of the model curve, the net benefit increases, the results showed that the model yielded net benefits both in training set (Fig. 5C) and validation set (Fig. 5D). CIC of the model in the training set (Fig. 5E) and validation set (Fig. 5F) showed that the predicted number of high-risk patients was always greater than with outcomes of HCC when the risk threshold was in the range of 0–0.3, and the cost–benefit ratios would be acceptable in the same range.

2.4 Comparison of the diagnostic efficacy of PIVKA-II and GAAP model

In order to investigate the different indicators’ accuracies, ROC curves were drawn and the AUC comparison was performed using DeLong's test (Fig. 6, Table 5). As demonstrated previously, the level of PIVKA-II in the HCC group was significantly higher than that in other subsets ($P < 0.001$, Fig. 2C). The AUC value of PIVKA-II alone to diagnose HCC in the training set was 0.851, which was lower than our model($P < 0.001$). The GAAP model provided an AUC value of 0.892 and it was lower than the AUC of our model($P = 0.012$).

Table 5
Comparison between the model and PIVKA-II and GAAP model

	Model/ Biomarker	AUC (95% CI)	*P* value	Cut-Off	Sensitivity	Specificity
Trainset set	Model	0.937(0.892–0.938)	—	0.591	0.902	0.854
	PIVKA-II	0.884(0.857–0.911)	< 0.001	0.518	0.846	0.837
	GAAP	0.763(0.682–0.844)	< 0.001	0.424	0.864	0.830
Validation set	Model	0.942(0.921–0.963)	—	0.510	0.917	0.886
	PIVKA-II	0.882(0.852–0.913)	< 0.001	0.525	0.851	0.838
	GAAP	0.882(0.851–0.9135)	< 0.001	0.450	0.829	0.794

Discussion

In this study, we constructed a nomogram prediction model including age, PIVKA-II, PLT and PT through LASSO, univariate and multivariate logistic regression analysis. An additional series of analysis was performed for model validation using training set and external validation set. It's well known that the AUC represents the diagnostic efficacy: AUC values of 0.5–0.7 indicate that the diagnostic value is limited, AUC values between 0.7 and 0.9 indicate a perfect diagnostic value, and AUC values greater than 0.9 indicate high accuracy. To further verify the diagnostic power of the model, we compared the discrimination of different control population (HC, CHB and LC). Thus, our data indicated that the model was deemed fit. Overall, the model presents excellent value in diagnose AFPN- HCC patients from healthy controls and benign liver disease (CHB, LC).

In this study, we included age in the final result and the trend matched the clinical and epidemiological observations that increased age are regarded as an independent HCC risk factor[30].HCC patients often have chronic hepatitis and liver cirrhosis background, so the liver function and liver reserve have been altered and damaged. Among the four key factors, PT and PLT are part of the biomarker referred to liver function. PT is a very important index reflecting liver synthesis function, reserve function and increased PT depends on the decreased synthesis of liver-derived coagulation factors[31, 32]. The nomogram shows that high level of PT represents a low point. One possible reason may be that the CHB hospitalized patients always occur with severe liver dysfunction which may leads to a high level of PT (Fig. 2G). Another reason may be due to that the data of HCC patients were collected before surgery with well-preserved hepatic function, so the PT tests were mostly within normal ranges. Numerous predictive models based on PLT counts have been built to identify HCC risk and the result showed that low numbers of platelets were associated with increased risk of HCC[33].Due to the liver cirrhosis background which could ultimately lead to portal hypertension and hypersplenism, HCC patients showed a subsequent lower platelet count than healthy individuals (Fig. 2E).

For a long time, there has been considerable effort devoted to searching for new biomarkers for diagnosis of HCC. In clinical practice, AFP is the most commonly used serum marker to screen for and diagnose HCC, but AFP-based methods are unsuitable for patients with AFPN-HCC. Several studies have discovered novel potential biomarkers such as PIVKA-II, AFP-L3, Golgi protein 73 (GP73), squamous cell carcinoma antigen (SCCA), centromere protein F autoantibody (anti-CENPF), glypican 3 and a number of DNA biomarkers, RNA biomarkers, protein biomarkers, but the clinical applications await future large-scale validation studies[34–39]. Among these, PIVKA-II has been detected to be elevated in HCC patients and has high sensitivity and specificity for differentiating HCC from patients with cirrhosis or chronic hepatitis[40]. Previous studies have indicated that the AUC of PIVKA-II for diagnosing HCC ranged from 0.701 to 0.854, sensitivity ranged from 0.51 to 0.77 and specificity from 0.678 to 0.912[41, 42]. In our study, the AUC for using PIVKA-II alone was 0.884 (95%CI, 0.857–0.911) with sensitivity of 0.46 and specificity of 0.837 using training set data, which was consistent with the previous studies. However, warfarin usage increased alongside the worldwide ageing population, the value of PIVKA-II in HCC diagnostic will gradually decrease. A number of lines of evidence have indicated that PIVKA-II had no sufficient sensitivity as single markers for routine use in HCC surveillance[43, 44].

Currently, the combination of multiple indicators has been investigated as candidate HCC biomarker. However, most of these models often missed the application value on AFPN-HCC which was composed of AFP and other biomarkers. For the detection of early stage HCC, the GALAD models has shown a specificity of 81.6%-93.3% and sensitivity of 80.2%-85.6% in different populations.[20] In addition, there was a consistent view that the combination of biomarkers is superior to the use of a single biomarker. The GAAP model had a 0.888 AUC in distinguish AFPN-HCC from chronic hepatitis liver disease and performed similarly to that of the GALAD score.[24] But in this study, the control groups concluded CLD only and required a multicenter large-scale validation to verify the results. In our study, we compared with GAAP model using training and validation set, and the results showed that the AUC of GAAP model was 0.763(95%CI,0.682–0.844) and 0.882((95%CI, 0.851–0.9135) which was lower than our model. Additionally, among the comparison of the diagnostic efficacy of PIVKA-II and the GAAP model, our model showed the best combined sensitivity and specificity.

Another advantage of our model is easy to compute without the use of complicated formula, after visualization as a nomogram. The nomogram provided easy-to-understand clinical tools which higher total points had a greater probability of being diagnosed as AFPN-HCC. For example, a 65-years-old (79 points), his laboratory examination result found that PIVKA-II is 455mAU/ml (assignment is 4, 100 points), PLT is 185×109/L (67 points), PT is 11.8s (78 points), then a total point value of 324points is given, which corresponds to a 0.989 probability being diagnostic as HCC.

There remain insufficiencies in this research. Firstly, we concede that our study was only a retrospective analysis, but this analysis will serve as a basis for a prospective trial. Then, further larger multi-centric studies are still required to ensure generalizability.

In conclusion, we assessed age, gender and 50 commonly used laboratory index and identified a combination of biomarkers that may be of use in the diagnosis of AFPN-HCC. The model partially filled the diagnostic blind area of AFPN-HCC and showed potential for further improving detection rates for AFPN- HCC.

Declarations

Disclosure

The authors declare no conflict of interest.

Funding

This work was supported by grants from National Natural Science Foundation of China (Grant number: 81802087), Joint Funds for the innovation of science and Technology, Fujian province (Grant number: 2019Y9012) and Fujian Provincial Health Technology Project (Grant number: 2021GGA023) Natural Science Foundation of Fujian Province (Grant number: 2022J01680).

References

1. Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. CA Cancer J Clin, 72(1), 7–33. doi: 10.3322/caac.21708
2. Gluer, A. M., Cocco, N., Laurence, J. M., Johnston, E. S., Hollands, M. J., Pleass, H. C.,.. . Lam, V. W. (2012). Systematic review of actual 10-year survival following resection for hepatocellular carcinoma. HPB (Oxford), 14(5), 285–290. doi: 10.1111/j.1477-2574.2012.00446.x
3. Lee, Y. T., Fujiwara, N., Yang, J. D., & Hoshida, Y. (2022). Risk stratification and early detection biomarkers for precision HCC screening. Hepatology (Baltimore, Md.), 10.1002/hep.32779. Advance online publication. https://doi.org/10.1002/hep.3277
4. Heimbach, J. K., Kulik, L. M., Finn, R. S., Sirlin, C. B., Abecassis, M. M., Roberts, L. R.,.. . Marrero, J. A. (2018). AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology, 67(1), 358–380. doi: 10.1002/hep.29086
5. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. (2012). J Hepatol, 56(4), 908–943. doi: 10.1016/j.jhep.2011.12.001
6. Simmons, O., Fetzer, D. T., Yokoo, T., Marrero, J. A., Yopp, A., Kono, Y.,.. . Singal, A. G. (2017). Predictors of adequate ultrasound quality for hepatocellular carcinoma surveillance in patients with cirrhosis. Aliment Pharmacol Ther, 45(1), 169–177. doi: 10.1111/apt.13841
7. Tzartzeva, K., Obi, J., Rich, N. E., Parikh, N. D., Marrero, J. A., Yopp, A.,.. . Singal, A. G. (2018). Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. Gastroenterology, 154(6), 1706–1718.e1701. doi: 10.1053/j.gastro.2018.01.064

8. Ichikawa, T., Sano, K., & Morisaka, H. (2014). Diagnosis of Pathologically Early HCC with EOB-MRI: Experiences and Current Consensus. Liver cancer, 3(2), 97–107. doi:10.1159/000343865\
9. Zheng, Y., Zhu, M., & Li, M. (2020). Effects of alpha-fetoprotein on the occurrence and progression of hepatocellular carcinoma. Journal of cancer research and clinical oncology, 146(10), 2439–2446. doi.:10.1007/s00432-020-03331-6
10. Baig, J. A., Alam, J. M., Mahmood, S. R., Baig, M., Shaheen, R., Sultana, I., & Waheed, A. (2009). Hepatocellular carcinoma (HCC) and diagnostic significance of A-fetoprotein (AFP). J Ayub Med Coll Abbottabad, 21(1), 72–75.
11. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. (2018). J Hepatol, 69(1), 182–236. doi: 10.1016/j.jhep.2018.03.019
12. Chen, H., Zhang, Y., Li, S., Li, N., Chen, Y., Zhang, B.,.. . Dai, M. (2018). Direct comparison of five serum biomarkers in early diagnosis of hepatocellular carcinoma. Cancer Manag Res, 10, 1947–1958. doi: 10.2147/cmar.S167036
13. Feng, H., Li, B., Li, Z., Wei, Q., & Ren, L. (2021). PIVKA-II serves as a potential biomarker that complements AFP for the diagnosis of hepatocellular carcinoma. BMC Cancer, 21(1), 401. doi: 10.1186/s12885-021-08138-3
14. Qi, F., Zhou, A., Yan, L., Yuan, X., Wang, D., Chang, R.,.. . Zhang, X. (2020). The diagnostic value of PIVKA-II, AFP, AFP-L3, CEA, and their combinations in primary and metastatic hepatocellular carcinoma. J Clin Lab Anal, 34(5), e23158. doi: 10.1002/jcla.23158
15. Kondo, A., Kondo, H., Nakagawa, Y., Ito, H., Shimomura, D., Hatanaka, N.,.. . Matsuo, S. (2020). Influence of Warfarin Therapy on Prothrombin Production and Its Posttranslational Modifications. J Appl Lab Med, 5(6), 1216–1227. doi: 10.1093/jalm/jfaa069
16. Toyoda, H., Kumada, T., & Tada, T. (2011). Highly sensitive Lens culinaris agglutinin-reactive α-fetoprotein: a new tool for the management of hepatocellular carcinoma. Oncology, 81 Suppl 1, 61–65. doi: 10.1159/000333263
17. Hu, B., Tian, X., Sun, J., & Meng, X. (2013). Evaluation of individual and combined applications of serum biomarkers for diagnosis of hepatocellular carcinoma: a meta-analysis. Int J Mol Sci, 14(12), 23559–23580. doi: 10.3390/ijms141223559
18. Toyoda, H., Kumada, T., & Tada, T. (2011). Highly sensitive Lens culinaris agglutinin-reactive α-fetoprotein: a new tool for the management of hepatocellular carcinoma. Oncology, 81 Suppl 1, 61–65. https://doi.org/10.1159/000333263
19. Caviglia, G. P., Abate, M. L., Petrini, E., Gaia, S., Rizzetto, M., & Smedile, A. (2016). Highly sensitive alpha-fetoprotein, Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein and des-gamma-carboxyprothrombin for hepatocellular carcinoma detection. Hepatol Res, 46(3), E130-135. doi: 10.1111/hepr.12544
20. Sterling, R. K., Jeffers, L., Gordon, F., Venook, A. P., Reddy, K. R., Satomura, S.,.. . Sherman, M. (2009). Utility of Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein and des-gamma-carboxy

prothrombin, alone or in combination, as biomarkers for hepatocellular carcinoma. Clin Gastroenterol Hepatol, 7(1), 104–113. doi: 10.1016/j.cgh.2008.08.041

21. Berhane S, Toyoda H, Tada T, Kumada T, Kagebayashi C, Satomura S, et al. Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. Clin Gastroenterol Hepatol. 2016;14(6):875 – 86.e6. doi: 10.1016/j.cgh.2015.12.042.
22. Berhane, S., Toyoda, H., Tada, T., Kumada, T., Kagebayashi, C., Satomura, S.,.. . Johnson, P. (2016). Role of the GALAD and BALAD-2 Serologic Models in Diagnosis of Hepatocellular Carcinoma and Prediction of Survival in Patients. Clin Gastroenterol Hepatol, 14(6), 875–886.e876. doi: 10.1016/j.cgh.2015.12.042
23. Singal, A. G., Tayob, N., Mehta, A., Marrero, J. A., El-Serag, H., Jin, Q.,.. . Parikh, N. D. (2022). GALAD demonstrates high sensitivity for HCC surveillance in a cohort of patients with cirrhosis. Hepatology, 75(3), 541–549. doi: 10.1002/hep.32185
24. Liu, M., Wu, R., Liu, X., Xu, H., Chi, X., Wang, X.,.. . Niu, J. (2020). Validation of the GALAD Model and Establishment of GAAP Model for Diagnosis of Hepatocellular Carcinoma in Chinese Patients. J Hepatocell Carcinoma, 7, 219–232. doi: 10.2147/jhc.S271790
25. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Bmj, 350, g7594. doi: 10.1136/bmj.g7594
26. Chinese Society of Hepatology, Chinese Medical Association; Chinese Society of Infectious Diseases, Chinese Medical Association. (2022). Guidelines for the prevention and treatment of chronic hepatitis B (version 2022) Zhonghua Gan Zang Bing Za Zhi. 2022;30(12):1309–1331. doi:10.3760/cma.j.cn501113-20221204-00607
27. Terrault, N. A., Lok, A. S. F., McMahon, B. J., Chang, K. M., Hwang, J. P., Jonas, M. M.,.. . Wong, J. B. (2018). Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. Hepatology, 67(4), 1560–1599. doi: 10.1002/hep.29800
28. Xu, X. Y., Ding, H. G., Li, W. G., Xu, J. H., Han, Y., Jia, J. D.,.. . Zhuang, H. (2020). Chinese guidelines on the management of liver cirrhosis (abbreviated version). World J Gastroenterol, 26(45), 7088–7103. doi: 10.3748/wjg.v26.i45.7088
29. Yoshiji, H., Nagoshi, S., Akahane, T., Asaoka, Y., Ueno, Y., Ogawa, K.,.. . Koike, K. (2021). Evidence-based clinical practice guidelines for Liver Cirrhosis 2020. J Gastroenterol, 56(7), 593–619. doi: 10.1007/s00535-021-01788-x
30. Kulik, L., & El-Serag, H. B. (2019). Epidemiology and Management of Hepatocellular Carcinoma. Gastroenterology. 2019;156(2):477 – 91.e1. doi: 10.1053/j.gastro.2018.08.065.
31. Amitrano, L., Guardascione, M. A., Brancaccio, V., & Balzano, A. (2002). Coagulation disorders in liver disease. Semin Liver Dis, 22(1), 83–96. doi: 10.1055/s-2002-23205
32. Matchar, D. B., Love, S. R., Jacobson, A. K., Edson, R., Uyeda, L., Phibbs, C. S., & Dolor, R. J. (2015). The impact of frequency of patient self-testing of prothrombin time on time in target range within VA

Cooperative Study #481: The Home INR Study (THINRS), a randomized, controlled trial. J Thromb Thrombolysis, 40(1), 17–25. doi: 10.1007/s11239-014-1128-8

33. Serag, H. B., Kanwal, F., Davila, J. A., Kramer, J., & Richardson, P. (2014). A new laboratory-based algorithm to predict development of hepatocellular carcinoma in patients with hepatitis C and cirrhosis. Gastroenterology, 146(5), 1249–1255.e1241. doi: 10.1053/j.gastro.2014.01.045
34. Ali, O. M., El Amin, H. A., Sharkawy, Y. L., Mohamed Ali, A. A., Kholef, E. F. M., & Elsewify, W. A. E. (2020). Golgi Protein 73 versus Alpha-Fetoprotein as a New Biomarker in Early Diagnosis of Hepatocellular Carcinoma. Int J Gen Med, 13, 193–200. doi: 10.2147/ijgm.S253622
35. Liu, C. H., Gil-Gómez, A., Ampuero, J., & Romero-Gómez, M. (2018). Diagnostic accuracy of SCCA and SCCA-IgM for hepatocellular carcinoma: A meta-analysis. Liver Int, 38(10), 1820–1831. doi: 10.1111/liv.13867
36. Hong, Y., Long, J., Li, H., Chen, S., Liu, Q., Zhang, B.,.. . Huang, J. (2015). An Analysis of Immunoreactive Signatures in Early Stage Hepatocellular Carcinoma. EBioMedicine, 2(5), 438–446. doi: 10.1016/j.ebiom.2015.03.010
37. Anatelli, F., Chuang, S. T., Yang, X. J., & Wang, H. L. (2008). Value of glypican 3 immunostaining in the diagnosis of hepatocellular carcinoma on needle biopsy. Am J Clin Pathol, 130(2), 219–223. doi: 10.1309/wmb5px57y4p8qcty
38. Chalasani NP, Ramasubramanian TS, Bhattacharya A, Olson MC, Edwards VD, Roberts LR, et al. (2021).A Novel Blood-Based Panel of Methylated DNA and Protein Markers for Detection of Early-Stage Hepatocellular Carcinoma. Clin Gastroenterol Hepatol. 19(12):2597 – 605.e4. doi: 10.1016/j.cgh.2020.08.065.
39. Liu, W. H., Ren, L. N., Wang, X., Wang, T., Zhang, N., Gao, Y., Luo, H., Navarro-Alvarez, N., & Tang, L. J. (2015). Combination of exosomes and circulating microRNAs may serve as a promising tumor marker complementary to alpha-fetoprotein for early-stage hepatocellular carcinoma diagnosis in rats. Journal of cancer research and clinical oncology, 141(10), 1767–1778. doi:10.1007/s00432-015-1943-0
40. Perne, M. G., Sitar-Tăut, A. V., Alexescu, T. G., Ciumărnean, L., Milaciu, M. V., Coste, S. C.,.. . Crăciun, A. (2023). Diagnostic Performance of Extrahepatic Protein Induced by Vitamin K Absence in the Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. Diagnostics (Basel), 13(5). doi: 10.3390/diagnostics13050816
41. Lim TS, Kim DY, Han KH, Kim HS, Shin SH, Jung KS, et al.(2016). Combined use of AFP, PIVKA-II, and AFP-L3 as tumor markers enhances diagnostic accuracy for hepatocellular carcinoma in cirrhotic patients. Scand J Gastroenterol. 51(3):344–53. doi: 10.3109/00365521.2015.1082190.
42. Poté, N., Cauchy, F., Albuquerque, M., Voitot, H., Belghiti, J., Castera, L.,.. . Paradis, V. (2015). Performance of PIVKA-II for early hepatocellular carcinoma diagnosis and prediction of microvascular invasion. J Hepatol, 62(4), 848–854. doi: 10.1016/j.jhep.2014.11.005
43. Park, S. J., Jang, J. Y., Jeong, S. W., Cho, Y. K., Lee, S. H., Kim, S. G.,.. . Bang, H. I. (2017). Usefulness of AFP, AFP-L3, and PIVKA-II, and their combinations in diagnosing hepatocellular carcinoma.

Medicine (Baltimore), 96(11), e5811. doi: 10.1097/md.0000000000005811

44. Piratvisuth, T., Tanwandee, T., Thongsawat, S., Sukeepaisarnjaroen, W., Esteban, J. I., Bes, M.,.. . Chan, H. L. (2022). Multimarker Panels for Detection of Early Stage Hepatocellular Carcinoma: A Prospective, Multicenter, Case-Control Study. Hepatol Commun, 6(4), 679–691. doi: 10.1002/hep4.1847

Figures

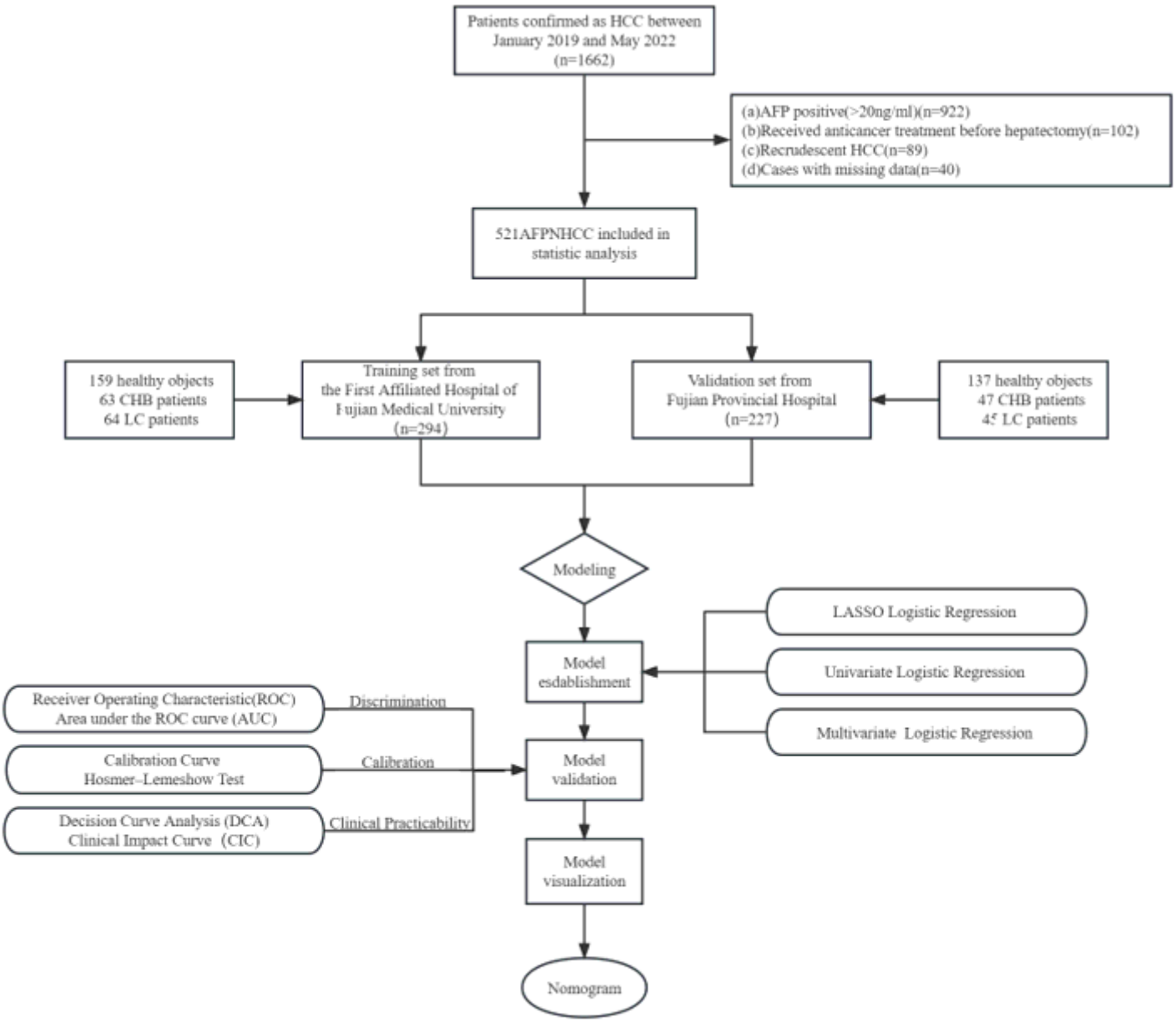

Figure 1

Procedure of study flowchart

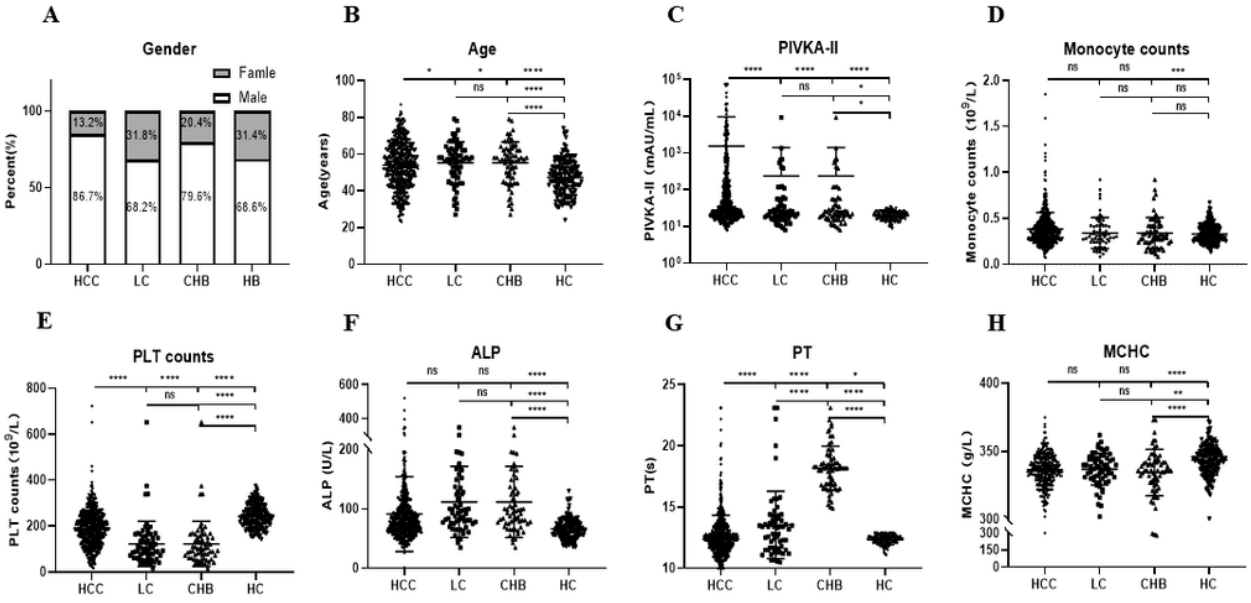

Figure 2

Selected candidate indicators in HCC and non-HCC groups

For gender(A), Chi-square tests were performed. For age(B), monocyte counts(D) , PLT counts(E) and ALP(F), the multiple pairwise comparisons were made using Tukey's Method. For PIVKA-II(C), PT(G) and MCHC(H), Kruskal–Wallis tests were used for comparisons among groups. ***P <0.001, **P <0.01, *P <0.05, ns P>0.05, ns, no significance.

Abbreviations: HCC, hepatocellular carcinoma; LC, liver cirrhosis; HC, healthy controls; CHB, chronic hepatitis B; PIVKA-II, protein induced by Vitamin K absence or antagonist-II; PLT, platelet; ALP, alkaline phosphatase; PT, prothrombin time; MCHC, mean corpuscular hemoglobin concentration

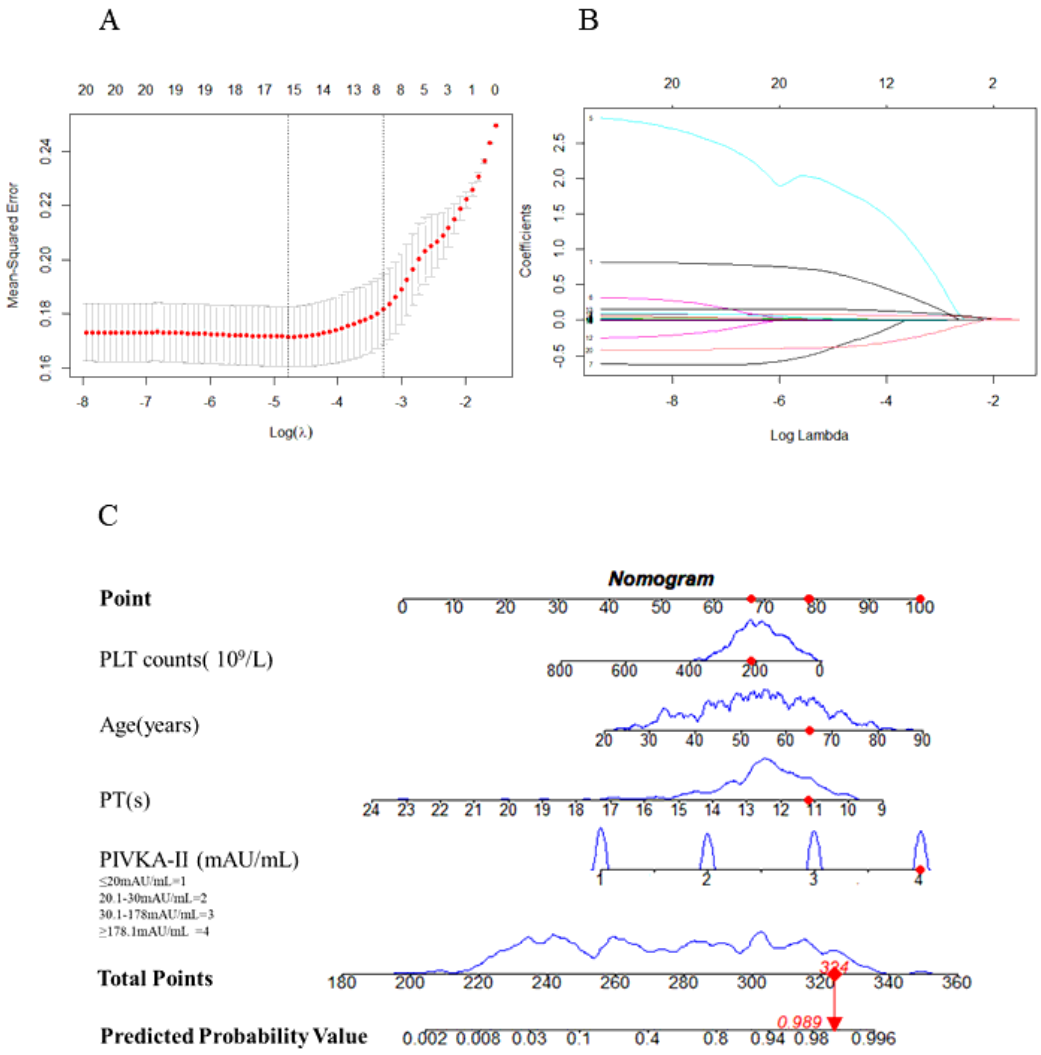

Figure 3

Developing and visualized of diagnosis model of AFPN-HCC

(A)Tuning parameter (λ) selection in the LASSO model used 10-fold cross-validation via minimum criteria. Dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1 standard error of the minimum criteria. (B) LASSO coefficient profiles of the 17 factors. A coefficient

profile plot was generated against the log (Lambda) sequence. Vertical line represents the values selected where optimal lambda resulted in 8 nonzero coefficients. (C) The nomogram of the model. The red dot represents the patient's characteristics on each variable axis and they are projected to the top line to get the corresponded point. Each summation point corresponds to a predicted probability value in the horizontal line on the bottom of the nomogram.

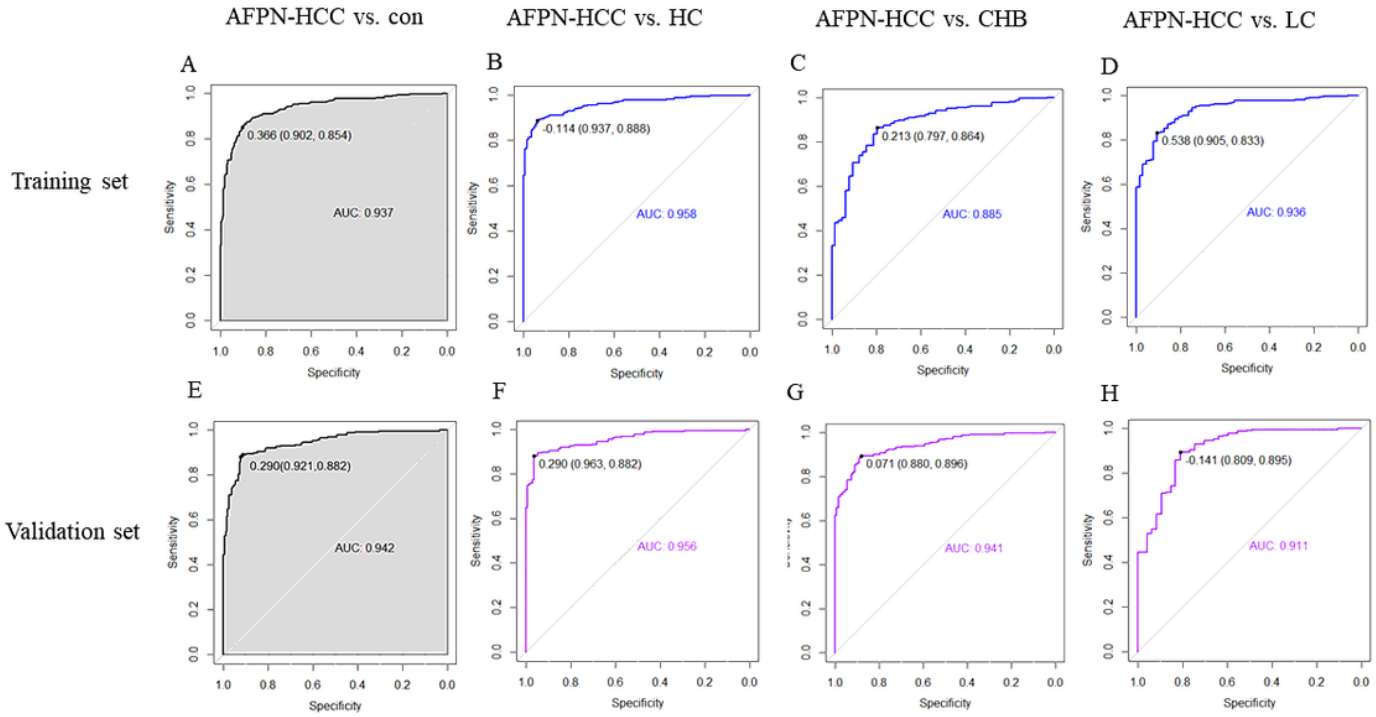

Figure 4

The ROC of the nomogram

The ROCs represent the performance of established nomogram model for discriminating AFPN-HCC from whole controls, HC, CHB and HC in training set. (A-D) and validation set(E-H). The AUC of each ROC has been marked in the figure. The number under the curve represents cut-off (Specificity, Sensitivity).

Abbreviations：AFPN-HCC, AFP-negative- hepatocellular carcinoma; HC, healthy control; LC, liver cirrhosis; CHB, chronic hepatitis B; AUC, area under the curve.

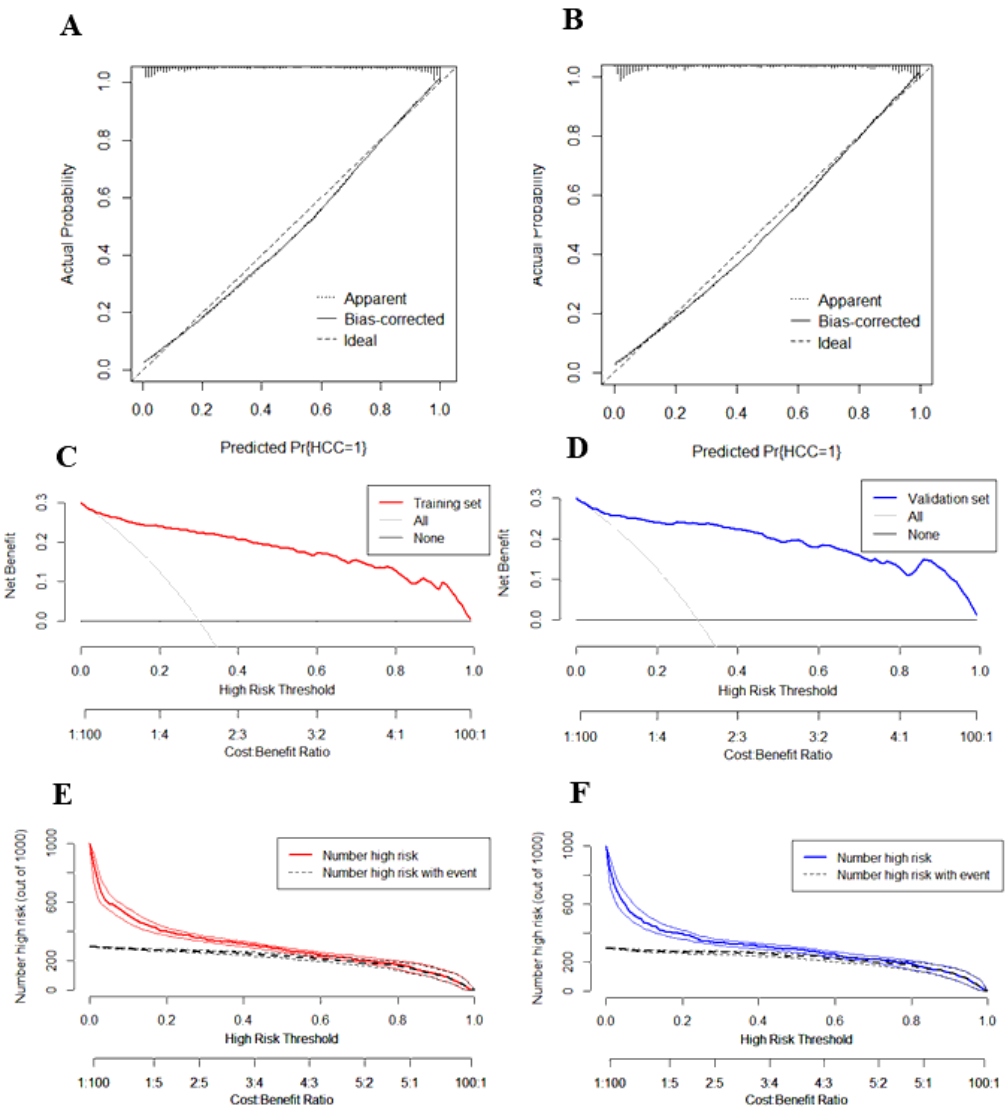

Figure 5

Calibration plot, DCA and CIC in training sets(A) and validation sets(B).

The dotted line represents a perfect prediction by an ideal model, and the solid black line shows the performance of the model. DCA in training set (C)and validation set(D). The x-axis represents the threshold probability and the y-axis shows the net benefit. The dark horizontal line means one extreme

situation that all samples are negative and not treated, with a net benefit of zero. The gray-dotted line indicates the other extreme situation that all patients suffered AFPN-HCC. Clinical impact curve (CIC) in training set training set(D) and validation set(E). The number of high-risk patients and the number of high-risk patients with event were plotted by different threshold probability in a population.

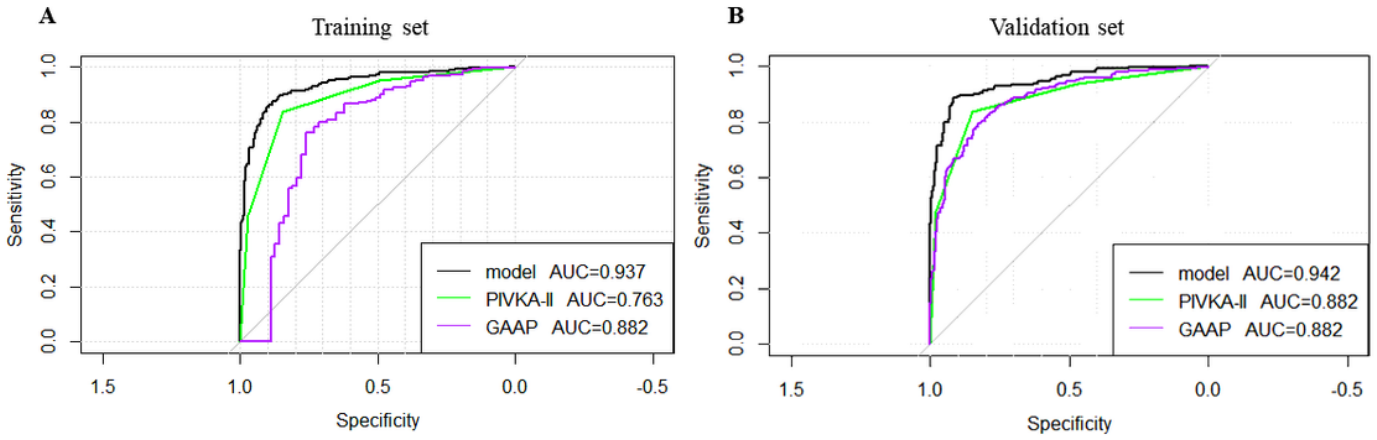

Figure 6

The ROC of the model, PIVKA-II and GAAP model in training sets(A) and validation sets(B).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.docx