ORIGINAL RESEARCH

Validation of the GALAD Model and Establishment of GAAP Model for Diagnosis of Hepatocellular Carcinoma in Chinese Patients

Miaoxia Liu[1,2,*]
Ruihong Wu [ID][1,3,*]
Xu Liu[1]
Hongqin Xu[1]
Xiumei Chi [ID][1,3]
Xiaomei Wang[1]
Mengru Zhan [ID][1]
Bao Wang[1]
Fei Peng [ID][1]
Xiuzhu Gao [ID][1,3]
Ying Shi[1]
Xiaoyu Wen[1]
Yali Ji[2]
Qinglong Jin[1]
Junqi Niu[1]

[1]Department of Hepatology, First Hospital of Jilin University, Changchun, Jilin Province 130021, People's Republic of China; [2]Hepatology Unit, Department of Infectious Diseases, Nanfang Hospital, Southern Medical University, Guangzhou 510515, People's Republic of China; [3]Phase I Clinical Research Center, First Hospital of Jilin University, Changchun, Jilin Province 130021, People's Republic of China

*These authors contributed equally to this work

Correspondence: Qinglong Jin
Department of Hepatology, First Hospital of Jilin University, 71 Xin Min Street, Changchun 130021, People's Republic of China
Tel/Fax +86-431-81875103
Email jinql2016@163.com

Purpose: GALAD is a statistical model for estimating the likelihood of having hepatocellular carcinoma (HCC) based on gender, age, AFP, AFP-L3, and PIVKA-II. We aimed to assess its performance and build new models in China, where hepatitis B virus (HBV) is the leading etiology of HCC.

Patients and Methods: We built the GALAD-C model with the same five variables in GALAD, and the GAAP model with gender, age, AFP, and PIVKA-II, using logistic regression based on 242 patients with HCC and 283 patients with chronic liver disease (CLD). We also collected 50 patients with other malignant liver tumors (OMTs) and 50 healthy controls (HCs). A test dataset (169 patients with HCC and 139 with CLD) was used to test the performance of GAAP.

Results: The GALAD-C and GAAP models achieved comparable performance (area under the receiver operating characteristic curve [AUC], 0.922 vs 0.914), and both were superior to GALAD, PIVKA-II, AFP, and AFP-L3% (AUCs, 0.891, 0.869, 0.750, and 0.711) for discrimination of HCC from CLD for the entire dataset. The AUCs of the GALAD, GALAD-C and GAAP models were excellent for the hepatitis C virus (HCV) subgroup (0.939, 0.958 and 0.954), and for discrimination HCC from HCs (0.988, 0.982, and 0.979), but were relatively lower for the HBV subgroup (0.855, 0.904, and 0.894), and for HCC within Milan Criteria (0.810, 0.841, and 0.840). They were not superior to AFP (0.873) for discrimination of HCC from OMT (0.873, 0.809, and 0.823). GAAP achieved an AUC of 0.922 in the test dataset.

Conclusion: GALAD was excellent for discrimination of HCC from CLD in the HCV subgroup of a cohort of Chinese patients. The GAAP and GALAD-C models achieved better performance compared with GALAD. These three models exhibited better performance in patients with an HCV etiology than those with HBV.

Keywords: hepatocellular carcinoma, alpha-fetoprotein, PIVKA-II, GALAD

Introduction

Liver cancer is predicted to be the sixth most commonly diagnosed cancer and the fourth leading cause of cancer death worldwide in 2018, with about 841,000 new cases and 782,000 deaths annually.[1] The etiology of hepatocellular carcinoma (HCC) varies in different regions worldwide. Non-viral hepatitis (alcoholic hepatitis and non-alcoholic fatty liver disease) accounts for most cases of HCC in Western countries. Most cases of HCC in Japan are associated with chronic hepatitis C (CHC), while most cases of HCC in China are associated with chronic hepatitis

B (CHB) and CHC. Cases in China accounted for about 50% of the new liver cancer cases and deaths that occurred worldwide during 2012.[2] About 25% of cirrhosis patients die of HCC because tumors are frequently detected at an advanced symptomatic stage at which effective treatment options are limited.[3] To achieve early detection of HCC in asymptomatic patients, guidelines recommend monitoring high-risk patients for HCC to improve outcomes.[4] Abdominal ultrasound is usually recommended for HCC surveillance, and serum alpha-fetoprotein (AFP) level is included in some guidelines. However, the sensitivity and specificity values of ultrasound for HCC detection range from 40% to 80% and 80% to 100%, respectively,[5] and the sensitivity (55–74%) and specificity (84–91%) values of AFP are moderate at best.[6] AFP-L3, *Lens culinaris* agglutinin-reactive fraction of AFP, is one of the three glycoforms of AFP, and is the major glycoform elevated in the serum of HCC patients.[7] The Japan Society of Hepatology HCC guidelines for surveillance recommend the use of AFP-L3%.[8]

Des-gamma-carboxy prothrombin (DCP) is also known as a protein induced by vitamin K absence or antagonists II (PIVKA-II). This abnormal prothrombin molecule is produced in malignant cells as the result of an acquired defect in posttranslational carboxylation of the prothrombin precursor.[9] PIVKA-II has higher sensitivity and specificity values than AFP for HCC diagnosis; combining PIVKA-II with AFP increases these values even more.[10,11] PIVKA-II is recommended as a tumor marker for HCC surveillance by the American Association for the Study of Liver Diseases (AASLD), the European Association for the Study of Liver (EASL), and the Japan Society of Hepatology, and has been widely used in those regions. AFP is not recommended by AASLD and EASL because of its lower sensitivity and specificity values. However, AFP has been widely used by most physicians in China, while PIVKA-II has been unrecognized and has had only limited use. There are limited data about PIVKA-II use for HCC caused by the two different chronic viral hepatitis diseases in China.

In 2013, Johnson P. J. et al developed a serum-based tool (ie, the GALAD model and associated GALAD score) for surveillance of HCC based on a UK cohort. The GALAD score is based on gender, age, and the three serologic biomarkers AFP, AFP-L3, and PIVKA-II.[12] Performance of this model has been validated in the UK, Germany, Japan,[13] and the USA (Mayo Clinic and the National Cancer Institute Early Detection Research Network).[14] The GALAD score can detect HCC with high levels of accuracy in patients with

non-alcoholic steatohepatitis, with and without cirrhosis.[15] The latest AASLD guidelines note the potential for use of the GALAD score for HCC surveillance.[16] However, the model has not been validated in China, where hepatitis B virus (HBV) infection is the most common etiology of HCC.

Our study aimed 1) to compare the levels of PIVKA-II, AFP, and AFP-L3 in patients with different liver diseases, 2) to evaluate the performance of the GALAD model for HCC detection, and 3) to build new models in Chinese patients.

Patients and Methods
Design and Patients
All subjects were enrolled between 2012 and 2016 when they were patients at the First Hospital of Jilin University. A dataset (ie, training set) consisting of 242 patients with HCC and 283 patients with chronic liver disease (CLD) was used to test the GALAD model and build the GALAD-C and GAAP models. The 283 patients with CLD consisted of 187 with cirrhosis and 96 with chronic hepatitis (CH). We also collected data from 50 patients with other malignant liver tumors (OMTs; 23 with intrahepatic cholangiocarcinoma (ICC) and 27 with liver metastases) and 50 healthy controls (HCs). Another dataset (ie, test set) consisting of 308 patients (169 patients with HCC and 139 patients with CLD) was used to test the new model.

Demographic data, clinical characteristics, diagnostic and laboratory results (AFP values for the majority of patients, liver function, and routine blood analysis) were acquired from the clinical database.

Informed consent was obtained from all subjects. The study protocol was approved by the Ethics Committee of the First Hospital of Jilin University.

Inclusion Criteria
The diagnostic criteria for HCC were: (1) histopathological diagnosis of HCC; or (2) results consistent with the diagnostic criteria recommended by the Asian Pacific Association for the Study of the Liver guidelines if no pathological examination results were available: A) computed tomography (CT), magnetic resonance imaging (MRI), or contrast-enhanced ultrasound results revealing typical imaging lesions of HCC, and the lesion site had typical blood flow changes; or B) suspected small nodules found using CT, MRI, or contrast-enhanced ultrasound were confirmed by positron emission tomography examination. Early-stage HCC was defined according to the

Milan Criteria (single lesion ≤5 cm or up to three separate lesions, none larger than 3 cm).[17]

The diagnosis of chronic hepatitis B (chronic hepatitis C) and related cirrhosis was in accordance with the updated guidelines for the prevention and treatment of CHB (hepatitis C) infection from the Chinese Society of Hepatology.[18,19] Briefly, the diagnostic criteria for chronic hepatitis patients were: HBV infection (HBsAg-positive or HBV DNA positive) or hepatitis C virus (HCV) infection (HCV RNA+ and HCV antibody+) for at least 6 months, alanine aminotransferase (ALT) is persistently or repeatedly elevated, or hepatitis lesions are identified by liver biopsy. The diagnostic criteria for liver cirrhosis were: (1) histopathological diagnosis, or (2) if there was no histological diagnosis, ultrasound, CT, or MRI imaging results indicated splenomegaly without liver space-occupying lesions. We graded the liver cirrhosis according to the Child-Pugh classification system, which is a scoring system to grade liver function and is currently included in most of the available HCC treatment guidelines. A score of 5–6 was defined as class A, 7–9 as class B, and 10–15 as class C liver function.[20]

The criteria for the diagnosis of liver metastasis were: A) a history of primary cancer and clinical manifestation of liver tumors; B) imaging examination revealed substantial hepatic space-occupying lesions that were mostly scattered or multiple; C) liver metastases were found during surgery for the primary disease and the diagnosis was based on histopathology examination results.

The criteria for the diagnosis of ICC were: (1) histopathological diagnosis of ICC; or (2) there were no pathology results, but there were typical clinical manifestations, biochemical changes, and elevated CA-199. Enhanced CT and MRI results were consistent with ICC characteristics.

The criteria for inclusion in the Healthy control group were: (1) no family history of cancer and no history of diagnosis and treatment of liver-related disease; and (2) serological markers indicated no HBV infection currently or previously and anti-HCV antibody test was negative; and (3) the results of liver function, kidney function, and routine blood tests indicated no abnormalities; and (4) ultrasound results indicated no abnormalities in liver or gallbladder systems; and (5) no abnormalities indicated by results of liver fibroscan.

Exclusion Criteria

The exclusion criteria were: (1) Current warfarin treatment; or (2) HCC, liver metastasis, and ICC treatment using surgery, ablation, radiotherapy, or chemotherapy; or (3) no frozen serum available or serum volume insufficient for biomarker assays; or (4) for patients with multiple admissions, the later time points were excluded. Only the first admission was included.

AFP, AFP-L3, and PIVKA-II Assays

We measured these biomarkers using different methods. This was different from the GALAD model, where the three biomarkers were assayed using the same method, the μTASWako i30 immunoanalyzer. In this study, serum AFP was measured using the Roche electrochemiluminescence immunoassay (ng/mL units). PIVKA-II was measured using the ARCHITECT immunoassay (mAU/Ml units) and frozen-thawed serum. AFP-L3 was measured using reagents from the Rejing Biotech company and frozen-thawed serum. The minimum quantitative limit of AFP-L3 was 1 ng/mL, and AFP-L3% was judged as negative if AFP-L3 <1 ng/mL. In the logistic regression analysis, we used 1% to represent AFP-L3 negative. AFP-L3 was not measured in the test dataset.

Statistical Methods

SPSS version 18.0 software (IBM/SPSS Inc., Chicago, IL), GraphPad Prism version 5.0 (for Windows, GraphPad Software, San Diego California, USA, www.graphpad.com), or R language (version 3.4.4) were used to perform the analyses or draw graphs, or both. Categorical data were expressed as numbers (percentages), and differences were tested using Chi-square tests. Skewed distribution continuous variables were presented as median and interquartile range values; differences were tested using the Mann–Whitney U-test or Kruskal–Wallis H-test. Post hoc Dunn's Multiple Comparison tests were used for pairwise comparisons. Spearman correlation analysis was used for correlation analyses.

Because the methods we used to measure the three serum markers were different from those in the original GALAD work, the units for PIVKA-II were mAU/mL, not ng/mL, and these units could not be exchanged, we rebuilt the model, called GALAD-C (C means China), using the same regression method (logistic regression) and the same five variables as covariates (Gender [1 for male, 0 for female], Age, AFP-L3%, log10[AFP] and log10[PIVKA-II]), with Gender as categorical variable and the other four as continuous ones. HCC is the dependent variable (1 for HCC, 0 for non-HCC). The estimated constant and coefficients for the five variables by logistic regression were used to construct the GALAD-C score. We built another model, called GAAP, only using four variables (Gender, Age, log10[AFP] and log10[PIVKA-II])

by the same procedure. The GAAP model was assessed in the test dataset.

Receiver operating characteristic (ROC) curve analysis was performed. The best cutoff values to discriminate HCC from CLD were determined based on the Youden index.[21] The subsequent analysis differentiating HCC from other conditions shared the same cutoff values. The results for sensitivity, specificity, positive predictive value, negative predictive value, and percent correctly classified were calculated. When ROC curves for models/markers/marker combinations were performed on the same individuals, the between-AUC comparisons were performed using the method proposed by Delong.[22] This nonparametric statistical test compared the areas under two or more correlated receiver operating curves. The roc.test function in pROC package in R language was used to test the ROC curves with the parameter paired = TRUE and method = "delong". When ROC curves were compared between HCV and HBV etiologies, the differences were tested by using the same function but the parameter paired = FALSE.

All statistical tests were two-sided. A P value ≤0.05 was considered to be statistically significant.

Results
Demographic and Clinical Characteristics of Enrolled Individuals
The results for demographic and clinical characteristics of the individuals in the training set and test set are presented in Tables 1 and 2, respectively.

Serum AFP, PIVKA-II, and AFP-L3% in HCC, CLD, OMT, and HC
The median level of AFP in the HCC group was significantly higher than that in the other groups (Table 1, Figure 1A). There was no significant difference between the Cirrhosis and Chronic Hepatitis groups, which were both significantly higher than in the OMT and HC groups. No significant difference was found between the OMT and HC groups.

The median level of PIVKA-II in the HCC group was significantly higher than that in all the other groups (Table 1, Figure 1B). No significant differences were found among Cirrhosis, Chronic Hepatitis, and HC groups. The PIVKA-II level in the OMT group was significantly higher than that in the Cirrhosis, Chronic Hepatitis, and HC groups. The ICC patients had relatively higher PIVKA-II values (median [interquartile range], 41.7 [29.1–80.2] mAU/mL, Figure 1C).

The percentage of AFP-L3 ≥10% in the HCC group was significantly higher than that in all other groups (Table 1, Figure 1D).

There was a weak correlation between AFP and PIVKA-II levels (Spearman r =0.323, P < 0.001, Supplementary Figure S1). These results suggested that the accuracy of HCC diagnosis could be improved by a combination of the two biomarkers.

Serum AFP and PIVKA-II in HCV-Related Liver Disease and HBV-Related Liver Disease
The results for serum AFP and PIVKA-II levels in the context of HCV and HBV etiologies are presented in Figure 1E and F; Supplementary Table S1. In the HCV subgroup, consistent with the entire dataset, the AFP and PIVKA-II levels in HCC group were significantly higher than those in the Cirrhosis and Hepatitis groups. In the HBV subgroup, the PIVKA-II results were consistent with the entire dataset, while the difference of AFP between the HCC and Hepatitis groups was not significant; the AFP level was significantly higher in the Hepatitis group than in the Cirrhosis group. We also compared the HCV and HBV groups. There were no significant differences between the HBV-related HCC and HCV-related HCC groups (25.28 vs 36.68 ng/mL, respectively, P > 0.05), or between the HBV-related Cirrhosis and HCV-related Cirrhosis groups (3.62 vs 5.01 ng/mL, respectively, P > 0.05). The AFP level in the CHB group was higher than that in the CHC group (12.0 vs 4.2 ng/mL, respectively, P < 0.05).

The higher AFP levels in the CHB group may be associated with liver inflammation. In this study, the ALT level in the CHB group was significantly higher than that in the HCC, HBV-related Cirrhosis, and CHC groups (median: 175 U/L, 42U/L, 35U/L, and 72 U/L, respectively, all P values <0.05) (supplementary Figure S2). There was also a weak correlation between ALT and AFP levels (Spearman r = 0.2067, P < 0.0001). However, no correlation was found between ALT and PIVKA-II levels (Spearman r = 0.0229, P = 0.5906) (supplementary Figure S3).

Serum AFP and PIVKA-II in HCC Group, with and without Cirrhosis and Different Child-Pugh Scores
No statistically significant differences were found in median AFP or PIVKA-II levels between the HCC with

Table 1 Characteristics of the Study Subjects Used to Evaluate the GALAD Model and Construct the GAAP Model

Variables	HCC	CLD			Other Malignant Liver Tumor	HC
		Total	Cirrhosis	Hepatitis		
N	242	283	187	96	50	50
Age, year	59 (54–65)	52 (44–60)	54 (48–62)	47(34–55)	60 (55–63)	49 (40–55)
Sex, male, n(%)	176 (72.7)	167 (59.0%)	111 (59.4)	56 (58.3)	32 (64.0)	7 (14.0)
Etiology, n (%)						
HBV	135 (55.8)	149 (52.7)	102 (54.5)	47 (49.0)	NA	0
HCV	106 (43.8)	133 (47.0)	84 (44.9)	49 (51.0)	NA	0
Alcohol	1 (0.4%)	1 (0.4)	1 (0.5%)	N/A	NA	0
ALT, U/L	42 (28–65)	51 (27–116)	39 (24–63)	121 (61–209)	55 (25–140)	27 (20–30)
AST, U/L	51 (34–76)	54 (34–96)	48 (31–85)	69 (42–142)	52 (28–124)	25 (20–28)
Total Bilirubin, μmol/L	21 (15–32)	22 (14–36)	24 (15–40)	18 (14–28)	22 (11–185)	13 (11–16)
Albumin, g/L	29 (34–38)	34 (28–39)	31 (27–37)	38 (35–41)	34 (31–39)	45 (43–48)
AFP, ng/mL	33.11 (7.17–492.9)	4.8 (2.7–17.1)	4.05 (2.43–14.0)	6.73 (3.31–26.68)	2.76 (2.06–4.27)	2.67 (1.53–3.87)
PIVKA-II, mAU/Ml	171.1 (33.8–1312.2)	21.4 (15.7–29.7)	18.4 (14.2–28.0)	23.8 (18.6–32.2)	37.4 (27.9–67.5)	24.6 (20.3–27.5)
AFP-L3, n(%)						
<10%,	155 (64.0)	273 (96.5)	182 (97.3)	91 (94.8)	49 (98.0)	50 (100)
[10–20) %,	29 (12.0)	6 (2.1)	3 (1.6)	3 (3.1)	1 (2.0)	0 (0)
[20–100)%,	58 (24.0)	4 (1.4)	2 (1.1)	2 (2.1)	0 (0)	0 (0)
Tumor characteristics						
Early HCC (within Milan criteria), n(%)	86 (35.5)	N/A	N/A	N/A	N/A	N/A
Solitary, n(%)	118(50.6), N=233	N/A	N/A	N/A	N/A	N/A
Maximum tumor size<5cm, n(%)	145(65.9), N=220	N/A	N/A	N/A	N/A	N/A
Portal Vein Invasion, n(%)	41(17.5), N=234	N/A	N/A	N/A	N/A	N/A
Metastasis, n(%)	10 (4.3), N=234	N/A	N/A	N/A	N/A	N/A

Note: All continuous variables are presented as median (interquartile range).
Abbreviations: HCC, hepatocellular carcinoma; CLD, chronic liver disease; HC, healthy controls; HBV, hepatitis B virus; HCV, hepatitis C virus; NA, not available; N/A, not applicable.

Cirrhosis and HCC without Cirrhosis groups, or among HCC groups with different Child-Pugh scores (supplementary Figure S4).

GALAD, GALAD-C, and GAAP Models for HCC Diagnosis

The GALAD-C model and the GAAP model were built based on the 242 HCC patients and the 283 CLD patients. The relevant parameters are presented in Supplementary Table S2.

The original GALAD model, the GALAD-C model and the GAAP model formulas were:

GALAD score = $-10.08 + 1.67 \times$ [Gender (1 for male, 0 for female)] $+ 0.09 \times$ [Age] $+ 0.04 \times$ [AFP-L3%] $+ 2.34 \times \log_{10}$ [AFP] $+ 1.33 \times \log_{10}$ [PIVKA-II].[12]

GALAD-C score = $-11.501 + 0.733 \times$ [Gender (1 for male, 0 for female)] $+ 0.099 \times$ [Age] $+ 0.073 \times$ [AFP-L3%] $+ 0.840 \times \log_{10}$ [AFP] $+ 2.364 \times \log_{10}$ [PIVKA-II].

GAAP score = $-11.203 + 0.699 \times$ [Gender (1 for male, 0 for female)] $+ 0.094 \times$ [Age] $+ 1.076 \times \log_{10}$ [AFP] $+ 2.376 \times \log_{10}$ [PIVKA-II].

The probability of an individual being classified as having HCC (range, 0 to 1) was calculated using: $P(HCC) = \exp[\text{score}]/(1 + \exp[\text{score}])$.

Table 2 Characteristics of Patients Used to Test the GAAP Model

Variables	HCC	CLD/Other Malignant Liver Tumor
N	169	139
Age, year	54 (48–61)	52 (43–59)
Sex, male, n(%)	128 (76)	83 (60)
Etiology, n(%)		
HBV	129 (76.3)	1 (0.7)
HCV	12 (7.1)	4 (2.9)
Alcohol	17 (10.1)	41 [28 Cirrhosis:13 hepatitis] (29.5)
PBC/PSC	0	26 (18.7)
NASH/NAFLD	0	16 (11.5)
Autoimmune	0	8 (5.8)
Parasite	0	9 (6.5)
DILI	0	24 (17.3)
ICC/MT	0	3 (2.2)
Unknown	11 (6.5)	7 (5.0)
Liver function tests		
ALT, U/L	40 (26–70)	38 (21–55)
AST, U/L	59 (40–105)	36 (29–65)
γ-GT, U/L	110 (48–253)	105 (45–261)
HCC biomarkers		
AFP, ng/mL	259 (7–2129)	3 (2–6)
PIVKA-II, mAU/mL	1158 (67–10,469)	25 (17–41)
Tumor characteristics		
Early HCC (within Milan criteria), n(%)	35 (22.6), N=154	N/A
Solitary, n(%)	84 (50.3), N=167	N/A
Maximum tumor size<5cm, n(%)	64 (41.6), N=154	N/A
Portal Vein Invasion, n(%)	81 (48.5), N=167	N/A
Metastasis, n(%)	18 (10.8), N=167	N/A

Note: All continuous variables are presented as median (interquartile range).
Abbreviations: HCC, hepatocellular carcinoma; CLD, chronic liver disease; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; NASH, non-alcoholic steatohepatitis; NAFLD, non-alcoholic fatty liver disease; DILI, drug-induced liver injury; ICC, intrahepatic cholangiocarcinoma; MT, liver metastases; NA, not available.

Performance of GALAD, GALAD-C, and GAAP for Discrimination of HCC and CLD

The GALAD model provided an area under the receiver operating characteristic curve (AUC) value of 0.891 (95% confidence interval (CI), 0.864–0.918) (Figure 2A, Table 3), which was much lower than the figure for the original UK cohort (AUC 0.97).[12] The GALAD-C model provided an AUC value of 0.922 (95% CI, 0.900–0.945). This result was significantly higher than that for the GALAD model ($P = 0.0005$) and was comparable with the GAAP model result (AUC = 0.914[0.889–0.938], $P = 0.0561$).

These three models found higher AUC values than the individual biomarkers, and their simple combinations, PIVKA-II (0.869), AFP (0.750), AFP-L3% (0.711), AFP+PIVKA-II (0.719), AFP+AFP+AFP-L3 (0.698), and AFP+PIVKA-II+AFP-L3 (0.722). Here we used "+" as "OR". In the simple "OR" model, we first determined the optimal cutoffs of AFP at 12.62 ng/mL, PIVKA-II at 28.23 mAU/mL, and AFP-L3% at 1.744 by maximizing the Youden's index for detection of HCC compared with CLD. Second, the levels of these three markers were dichotomized into 1 ($\geq$ cutoff) or 0 ($<$ cutoff). Finally, the simple "OR" models were constructed. As these three serum biomarkers combined with the simple "OR" models had relatively low performance, they were not analyzed in the subsequent subgroup analysis.

In subgroup analyses, HCC vs cirrhosis and HCC vs hepatitis, the GALAD-C, and GAAP models achieved
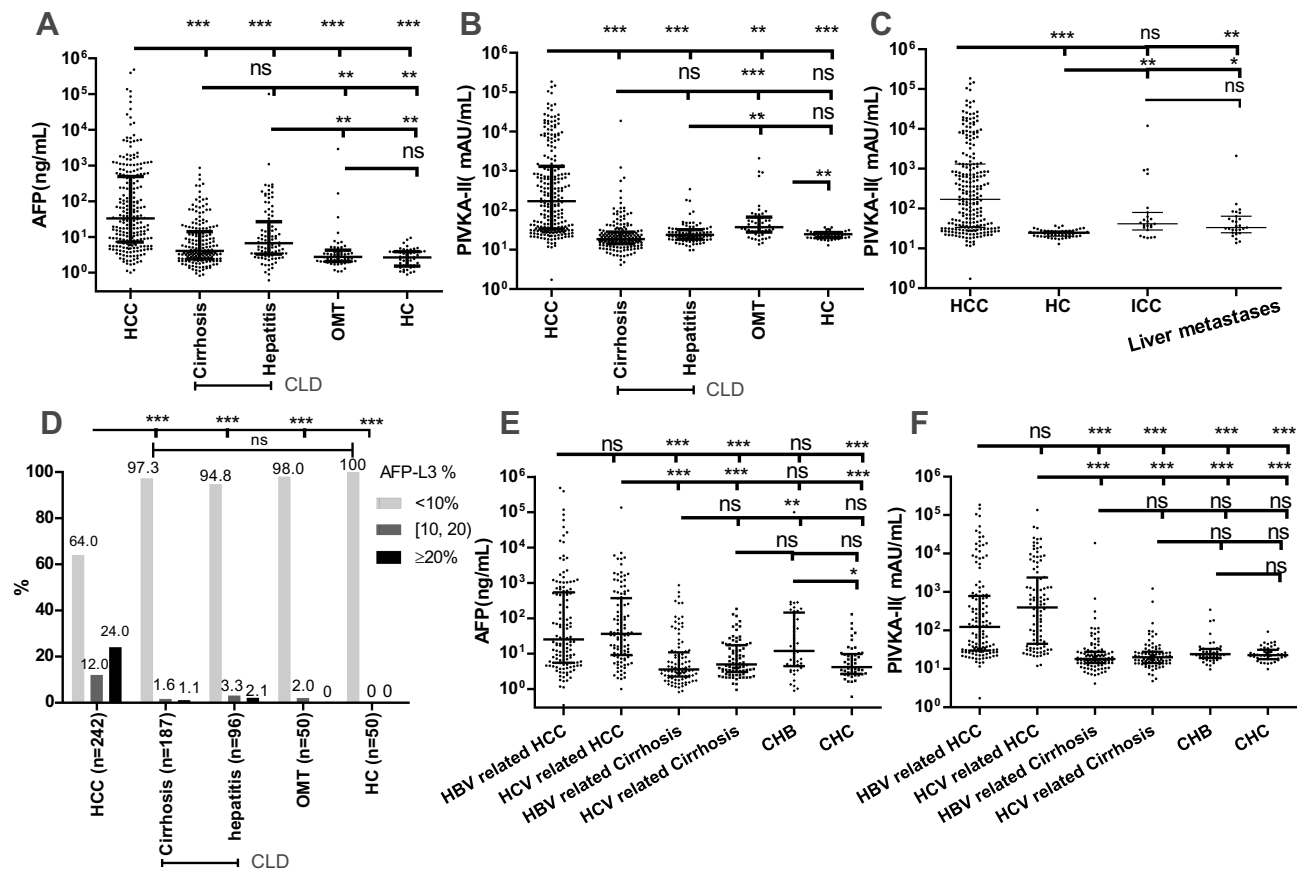
Figure 1 Serum AFP, PIVKA-II, and AFP-L3% in HCC and non-HCC groups. Comparison of AFP (**A**), PIVKA-II (**B**), and AFP-L3% (**D**) among HCC, Cirrhosis, Hepatitis, OMT, and HC groups. Comparison of PIVKA-II (**C**) among ICC, HCC, liver metastasis, and HCs. Comparison of AFP (**E**) and PIVKA-II (**F**) among HBV-related liver disease and HCV-related liver disease groups. The three horizontal bars in A, B, D, E, and F represent median with interquartile range values. For AFP and PIVKA-II, Kruskal–wallis *H*-tests were used for comparisons among groups; post hoc Dunn's Multiple Comparison tests were performed for pairwise comparisons. For AFP-L3%, Chi-square tests were performed. ***P <0.001, **P <0.01, *P <0.05, ns *P*>0.05.
Abbreviations: HCC, hepatocellular carcinoma; HC, healthy controls; CLD, chronic liver disease; OMT, other malignant liver tumors; HBV, hepatitis B virus; HCV, hepatitis C virus; CHB, chronic hepatitis B; CHC, chronic hepatitis C.

AUC values >0.90. These values were higher than those for the original GALAD model, PIVKA-II, AFP, and AFP-L3% results (Figure 2B and C, supplementary Table S3-4).

Discrimination of HCC Subgroups and CLD

The analyses of the HCC within Milan Criteria group vs the CLD group found that the GALAD-C and GAAP models had comparable AUC values (0.841 vs 0.840, respectively, *P* = 0.921). Both models had higher AUC values than the GALAD model and the individual biomarkers PIVKA-II, AFP, and AFP-L3 (0.810, 0.801, 0.679, and 0.602, respectively) (Figure 2D, supplementary Table S5).

The analyses of the HCC (maximum tumor diameter <5 cm) group vs the CLD group found that the GALAD-C and GAAP models had comparable AUC values (0.889 vs

0.882, respectively, *P* = 0.363). Both models had higher AUCs than the GALAD model and the individual biomarkers PIVKA-II, AFP, and AFP-L3% (Figure 2E, supplementary Table S6).

The analyses of the HCC (maximum tumor diameter ≥5 cm) vs the CLD groups revealed that the GALAD-C and GAAP models had comparable AUC values (0.983 vs 0.982, respectively, *P*= 0.1298). Both models had higher AUC values than the GALAD model and the individual biomarkers PIVKA-II, AFP, and AFP-L3% (Figure 2F, supplementary Table S7).

Discrimination of HCC from CLD, Stratified by Etiology

In those patients with an HCV etiology, all three models achieved high performances, with AUCs of 0.930, 0.958, and 0.954 for the GALAD, GALAD-C, and GAAP models,
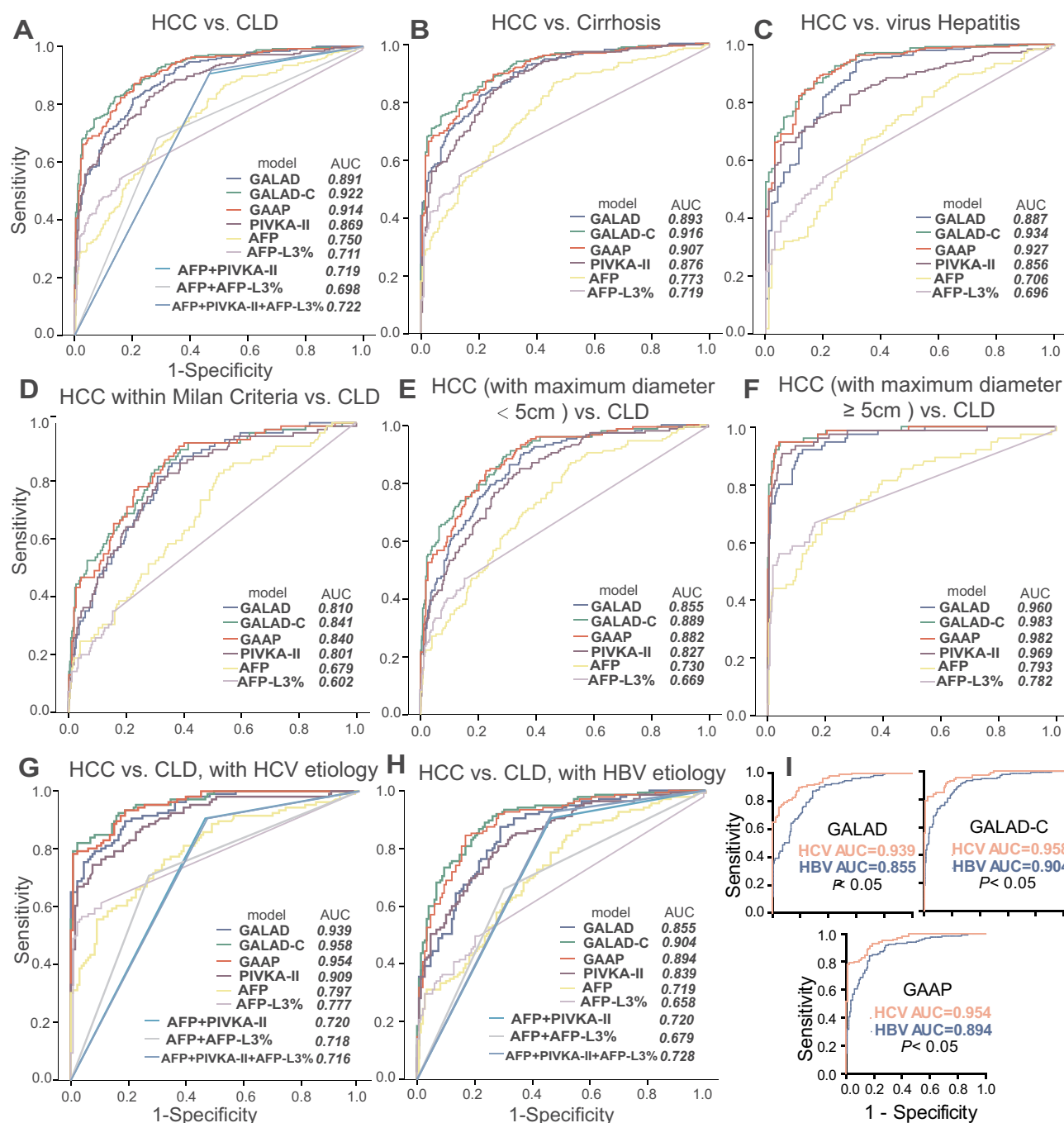
Figure 2 ROC curves comparing performance between GALAD, GALAD-C, GAAP, individual biomarkers, and combinations for discriminating HCC from CLD (**A**), HCC from Cirrhosis (**B**), HCC from Hepatitis (**C**), HCC within Milan Criteria from CLD (**D**), HCC (maximum diameter < 5 cm) from CLD (**E**), HCC (maximum diameter ≥ 5cm) from CLD (**F**), HCC from CLD, HCV etiology (**G**), HCC from CLD, HBV etiology (**H**), and comparisons between HCV and HBV etiology for the three models (**I**). HCC, hepatocellular carcinoma; CLD, chronic liver disease. In the marker combinations, "+" means "OR", and the cutoffs were 28.23 mAU/mL, 12.62 ng/mL, and 1.744% for PIVKA-II, AFP, and AFP-L3%, respectively.

respectively (Figure 2G). In patients with an HBV etiology, the corresponding AUC values were 0.855, 0.904, and 0.894, respectively, for the three models (Figure 2H) and the AUCs of the GALAD-C and GAAP models were significantly higher than that of the GALAD model. These three models had a significantly higher performance for HCV etiology compared with HBV etiology (Figure 2I). For the single-marker models, in both the HCV and HBV etiology subsets

Table 3 ROC Curve Analysis of Serum Biomarkers Alone and Combination, GALAD-C and GAAP for Discriminating HCC (n=242) and CLD (n=283)

Model/ Biomarker N=575	AUC (95% CI)	P value (GALAD-C vs Others)	Cut-Off Value	Sensitivity %	Specificity %	PPV %	NPV %	Correctly Classified %
GALAD	0.891 (0.864, 0.918)	0.0005	0.946	81.8	79.9	77.6	83.7	80.8
GALAD-C	0.922 (0.900, 0.945)	/	−0.374	82.6	85.9	83.3	85.3	84.4
GAAP	0.914 (0.889, 0.938)	0.0561	−0.650	87.2	79.2	78.1	87.8	82.9
PIVKA-II mAU/mL	0.869 (0.839, 0.900)	<0.0001	28.23	82.6	74.2	73.3	83.3	78.1
AFP ng/mL	0.750 (0.709, 0.792)	<0.0001	12.62	64.5	72.1	66.4	70.3	68.6
AFP-L3%	0.711 (0.666, 0.757)	<0.0001	1.744	54.1	84.5	74.9	68.3	70.5
AFP+PIVKA-II	0.719 (0.675, 0.763)	<0.0001	Same as above	90.5	53.4	62.4	86.8	70.5
AFP+AFP-L3%	0.698 (0.652, 0.743)	<0.0001	Same as above	68.2	71.4	67.1	72.4	69.9
AFP+PIVKA-II +AFP-L3%	0.722 (0.678, 0.766)	<0.0001	Same as above	91.7	52.7	62.4	88.2	70.7

Notes: "+" means "OR". GALAD-C vs GAAP, the P value is 0.0561. The construction of GALAD-C and GAAP models, and all the results obtained are based on all HCC patients (n=242) and all CLD patients (n=283) present in Table 1.
Abbreviations: AUC, area under receiver operating characteristic curve; HCC, hepatocellular carcinoma; CLD, chronic liver disease; PPV, positive predictive value; NPV, negative predictive value.

PIVKA-II achieved the highest AUC values (0.909 for HCV and 0.839 for HBV), followed by AFP (0.797 and 0.719) and AFP-L3% (0.777 and 0.658). All three single markers had higher AUC values for the HCV subset compared with the HBV subset. The results for sensitivity, specificity, positive predictive value, negative predictive value, and percent correctly classified are presented in supplementary Table S8.

Discrimination of HCC from CLD with Negative AFP

AFP-negative was defined as <20 ng/mL in this study, which is a commonly used cutoff value in studies and clinical practice. The GALAD, GALAD-C and GAAP models achieved similar performance in each of the three datasets, the entire subset (all HCC and CLD patients with negative AFP), the HCV subset, and the HBV subset; the range in AUC values was 0.882 to 0.924 (supplementary Figure S5). The results for sensitivity, specificity, positive predictive value, negative predictive value,

and percent correctly classified, using different cutoff points, are also presented in supplementary Figure S5. For example, the GAAP model had AUC values of 0.888 (0.851–0.924), 0.924 (0.883–0.965), and 0.892 (0.844–0.940) for the entire subset, the HCV subset, and the HBV subset, respectively. The corresponding values for sensitivity for HCC diagnosis were 79.4%, 88.1%, and 87.5%, respectively, when keeping specificity at 80%. As expected, lower values for sensitivity (64.5%, 60.3%, and 70.3%, respectively) were obtained when keeping specificity at 90%. Using the cutoff point maximizing the sum of sensitivity and specificity, the GAAP model achieved sensitivity values of 92.5%, 88.1%, and 87.5%, respectively, and corresponding specificity values of 70.9%, 83.8%, and 81.1%, respectively. Performance of each of the three models was higher than that of the individual markers. PIVKA-II had the best performance among the single-marker models; the values for sensitivity were 67.3%, 69%, and 65.6%, respectively, at a specificity of 80%, and were 50.5%, 59.5%, and 45.3% at a specificity of 90%.

Discrimination Between HCC and Other Malignant Liver Tumor

The GALAD-C and GAAP models had comparable AUC values (0.823 vs 0.809, respectively, $P = 0.067$). These AUCs were higher than for the PIVKA-II (0.704) and AFP-L3% (0.747) results, but lower than the original GALAD (0.873) and AFP (0.873) results (supplementary Figure S6A). We performed separate subgroup analyses of ICC and liver metastases (supplementary Figure S6B, S6C). For differentiating HCC from ICC, AFP had the highest and most acceptable performance (AUC=0.899), with a high specificity of 91.3% and a moderate sensitivity of 64.5% at a cutoff value of 12.62 ng/mL. For differentiating HCC from liver metastases, the original GALAD model had the highest AUC value (0.876), but it was only slightly higher than the AFP result (AUC=0.850). The use of AFP resulted in a high specificity of 92.6% and a moderate sensitivity of 64.5% at a cutoff value of 12.62 ng/mL. This result indicated these models were not suitable for use, while the use of the single marker AFP > 12.62 had a high specificity to exclude the possibility of ICC or liver metastases.

Discrimination Between HCC and Healthy Control Group

The GALAD, GALAD-C, and GAAP models had comparable AUC values (0.988 vs 0.982 vs 0.979, respectively, pairwise P values <0.05). They all had AUC values greater than the values for the individual biomarkers PIVKA-II, AFP, and AFP-L3% (supplementary Figure S6D).

AFP, PIVKA-II, GALAD-C, and GAAP with Maximum Tumor Diameter and Portal Vein Invasion

PIVKA-II level, the GALAD-C score, and the GAAP score were significantly positively correlated with the maximum diameters of HCC tumors (r=0.6453, r=0.6133, r=0.6332, respectively), while the AFP level was weakly correlated with that (r=0.1815) (supplementary Figure S7).

The median values of AFP, PIVKA-II, GALAD-C score and GAAP score in the HCC group with portal vein invasion were significantly higher than those in the HCC group without portal vein invasion (all P values<0.001) (supplementary Figure S8).

Diagnostic Effectiveness of GAAP Model in Test Dataset

The GAAP model achieved an AUC of 0.924 (95% CI, 0.895–0.952), and the values for sensitivity, specificity, and percent correctly classified were 88%, 80%, and 84%, respectively, at a cutoff of ≥ -0.650 (determined in the training dataset), when discriminating of HCC from CLD in the test dataset (supplementary Table S9).

Discussion

We validated the use of the GALAD model for discriminating HCC from chronic hepatitis liver disease with the HCV etiology in China, although the performance in the entire dataset and the subgroup with the HBV etiology was significantly lower than for previous studies. By using the original GALAD formula to discriminate HCC from chronic liver disease, we obtained a significantly lower AUC (0.891) than that in the initial study based on a UK cohort (0.97),[12] and in other studies (AUC values up to 0.93 to 0.98).[13–15,23] The methods we used to test the three markers were different. The units for PIVKA-II were also different, and the two units could not be interchanged. We refitted the model using the same five variables and the same regression method (logistic regression) based on our dataset. As expected, we obtained different coefficients and named it GALAD-C. The coefficients changed substantially, especially the coefficient for AFP. It was approximately two-fold of that for PIVKA-II (2.34 vs 1.33) in the original GALAD model, while it became about one-third of the coefficient for PIVKA-II (0.840 vs 2.364) in our dataset. This change indicated that the relative contributions of the two markers changed. Although it did not result in a value as high as in the original study, the GALAD-C model obtained an excellent, and higher, AUC value (0.922) compared with the GALAD model (0.891). This may be related to the ethnic factor and characteristics of population. A previous study showed that the GALAD model achieved the highest AUC values in the UK (AUC=0.97), followed by Germany (0.94) and Japan (0.93) using the same test method.[13] Our result based on the GALAD-C model was most similar to that of Japan. We constructed the GAAP model by excluding AFP-L3, because measurement of this marker is complex, time-consuming, expensive, and requires up to a 400-uL serum sample, and the contribution of AFP-L3 was low in our study. The GAAP model performed comparably to the GALAD-C model.

In our study, the performance of the GALAD model for an HCV etiology (AUC, 0.939) was high and was similar to that

of previous studies with AUC values ranging from 0.92 to 0.98.[13,14] Higher AUCs were obtained using the GALAD-C (0.958) and GAAP (0.954) models. However, the performance of these three models for the HBV subset (AUC, 0.855–0.904) was significantly lower than the performance for the HCV subset (AUC, 0.939–0.958). It was also lower than in previous studies (AUCs ranging from 0.93 to 0.99,[13] even to 1.00 in a study using small sample sizes (11 HCC vs 27 CLD)).[14] That 55.8% of HCC patients were with an HBV etiology results in the lower performances of these three models in the entire dataset of our study. One possible reason the models had a low performance for the HBV subgroup was that AFP was elevated in the patients with chronic HBV, and AFP elevation can be associated with ALT elevation. There was a weak correlation between AFP and ALT values in our study and a moderate correlation in a previous study.[24] A previous study found that the GALAD model exhibited very similar performance for HCV vs HBV subgroups, with AUC values of 0.98 vs 0.99, respectively, in the UK, 0.93 vs 0.94, respectively, in Germany, and 0.92 vs 0.93, respectively, in Japan.[13] But, because ALT values were not available, we did not determine whether the values for this group of patients with HBV were relatively low. Another possible reason for the low model performance was that the patients with HBV HCC had higher percentages of unifocal small tumors, compared with the patients with HCV HCC (supplementary Table S10). The performance of these models and individual markers apparently decreased with the decrease of tumor size, independent of whether the entire dataset or the HBV or HCV subsets were analyzed (supplementary Table S11). This result was not consistent with the results of Sarah Berhane's study, which found that the performance of the GALAD model was almost consistent, from a unifocal tumor size <5 cm down to <2 cm, although there are differences among different countries (0.93 to 0.92 for UK, 0.91 to 0.89 for Japan, 0.85 to 0.87 for Germany [only to 3 cm, since only seven HCC < 2 cm and its AUC value increases to 0.93]).[13]

We found that model performance for patients with early-stage HCC (Milan Criteria) was only moderate (AUC value of 0.84 for GALAD-C and GAAP models), which was lower than for Japan and UK early-stage HCC (Milan Criteria) (0.91 and 0.93, respectively, GALAD model),[13] and for a cohort in Germany (0.92) and a Mayo Clinic cohort with early-stage defined as Barcelona clinic liver cancer 0/A stage (0.92).[14,25] The analysis by etiology found that performance was better for an HCV etiology than for an HBV etiology (AUC, 0.895 vs.0.832 for both GALAD and GAAP models, supplementary Table S10).

In our study, among the individual markers, PIVKA-II achieved the best performance for discrimination of HCC from CLD. Study results are inconsistent in terms of which marker is better; some findings support the PIVKA-II,[26–29] and some support the AFP.[30–32] But one thing was unanimous that the combination of two biomarkers is superior to the use of a single biomarker.[24,32–35] Our study found there was a weak correlation between the two markers. Among HCC patients with negative AFP (defined as <20 ng/mL, 44.2% (107/242)), 57.9% were PIVKA-II positive (defined as ≥40 mAU/mL); among PIVKA-II-negative HCC patients (30.2% (73/242)), 38.4% were AFP positive.

That combination of AFP and PIVKA-II increased performance was supported by the results of the analysis of models/individual markers in patients with negative AFP. Three models achieved similar and acceptable performances in patients with negative AFP; AUCs ranged from 0.882 to 0.924. The GAAP model achieved sensitivities of 79.4% and 64.5% at a specificity of 80% and 90%, respectively. The GALAD-C and GAAP models achieved sensitivities of 67.3% and 75.7% at cutoffs of −0.374 and −0.65, respectively (supplementary Figure S5). This result was similar to a previous study obtaining a sensitivity of 67.3% at a GALAD score cutoff of −0.65.[25] Performance of these three models was significantly higher than that of individual markers. Among single markers, PIVKA-II had the best performance (AUC, 0.834), achieving sensitivities of 67.3% and 50.5% at specificity values of 80% and 90%, respectively. This sensitivity of PIVKA-II was consistent with previous study findings of sensitivities of 51.02–56.5% at specificities of 84%–90%.[36,37]

This study found that PIVKA-II was elevated in patients with ICC. The GALAD, GALAD-C, and GAAP models including PIVKA-II as a component achieved relatively lower AUC values (0.87, 0.791, and 0.78, respectively) for discrimination of patients with HCC from those with ICC. These values were significantly lower than those found in a previous study (AUC, 0.95)[13] that included 13 ICC patients. That study and our study enrolled in very small numbers of patients. No other studies comparing PIVKA-II/DCP levels between patients with HCC and ICC were found. Another study showed that DCP had no efficiency (AUC, 0.512) for differentiation of patients with ICC from those without cancer; it had moderate performance (AUC, 0.846) for differentiation of patients with HCC from patients without cancer.[38] This result indirectly suggests that DCP has diagnostic value for differential diagnosis of HCC and ICC. However, this study did not give the number of patients included in the study population. More studies with greater sample size numbers

are needed to validate the performance of PIVKA-II for differentiation of HCC from ICC. We found the median level of conjugated bilirubin was significantly higher in the ICC (74.65 μmol/L) than the HCC (7.20 μmol/L, $P<0.05$), and other groups (supplementary Figure S9). The possible mechanism of PIVKA-II elevation in patients with ICC could be similar to obstructive jaundice. Studies found that PIVKA-II increases in non-HCC patients with obstructive jaundice, alcoholic liver disease, warfarin use, and antibiotic use.[39–44] Elevation of PIVKA-II in patients with obstructive jaundice is associated with vitamin K deficiency.[40] Obstructive jaundice can affect bile excretion, which decreases the absorption of fat-soluble vitamin K. The resulting vitamin K deficiency leads to PIVKA-II elevation. Theoretically, obstructive jaundice and cholestasis were present in the patients with ICC, which led to PIVKA-II elevation.

This study had some limitations. It was based on a retrospective design and the data were obtained from a single medical center.

Conclusions

In conclusion, GALAD was excellent for discrimination of HCC from CLD in the HCV subgroup of a cohort of Chinese patients. The GAAP model we constructed used only four variables had performance comparable to the GALAD-C model we constructed by optimizing the parameters of the GALAD. The GAAP and GALAD-C models achieved better performance compared with GALAD. These three models exhibited better performance in patients with an HCV etiology than those with HBV. AFP alone was the best marker for differentiating HCC from a group of other tumors consisting of ICC and liver metastases. The GALAD, GALAD-C, and GAAP models had excellent performance for differentiation of the HCC from the healthy controls. More studies are needed to further confirm the effectiveness of the GALAD and GAAP models for use in populations in China.

Abbreviations

AFP, α-fetoprotein; AFP-L3, *Lens culinaris* agglutinin-reactive fraction of AFP; AUC, area under the receiver operating characteristic curve; CI, confidence interval; CLD, chronic liver disease; DCP, des-g carboxyprothrombin; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; PBC, primary biliary cholangitis; ROC, receiver operating characteristic; ICC, intrahepatic cholangiocarcinoma; PIVKA-Ⅱ, protein induced by vitamin K absence or antagonist-II.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87–108.
3. Yang JD, Roberts LR. Hepatocellular carcinoma: a global view. *Nat Rev Gastroenterol Hepatol*. 2010;7(8):448–458. doi:10.1038/nrgastro.2010.100
4. 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(4):317–370.
5. Singal AG, Conjeevaram HS, Volk ML, et al. Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. *Cancer Epidemiol Biomarkers Prev*. 2012;21(5):793–799. doi:10.1158/1055-9965.EPI-11-1005
6. Zhao S, Long M, Zhang X, et al. The diagnostic value of the combination of Golgi protein 73, glypican-3 and alpha-fetoprotein in hepatocellular carcinoma: a diagnostic meta-analysis. *Ann Transl Med*. 2020;8(8):536. doi:10.21037/atm.2020.02.89
7. Oda K, Ido A, Tamai T, et al. Highly sensitive lens culinaris agglutinin-reactive alpha-fetoprotein is useful for early detection of hepatocellular carcinoma in patients with chronic liver disease. *Oncol Rep*. 2011;26(5):1227–1233.
8. Kokudo N, Takemura N, Hasegawa K, et al. Clinical practice guidelines for hepatocellular carcinoma: the Japan society of hepatology 2017 (4th JSH-HCC guidelines) 2019 update. *Hepatol Res*. 2019;49(10):1109–1113. doi:10.1111/hepr.13411
9. Inagaki Y, Tang W, Makuuchi M, Hasegawa K, Sugawara Y, Kokudo N. Clinical and molecular insights into the hepatocellular carcinoma tumour marker des-gamma-carboxyprothrombin. *Liver Int*. 2011;31(1):22–35. doi:10.1111/j.1478-3231.2010.02348.x

10. Saitta C, Raffa G, Alibrandi A, et al. PIVKA-II is a useful tool for diagnostic characterization of ultrasound-detected liver nodules in cirrhotic patients. *Medicine*. 2017;96(26):e7266. doi:10.1097/MD.0000000000007266

11. Huang S, Jiang F, Wang Y, et al. Diagnostic performance of tumor markers AFP and PIVKA-II in Chinese hepatocellular carcinoma patients. *Tumour Biol*. 2017;39(6):1010428317705763. doi:10.1177/1010428317705763

12. Johnson PJ, Pirrie SJ, Cox TF, et al. The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers. *Cancer Epidemiol Biomarkers Prev*. 2014;23(1):144–153. doi:10.1158/1055-9965.EPI-13-0870

13. Berhane S, Toyoda H, Tada T, 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):875–886e876. doi:10.1016/j.cgh.2015.12.042

14. Yang JD, Addissie BD, Mara KC, et al. GALAD score for hepatocellular carcinoma detection in comparison with liver ultrasound and proposal of GALADUS score. *Cancer Epidemiol Biomarkers Prev*. 2019;28(3):531–538. doi:10.1158/1055-9965.EPI-18-0281

15. Best J, Bechmann LP, Sowa JP, et al. GALAD score detects early hepatocellular carcinoma in an international cohort of patients with nonalcoholic steatohepatitis. *Clin Gastroenterol Hepatol*. 2019.

16. Heimbach JK, Kulik LM, Finn RS, et al. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology*. 2018;67(1):358–380. doi:10.1002/hep.29086

17. Mazzaferro V, Regalia E, Doci R, et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med*. 1996;334(11):693–699. doi:10.1056/NEJM199603143341104

18. Chinese Society of Hepatology CMA, Chinese Society of Infectious Diseases CMA, Hou JL, lai W. The guideline of prevention and treatment for chronic hepatitis B: a 2015 update. *Zhonghua Gan Zang Bing Za Zhi*. 2015;23(12):888–905.

19. Chinese Society of Hepatology CMA, Wei L, Hou JL. The guideline of prevention and treatment for hepatitis C: a 2015 update. *Zhonghua Gan Zang Bing Za Zhi*. 2015;23(12):906–923.

20. Pugh RN, Murray-Lyon IM, Dawson JL, Pietroni MC, Williams R. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg*. 1973;60(8):646–649. doi:10.1002/bjs.1800600817

21. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32–35. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3

22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845. doi:10.2307/2531595

23. Caviglia GP, Abate ML, Petrini E, Gaia S, Rizzetto M, Smedile A. Highly sensitive alpha-fetoprotein, Lens culinaris agglutinin-reactive fraction of alpha-fetoprotein and des-gamma-carboxyprothrombin for hepatocellular carcinoma detection. *Hepatol Res*. 2016;46(3):E130–135. doi:10.1111/hepr.12544

24. Ricco G, Cavallone D, Cosma C, et al. Impact of etiology of chronic liver disease on hepatocellular carcinoma biomarkers. *Cancer Biomark*. 2018;21(3):603–612. doi:10.3233/CBM-170551

25. Best J, Bilgi H, Heider D, et al. The GALAD scoring algorithm based on AFP, AFP-L3, and DCP significantly improves detection of BCLC early stage hepatocellular carcinoma. *Z Gastroenterol*. 2016;54(12):1296–1305. doi:10.1055/s-0042-119529

26. Yu R, Tan Z, Xiang X, Dan Y, Deng G. Effectiveness of PIVKA-II in the detection of hepatocellular carcinoma based on real-world clinical data. *BMC Cancer*. 2017;17(1):608. doi:10.1186/s12885-017-3609-6

27. Seo SI, Kim HS, Kim WJ, et al. Diagnostic value of PIVKA-II and alpha-fetoprotein in hepatitis B virus-associated hepatocellular carcinoma. *World J Gastroenterol*. 2015;21(13):3928–3935. doi:10.3748/wjg.v21.i13.3928

28. Ji J, Wang H, Li Y, et al. Diagnostic evaluation of des-gamma-carboxy prothrombin versus alpha-fetoprotein for hepatitis B virus-related hepatocellular carcinoma in china: a large-scale, multicentre study. *PLoS One*. 2016;11(4):e0153227. doi:10.1371/journal.pone.0153227

29. Pote N, Cauchy F, Albuquerque M, et al. Performance of PIVKA-II for early hepatocellular carcinoma diagnosis and prediction of microvascular invasion. *J Hepatol*. 2015;62(4):848–854. doi:10.1016/j.jhep.2014.11.005

30. Choi J, Kim GA, Han S, Lee W, Chun S, Lim YS. Longitudinal assessment of three serum biomarkers to detect very early-stage hepatocellular carcinoma. *Hepatology*. 2019;69(5):1983–1994. doi:10.1002/hep.30233

31. Wu M, Liu Z, Li X, Zhang A, Li N. Dynamic changes in serum markers and their utility in the early diagnosis of all stages of hepatitis B-associated hepatocellular carcinoma. *Onco Targets Ther*. 2020;13:827–840. doi:10.2147/OTT.S229835

32. Yu R, Xiang X, Tan Z, Zhou Y, Wang H, Deng G. Efficacy of PIVKA-II in prediction and early detection of hepatocellular carcinoma: a nested case-control study in Chinese patients. *Sci Rep*. 2016;6(1):35050. doi:10.1038/srep35050

33. Ertle JM, Heider D, Wichert M, et al. A combination of alpha-fetoprotein and des-gamma-carboxy prothrombin is superior in detection of hepatocellular carcinoma. *Digestion*. 2013;87(2):121–131. doi:10.1159/000346080

34. Caviglia GP, Ribaldone DG, Abate ML, et al. Performance of protein induced by vitamin K absence or antagonist-II assessed by chemiluminescence enzyme immunoassay for hepatocellular carcinoma detection: a meta-analysis. *Scand J Gastroenterol*. 2018;53(6):734–740. doi:10.1080/00365521.2018.1459824

35. Fu J, Li Y, Li Z, Li N. Clinical utility of decarboxylation prothrombin combined with alpha-fetoprotein for diagnosing primary hepatocellular carcinoma. *Biosci Rep*. 2018;38(5). doi:10.1042/BSR20180044

36. Chen H, Zhang Y, Li S, et al. Direct comparison of five serum biomarkers in early diagnosis of hepatocellular carcinoma. *Cancer Manag Res*. 2018;10:1947–1958. doi:10.2147/CMAR.S167036

37. Wang X, Zhang W, Liu Y, et al. Diagnostic value of prothrombin induced by the absence of vitamin K or antagonist-II (PIVKA-II) for early stage HBV related hepatocellular carcinoma. *Infect Agent Cancer*. 2017;12(1):47. doi:10.1186/s13027-017-0153-6

38. Li Y, Chen J. Serum des-gamma-carboxy prothrombin for diagnosis of adult primary cancer in liver. *J Coll Physicians Surg Pak*. 2019;29(10):972–976. doi:10.29271/jcpsp.2019.10.972

39. Kang KH, Kim JH, Kang SH, et al. The influence of alcoholic liver disease on serum PIVKA-II levels in patients without hepatocellular carcinoma. *Gut Liver*. 2015;9(2):224–230. doi:10.5009/gnl14047

40. Tanaka T, Taniguchi T, Sannomiya K, et al. Novel des-gamma-carboxy prothrombin in serum for the diagnosis of hepatocellular carcinoma. *J Gastroenterol Hepatol*. 2013;28(8):1348–1355. doi:10.1111/jgh.12166

41. Nakao A, Suzuki Y, Isshiki K, et al. Clinical evaluation of plasma abnormal prothrombin (des-gamma-carboxy prothrombin) in hepatobiliary malignancies and other diseases. *Am J Gastroenterol*. 1991;86(1):62–66.

42. Oka T, Touchi A, Harauchi T, Takano K, Yoshizaki T, Matsubara T. In vivo effects of beta-lactam antibiotics and heterocyclic thiol compounds on vitamin K-dependent carboxylation activity and blood coagulation factors in vitamin K-deficient rats. *Biochem Pharmacol*. 1988;37(10):2091–2095. doi:10.1016/0006-2952(88)90561-8

43. Sakizono K, Oita T, Eto M, Bito S, Takegawa H, Kasakura S. Studies on the mechanism of elevation of serum PIVKA-II levels in alcoholic liver cirrhosis. *Rinsho Byori*. 2002;50(3):289–295.

44. Ohhira M, Ohtake T, Saito H, et al. Increase of serum des-gamma-carboxy prothrombin in alcoholic liver disease without hepatocellular carcinoma. *Alcohol Clin Exp Res*. 1999;23(4 Suppl):67S–70S. doi:10.1111/j.1530-0277.1999.tb04537.x