Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Diagnostic Performance Comparison of AFP, PIVKA-II, GALAD Model, and ASAP Model Across Two Chemiluminescence Immunoassay Platforms for Hepatocellular Carcinoma

Yuan Huang[1,*], Rui Ding[1,*], Yue Cui[2,*], Peng Li[1], Jie Niu[3], Guan-Hua Wang[1], Xu-Zhen Qin[1]

[1]Department of Clinical Laboratory, Chinese Academy of Medical Sciences & Peking Union Medical College Hospital, Beijing, 100730, People's Republic of China; [2]Department of Clinical Laboratory, Chongqing Traditional Chinese Medicine Hospital, Chongqing, 400021, People's Republic of China; [3]Department of Clinical Laboratory, Capital Center for Children's Health, Capital Medical University, Beijing, 100020, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xu-Zhen Qin, Email qxz_01@163.com

Objective: This study aimed to evaluate the diagnostic performance of individual serum biomarkers [alpha-fetoprotein (AFP), protein induced by vitamin K absence or antagonist-II (PIVKA-II)] and composite models (GALAD, ASAP) for hepatocellular carcinoma (HCC) across two immunoassay platforms.

Methods: From 2011 to 2021, 518 serum samples were selected from a liver-related disease biobank at Peking Union Medical College Hospital (Beijing, China), including 102 HCC patients, 117 with benign liver disease, 38 with cholangiocarcinoma, 96 with colorectal cancer, 65 with metastatic hepatic carcinoma, and 100 healthy controls. AFP and PIVKA-II levels were measured on both the Hotgen and Abbott ARCHITECT platforms. The GALAD and ASAP scores were calculated based on the data from each platform. Receiver operating characteristic (ROC) curve analysis and the corresponding areas under the curves (AUCs) were used to evaluate and compare the diagnostic value of the individual biomarkers and the two composite models.

Results: For HCC diagnosis, AFP exhibited comparable efficacy between Hotgen (AUC: 0.821) and Abbott (AUC: 0.846), whereas PIVKA-II performed better on Abbott (AUC: 0.863) than Hotgen (AUC: 0.787). GALAD and ASAP models exhibited significantly better diagnostic performance than individual serum biomarkers on both platforms ($P < 0.05$): on Hotgen, both models achieved an AUC of 0.872, while on Abbott, ASAP (AUC: 0.913) was marginally superior to GALAD (AUC: 0.901, $P = 0.0569$). Notably, both models performed better on Abbott than Hotgen (GALAD: 0.901 vs 0.872, $P = 0.0001$; ASAP: 0.913 vs 0.872, $P = 0.0003$). Spearman correlation analysis showed moderate inter-platform correlations for AFP ($r = 0.573$) and PIVKA-II ($r = 0.460$). Bland-Altman analysis indicated poor inter-platform consistency, with mean biases of 44.32% (AFP) and −92.02% (PIVKA-II).

Conclusion: GALAD and ASAP models demonstrate superior diagnostic efficacy for HCC compared to individual biomarkers, and their performance is significantly influenced by the immunoassay platform employed.

Keywords: hepatocellular carcinoma, AFP, PIVKA-II, GALAD, ASAP, hotgen biotech, Abbott Architect

Introduction

In 2020, approximately 905,677 new cases of liver cancer were diagnosed globally, resulting in over 830,000 deaths.[1] In China, liver cancer is the second leading cause of cancer-related mortality.[2] Hepatocellular carcinoma (HCC), which accounts for more than 80% of primary liver cancer cases, poses a conundrum of rapid progression, early metastasis, and a lack of early-stage symptoms.[3] The overall five-year survival rate for HCC is only 18%;[4] however, this increases to over 70% for patients diagnosed and treated at an early stage.[5] Therefore, early diagnosis and effective monitoring of HCC progression are crucial for improving patient outcomes and reducing mortality.

https://doi.org/10.2147/JHC.S554305
Received: 19 August 2025
Accepted: 6 December 2025
Published: 8 January 2026

Serum markers such as alpha-fetoprotein (AFP) and protein induced by vitamin K absence or antagonist-II (PIVKA-II) are widely employed for diagnosing and monitoring HCC.[6] These biomarkers provide non-invasive options for detecting HCC, but their individual diagnostic performance remains limited: meta-analyses have reported moderate sensitivity (61% for AFP, 66% for PIVKA-II) and specificity (87% for AFP, 88% for PIVKA-II).[7–9]

To mitigate the limitations of inadequate performance of individual serum biomarkers, composite diagnostic models integrating clinical parameters have been developed. A prominent example is the GALAD model, first introduced in 2014.[10] This model, constructed using the logistic regression method, incorporates five parameters (gender, age, AFP-L3, AFP, and PIVKA-II) and was initially established using the mTASWako i30 autoanalyzer. Since its development, the GALAD model has been extensively validated across diverse populations and etiologies, consistently demonstrating high sensitivity and specificity.[10–14] Another representative model is the ASAP model, proposed in 2019.[15] Also built based on the logistic regression method, ASAP differs from GALAD in that it excludes AFP-L3 and relies on four parameters (gender, age, AFP, and PIVKA-II); it was originally developed and tested on the ARCHITECT immunoassay platform. Subsequent external validations have further confirmed the ASAP model's robust diagnostic utility.[15–18]

Notably, recent studies have primarily focused on comparing diagnostic performance for HCC: between these models, between the models and individual serum biomarkers, and across different models in etiologically distinct populations.[19–22] However, no systematic studies have investigated whether the performance of these models varies when applied to biomarker data generated by different analytical platforms. Importantly, these distinct detection platforms differ in key technical parameters, including reagent formulations, detection thresholds, signal amplification mechanisms, and instrument calibration protocols. Such platform-related differences may introduce quantitative biases in biomarker measurements, which can alter model outputs and further affect diagnostic decisions. The primary objective of this study is to compare the clinical performance of AFP, PIVKA-II, the GALAD model, and the ASAP model for HCC diagnosis across two distinct detection platforms.

Methods

Participants

A total of 518 participants were enrolled from Peking Union Medical College Hospital (Beijing, China) between 2011 and 2021. The cohort comprised 102 HCC patients, 117 with benign liver disease (BLD), 38 with cholangiocarcinoma (CCA), 96 with colorectal cancer (CRC), 65 with metastatic hepatic carcinoma (MHC), and 100 healthy controls. HCC was defined according to the American Association for the Study of Liver Diseases (AASLD) Practice Guidelines for the management of HCC (updated version, 2010).[23] BLD cases included patients with hepatic cysts or hepatic hemangiomas confirmed by ultrasound or computed tomography. CCA, CRC, and MHC diagnoses were pathologically confirmed. Demographic and clinical data from electronic medical records were retrieved, including patient gender, age, and alanine aminotransferase (ALT) levels.

Measurement of Serum AFP, PIVKA-II and AFP-L3

Three to five milliliters of venous blood were collected from each participant. Samples were centrifuged at 1,500 ×g for 10 min, and the obtained serum was aliquoted and stored at −80°C until analysis. Frozen samples were thawed at room temperature prior to testing. Serum AFP and PIVKA-II levels were measured using the Abbott ARCHITECT i2000 automated analyzer (Illinois, USA) and the Hotgen Biotech C2000 automated analyzer (Beijing, China). Serum AFP-L3 was measured using the Hotgen Biotech C2000 automated analyzer. The detection methods employed on both the Abbott and Hotgen platforms are chemiluminescence microparticle immunoassays. All tests were performed as per the manufacturers' specifications. Results from the Hotgen Biotech C2000 were designated as AFP_Hotgen, PIVKA-II _Hotgen, and AFP-L3_Hotgen; Results from the Abbott ARCHITECT i2000 were indicated as AFP_Abbott and PIVKA-II_Abbott. PIVKA-II_Hotgen was expressed in ng/mL, whereas PIVKA-II_Abbott was reported in mAU/mL (1 mAU/mL = 1 ng/mL per manufacturer's specifications).

Definition of GALAD and ASAP Models

GALAD and ASAP models were calculated according to previous studies to evaluate their diagnostic value for HCC.[10,15] GALAD score= −10.08 + 0.09 × [Age] + 1.67 × [Gender (1 for males, 0 for females)] + 0.04 × [AFP-L3%] + 2.34 × $\log_{10}$ [AFP] + 1.33 × $\log_{10}$ [PIVKA-II]. ASAP score= −7.57711770 + 0.04666357 × Age - 0.57611693 × [Gender (0 for males, 1 for females)] + 0.42243533 × ln [AFP] + 1.10518910 × ln[PIVKA-II].

Statistical Analysis

For the evaluation of HCC diagnostic performance and the calculation of GALAD and ASAP scores, all specimens were included. Values exceeding the upper limit of detection for the Abbott platform (>2000 ng/mL for AFP and >30000 mAU/mL for PIVKA-II) were censored at these respective thresholds. Subsequently, for the inter-platform comparison, specimens with analyte levels exceeding the Abbott system's upper limits (13 for AFP and 3 for PIVKA-II) were excluded, yielding 505 and 515 valid paired measurements for AFP and PIVKA-II, respectively.

Statistical analyses were performed using SPSS 18.0, GraphPad Prism 10.0, Origin 2021, and MedCalc for data analyses and graphical representations. The Kolmogorov–Smirnov test was employed to assess the distribution of variables. The Mann–Whitney *U*-test and Kruskal–Wallis *H*-test were utilized to compare differences between groups. Receiver operating characteristic (ROC) curves and area under the curve (AUC) with 95% confidence intervals (CIs) were conducted to elucidate the diagnostic values of AFP, PIVKA-II, AFP-L3 and composite models. DeLong's test was used for AUC comparisons. Parameters related to diagnostic efficiency, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated.

Spearman correlation analysis was utilized to evaluate the correlation between the two platforms. Passing-Bablok regression was applied to assess systematic and proportional bias. Bland-Altman plots were used to determine the consistency between two variables. P values < 0.05 were regarded as statistically significant.

Results

Clinical Characteristics of Participants

In this study, the clinical characteristics of 518 enrolled participants, including age, gender, serum ALT levels, and concentrations of serum AFP, PIVKA-II, and AFP-L3, are presented in Table 1. In the HCC group, the mean age of patients was 56.91 years, with 75.5% of male patients, and 55.9% of patients with elevated ALT levels (>40 IU/L). AFP, PIVKA-IIand AFP-L3 levels in the HCC group, measured on both platforms, were significantly elevated compared to all other groups (P < 0.001, Figure 1). Inter-group comparisons were performed using Mann–Whitney *U*-test and Kruskal–Wallis H-test for non-parametric data. Figure 1 visually compares the distributions of AFP, PIVKA-II, and AFP-L3 across all groups.

Diagnostic Performance of AFP, PIVKA-II, AFP-L3, GALAD Model, and ASAP Model Determined by Two Platforms

ROC analysis was used to evaluate the diagnostic performance of AFP, PIVKA-II, AFP-L3, GALAD model, and ASAP model for HCC, with results shown in Table 2. Notably, the table provides sensitivity, specificity, PPV, and NPV corresponding to both ROC-derived optimal cut-off values and commonly used clinical cut-offs,[24] enabling clinicians to better evaluate and apply these biomarkers in practice. DeLong's test was used for AUC comparisons. The AUC of AFP_Hotgen was 0.821 (95% CI: 0.786–0.853), and that of AFP_Abbott was 0.846 (95% CI: 0.811–0.876). The diagnostic efficiency of Hotgen Biotech was similar to Abbott Architect (P = 0.0873). For PIVKA-II, AUC comparison revealed that the diagnostic efficiency of the Abbott platform was mildly higher than that of the Hotgen platform (P = 0.0017). The AUC of AFP-L3_Hotgen was 0.714 (95% CI: 0.673–0.752).

The GALAD_Hotgen model and the ASAP_Hotgen model exhibited significantly higher AUC values for HCC diagnosis compared to the individual serum markers AFP_Hotgen, PIVKA-II_Hotgen, and AFP-L3_Hotgen (P < 0.05). Meanwhile, the GALAD_Abbott model and ASAP_Abbott model also demonstrated a notable increase in AUC values for HCC diagnosis compared to the individual serum markers AFP_Abbott and PIVKA-II_Abbott (P < 0.05).

Table 1 Clinical Features of Participants Enrolled in This Study

	HCC (n=102)	CCA (n=38)	CRC (n=96)	MHC (n=65)	BLD (n=117)	Controls (n=100)	*P* Value
Male/Female (%)[a]	77/25 (75.5/24.5)	25/13 (65.8/34.2)	60/36 (62.5/37.5)	25/40 (38.5/61.5)	61/56 (52.1/47.9)	31/69 (31.0/69.0)	<0.001
Age (y)[b]	56.91 ± 12.67	57.89 ± 11.78	65.08 ± 10.38	56.17 ± 10.12	44.06 ± 13.91	36.11 ± 11.21	<0.001
ALT, IU/L[c]	57.50 (24.75, 171.75)	167.00 (29.50, 320.50)	13.00 (9.25, 20.00)	32.00 (16.00, 174.00)	32.00 (20.00, 84.50)	13.00 (10.00, 19.00)	<0.001
AFP_Hotgen (ng/mL)[c]	8.74 (3.65, 284.75)	3.48 (2.76, 6.43)	2.38 (1.71, 3.29)	2.57 (1.32, 3.58)	3.00 (2.03, 4.08)	2.71 (1.76, 3.47)	<0.001
PIVKA-II_Hotgen (ng/mL)[c]	12.80 (4.28, 259.15)	4.12 (2.81, 5.14)	3.94 (2.84, 4.56)	3.71 (2.36, 4.51)	3.97 (2.94, 5.06)	3.79 (2.91, 4.56)	<0.001
AFP-L3_Hotgen (%)[c]	5.00 (5.00, 16.40)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	5.00 (5.00, 5.00)	<0.001
AFP_Abbott (ng/mL)[c]	6.58 (2.96, 211.85)	2.85 (1.29, 4.85)	2.22 (1.57, 3.20)	0.49 (0.28, 0.94)	0.81 (0.38, 2.18)	2.64 (1.94, 3.63)	<0.001
PIVKA-II_Abbott (mAU/mL)[c]	70.40 (24.82, 1214.40)	18.55 (12.63, 32.16)	20.87 (14.51, 25.32)	3.94 (2.62, 7.30)	6.50 (3.75, 19.13)	21.89 (18.29, 25.73)	<0.001

Notes: [a]Values were expressed as number (percentage) (Pearson Chi-square test). [b]Values were expressed as mean ± SD (One-Way ANOVA test). [c]Values were expressed as median (25th percentile, 75th percentile) (Kruskal–Wallis test).

Abbreviations: HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma; CRC, colorectal cancer; MHC, metastatic hepatic carcinoma; BLD, benign liver disease; ALT, alanine aminotransferase; AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence or antagonist-II; AFP-L3, alpha-fetoprotein lens culinaris agglutinin 3.

Furthermore, when comparing different models within the same platform, the AUCs of the GALAD_Hotgen and ASAP_Hotgen models were identical, with both achieving an AUC of 0.872. The ASAP_Abbott model (AUC = 0.913) exhibited a slightly higher AUC than the GALAD_Abbott model (AUC = 0.901), though this difference did not reach statistical significance (p = 0.0569). For comparisons of the same model across different platforms, the ASAP_Abbott model (AUC = 0.913) demonstrated a significantly higher AUC than the ASAP_Hotgen model (AUC = 0.872) (P = 0.0003). Similarly, the GALAD_Abbott model (AUC = 0.901) yielded a significantly higher AUC compared to the GALAD_Hotgen model (AUC = 0.872) (P = 0.0001).

It is worth explaining that due to the unavailability of AFP-L3 testing on the Abbott platform, when evaluating the GALAD_Abbott model, AFP-L3 was measured using the Hotgen platform, while other biomarkers were tested on the Abbott platform.

Comparison of Serum AFP and PIVKA-II Levels Between Hotgen Biotech and Abbott Architect

Passing-Bablok regression analysis revealed that the regression equation for AFP is AFP_Hotgen = 1.116 * AFP_Abbott + 0.221 ($P = 0.16$). The slope's 95% confidence interval (CI) was 1.066 to 1.162, and the intercept's 95% CI ranged from 0.111 to 0.373. Spearman correlation analysis revealed a correlation coefficient was 0.573 ($P < 0.001$). The regression equation for PIVKA-II is PIVKA-II_Hotgen = 0.093 * PIVKA-II_Abbott + 2.344 ($P < 0.01$). The slope's 95% CI was 0.078 to 0.108, and the intercept's 95% CI ranged from 2.052 to 2.627. The correlation coefficient from Spearman correlation analysis was 0.460 ($P < 0.001$). The scatter plots comparing AFP and PIVKA-II between the two platforms were displayed in Figure 2a and 2b, respectively.

Bland-Altman plots were used to compare AFP and PIVKA-II differences between the two platforms. According to the external quality control criteria of the China National Center for Clinical Laboratories, the allowable bias for both markers is ±30%, with 80% of data required to fall within this range. For AFP, the mean bias between the Hotgen and Abbott platforms was 44.32%, and only 57.43% (290/505) of the data points fell within the ±30% allowable bias range (Figure 2c), indicating poor consistency. For PIVKA-II, the mean bias between Hotgen and Abbott was −92.02%, with merely 9.32% (48/515) of data within the allowable range (Figure 2d). Notably, PIVKA-II concentrations measured by Hotgen were lower than those by Abbott in most samples.

Discussion

In this study, we evaluated the diagnostic performance of individual serum biomarkers (AFP, PIVKA-II, AFP-L3), and the composite models (GALAD and ASAP) for HCC across two distinct immunoassay platforms.

AFP levels were significantly higher in the HCC group than in all non-HCC groups across both platforms. The diagnostic efficacy of AFP was clinically favorable on both systems, with AUCs of 0.821 (Hotgen) and 0.846 (Abbott), showing no significant inter-platform difference (p > 0.05). However, its sensitivity remained insufficient even at the newly identified optimal

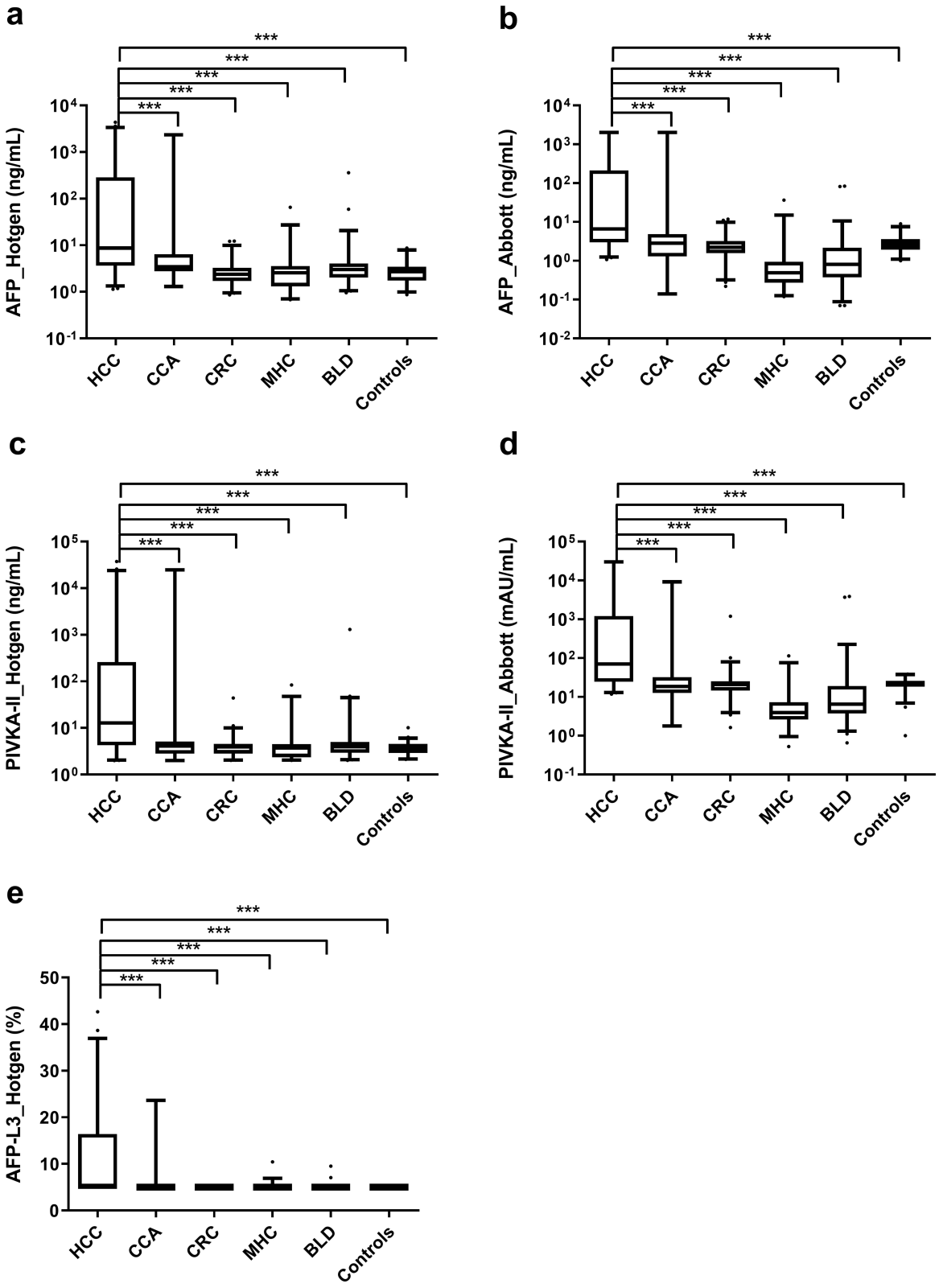

Figure 1 Expression levels of AFP_Hotgen (**a**), AFP_Abbott (**b**), PIVKA-II_Hotgen (**c**), PIVKA-II_Abbott (**d**) and AFP-L3_Hotgen (**e**) in HCC group and five non-HCC groups.

Notes: ***$P<0.001$.

Abbreviations: HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma; CRC, colorectal cancer; MHC, metastatic hepatic carcinoma; BLD, benign liver disease; AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence or antagonist-II; AFP-L3, alpha-fetoprotein lens culinaris agglutinin 3.

https://doi.org/10.2147/JHC.S554305

Table 2 Diagnostic Performance of Biomarkers and Models for HCC: ROC-Derived Optimal Cut-Offs and Commonly Used Clinical Cut-Offs

		Cut-off Value	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Hotgen	AFP_Hotgen (ng/mL)	> 4.58 [c]	0.821 (0.786–0.853)	69.61	81.25	47.65	91.60
	AFP_Hotgen (ng/mL)	> 20 [d]		42.16	98.56	87.77	87.42
	PIVKA-II_Hotgen (ng/mL)	> 5.85 [c]	0.787 (0.749–0.822)	66.67	89.90	61.81	91.67
	PIVKA-II_Hotgen (ng/mL)	> 40 [d]		44.12	97.84	83.36	87.72
	AFP-L3_Hotgen (%)	> 5.00 [c]	0.714 (0.673–0.752)	44.12	98.32	86.56	87.77
	AFP-L3_Hotgen (%)	> 10 [d]		38.24	99.04	90.71	86.74
	GALAD[a]_Hotgen	> −0.51 [c]	0.872 (0.839–0.899)	71.57	86.06	55.73	92.51
	ASAP[b]_Hotgen	> −2.20 [c]	0.872 (0.840–0.899)	68.63	91.11	65.43	92.22
Abbott	AFP_Abbott (ng/mL)	> 3.25 [c]	0.846 (0.811–0.876)	71.57	79.57	46.20	91.95
	AFP_Hotgen (ng/mL)	> 20 [d]		43.14	98.32	86.29	87.58
	PIVKA-II_Abbott (mAU/mL)	> 29.59 [c]	0.863 (0.831–0.892)	69.61	89.42	61.73	92.31
	PIVKA-II_Hotgen (ng/mL)	> 40 [d]		61.76	95.67	77.76	91.07
	GALAD[a]_Abbott	> −1.03 [c]	0.901 (0.872–0.925)	89.22	75.90	47.58	96.64
	ASAP[b]_Abbott	> −1.30 [c]	0.913 (0.885–0.936)	87.25	77.11	48.31	96.10

Notes: [a]The GALAD model, a composite model for HCC diagnosis, incorporates five parameters: gender, age, AFP-L3, AFP, and PIVKA-II.[10] [b]The ASAP model, a composite model for HCC diagnosis, incorporates four parameters: gender, age, AFP, and PIVKA-II.[15] [c]ROC-derived optimal cut-offs. [d]commonly used clinical cut-offs.
Abbreviations: HCC, hepatocellular carcinoma; ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence or antagonist-II; AFP-L3, alpha-fetoprotein lens culinaris agglutinin 3.

cut-off (3~5 ng/mL), reaching only ~70%. When using the commonly used clinical cut-off of 20 ng/mL, sensitivity further decreased to approximately 40%. Consistent with previous literature,[6,7] these results confirm that AFP alone provides inadequate sensitivity for HCC screening, leading to considerable missed diagnoses. Similarly, HCC patients exhibited elevated PIVKA-II levels, and Abbott's performance was marginally better than Hotgen's [(AUC: 0.863 vs 0.787, $P < 0.05$)]. Despite this, both platforms exhibited limited sensitivity and specificity for diagnosing HCC—whether at the optimal cut-offs or commonly used clinical cut-offs—as shown in Table 2. AFP-L3 consistently exhibited low sensitivity. This makes AFP-L3 more useful for excluding HCC than screening it, which is consistent with Debes JD's observations.[25] These findings underscore the constraints of relying on single biomarkers for diagnosing HCC,[26] highlighting that the application of composite models is the way forward for the serological diagnosis of HCC.

Nowadays, several composite models—including the GALAD and ASAP models—have emerged as promising tools for HCC diagnosis.[10,15] Consistent with previous literature, our findings demonstrate that both the GALAD and ASAP models exhibited significantly higher AUC values for HCC diagnosis compared to the individual serum markers AFP, PIVKA-II, and AFP-L3.[10–18] Numerous studies have compared the diagnostic performance of the GALAD and ASAP models for HCC, yet consistent conclusions remain elusive. In a multicenter study by Sun et al, the ASAP model (AUC = 0.886) demonstrated superior HCC detection efficacy compared to the GALAD model (AUC = 0.853), particularly in patients with chronic liver diseases and early-stage HCC.[19] For biomarker testing, AFP and PIVKA-II were measured via Abbott ARCHITECT kits, while AFP-L3 was assayed using either Fujirebio's Immunological Test System or Hotgen Biotech's micro centrifugal column. In contrast, Le et al reported comparable diagnostic performance between the ASAP (AUC = 0.96) and GALAD (AUC = 0.95) models for overall and early-stage HCC detection; the biomarker detection platform was unspecified.[20] Studies by Demirtas, Thanapirom, and their colleagues have focused on model performance across etiological subgroups, with both teams measuring biomarkers using the Wako uTASWako i30 analyzer.[21,22] Both research teams noted that the GALAD model achieved higher sensitivity in non-viral HCC, but discrepancies emerged regarding viral liver disease patients: Demirtas et al found comparable performance between the two models, while Thanapirom et al observed superior efficacy with the ASAP model.

In our study, and consistent with some previous reports, no significant difference in diagnostic performance was found between the GALAD and ASAP models when analyzed on the same platform. However, significant performance differences emerged when the same model was applied across different platforms. For the ASAP model, the AUC on the Abbott platform (0.913) was significantly higher than that on the Hotgen platform (0.871, $P = 0.0003$). A similar pattern was observed for the GALAD model, with its AUC on the Abbott platform (0.901) being significantly greater than that on the Hotgen platform (0.871,

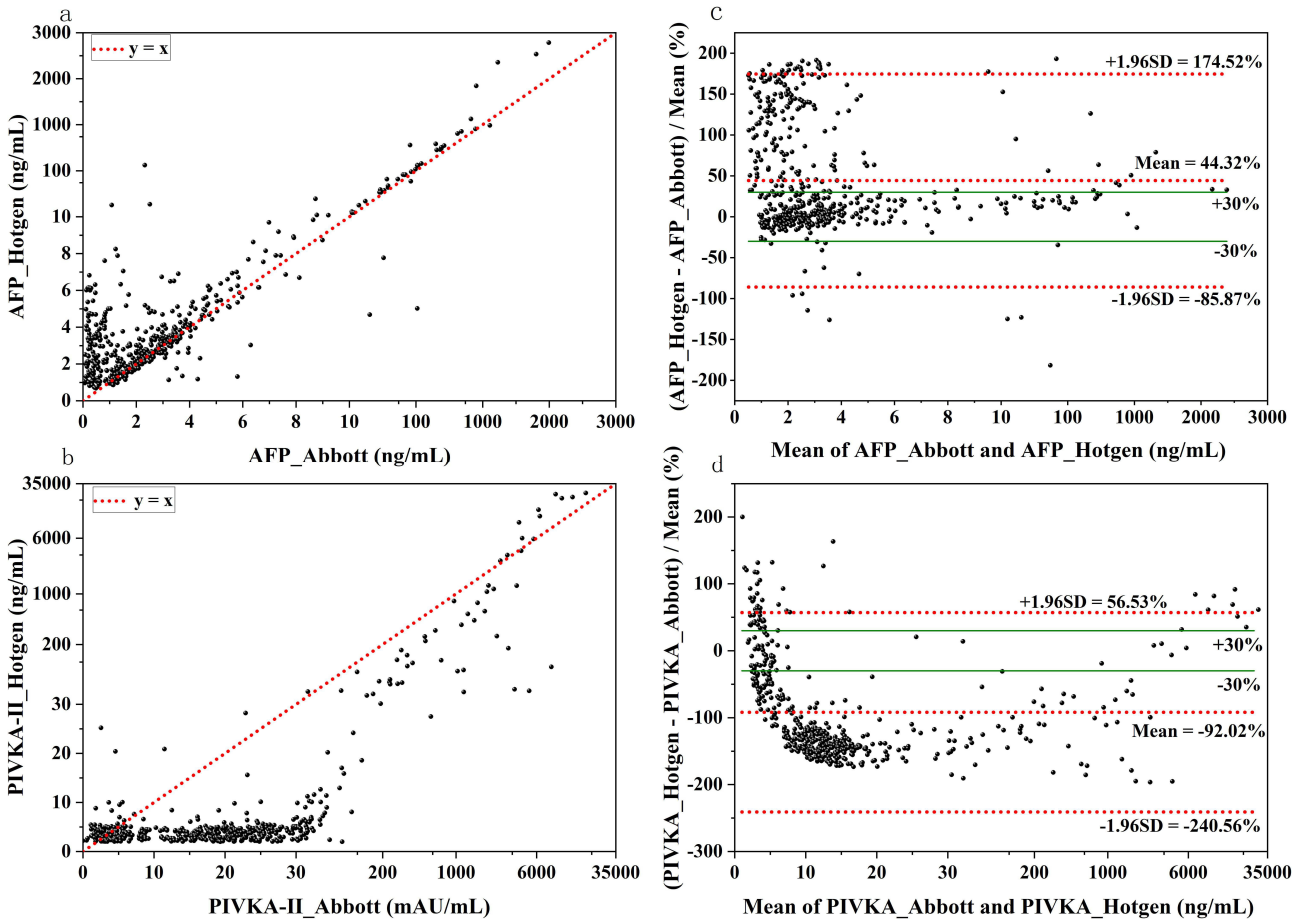

Figure 2 Quantitative comparison of serum AFP and PIVKA-II levels between Hotgen Biotech and Abbott Architect. (**a**) Scatter plot between AFP_Hotgen and AFP_Abbott. (**b**) Scatter plot between PIVKA-II_Hotgen and PIVKA-II_Abbott. (**c**) Bland-Altman plot between AFP_Hotgen and AFP_Abbott. (**d**) Bland-Altman plot between PIVKA-II_Hotgen and PIVKA-II_Abbott.

Abbreviations: AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence or antagonist-II; SD, standard deviation.

P = 0.0001). Previous studies have primarily focused on the impact of model-specific differences or study population variations, while the critical variable of detection platform has been overlooked. This oversight may partially explain the discrepancies in previous reports regarding which model has superior diagnostic efficacy. This study clearly demonstrates that the detection platform exerts a significant influence on the diagnostic performance of both the GALAD and ASAP models. This finding highlights the need to consider detection platform differences in clinical practice and cross-study comparisons. It also provides valuable insights for the standardized application and accurate interpretation of HCC diagnostic models.

In addition, we systematically compared AFP and PIVKA-II between the Hotgen Biotech and Abbott Architect platforms using various statistical methods. Passing-Bablok regression identified a significant linear relationship for AFP (slope = 1.116, intercept = 0.221, $P = 0.16$), indicating proportional and systematic biases, whereas PIVKA-II showed no linearity between platforms ($P < 0.001$). Spearman analysis confirmed moderate correlations (correlation coefficient = 0.573 for AFP, 0.460 for PIVKA-II), but Bland-Altman plots revealed clinically significant biases (mean bias: AFP + 44.32%, PIVKA-II −92.02%). Notably, despite these analytical differences, ROC analyses demonstrated comparable diagnostic efficacy for HCC across platforms—with AFP's AUC showing approximate consistency, while PIVKA-II had a minor inter-platform difference. This suggests that while absolute values are platform-dependent, the discriminative power remains stable However, it is recommended that the same detection system be used for the same patient throughout treatment and follow-up to avoid discrepancies due to platform differences.[27]

The key strength of this study lies in its novel dual-platform comparison of both individual biomarkers (AFP, PIVKA-II) and composite models (GALAD, ASAP) for HCC diagnosis. Concurrently, it has been revealed that the detection platform employed for biomarker quantification exerts a significant impact on the diagnostic performance of these models. However,

https://doi.org/10.2147/JHC.S554305

several limitations should be acknowledged: Firstly, tumor staging data for HCC patients were not obtained, making it impossible to assess the value of serum biomarkers and composite models for early HCC screening and diagnosis. Secondly, as a single-center study, it may be subject to selection bias. Thirdly, the study cohort exhibits class imbalance across different subgroups, and no specific correction strategies were implemented during data analysis. Future studies will expand the sample size of minority subgroups and adopt appropriate statistical methods to address this imbalance.

Data Confidentiality Statement

All personally identifiable information (eg, names, national ID numbers) was removed during data collection and replaced with unique anonymized codes. Access to research data was restricted to authorized study personnel, and data were stored on a password-protected server for the sole purpose of this scientific research.

Abbreviations

AFP, alpha fetoprotein; PIVKA-II, protein induced by vitamin K absence or antagonist-II; AFP-L3, alpha-fetoprotein lens culinaris agglutinin 3; HCC, hepatocellular carcinoma; CCA, cholangiocarcinoma; CRC, colorectal cancer; MHC, metastatic hepatic carcinoma; BLD, benign liver disease; ROC, receiver operating characteristic; AUC, area under the curve; GALAD, a HCC risk assessment model incorporating gender, age, AFP-L3, AFP, and PIVKA-II; ASAP, a HCC risk assessment model incorporating gender, age, AFP, and PIVKA-II; ALT, alanine aminotransferase; AASLD, American Association for the Study of Liver Diseases; CIs, confidence intervals; PPV, positive predictive value; NPV, negative predictive value.

Data Sharing Statement

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

This study was approved by the Ethical Committee of Peking Union Medical College Hospital (HS-2387) and conducted in accordance with the Declaration of Helsinki. Informed consent—including consent to review medical records—was waived by the Committee for the following reasons: This is a single-center study using anonymized residual clinical serum samples. Test results are for academic purposes only, not for clinical diagnosis or reagent registration. The study poses extremely low risk, as samples are in vitro and no direct participant contact is involved. All data are identified by sample IDs without personal identifying information, ensuring privacy protection.

Acknowledgments

The authors are grateful to all the patients, researchers, and institutions that participated in this study.

Funding

National Natural Science Foundation of China (No.62331025), Non-profit Central Research Institute Fund of the Chinese Academy of Medical Sciences (Grant No.: 2025-JKCS-10).

Disclosure

Dr Xu-Zhen Qin reports personal fees, free reagents from Abbott Company, during the conduct of the study. The authors declare no other competing interests in this work.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA*. 2021;71(3):209–249. doi:10.3322/caac.21660
2. Qi J, Li M, Wang L, et al. National and subnational trends in cancer burden in China, 2005-20: an analysis of national mortality surveillance data. *Lancet Public Health*. 2023;8(12):e943–e55. doi:10.1016/S2468-2667(23)00211-6
3. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology*. 2007;132(7):2557–2576. doi:10.1053/j.gastro.2007.04.061
4. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med*. 2019;380(15):1450–1462. doi:10.1056/NEJMra1713263

5. Wolf E, Rich NE, Marrero JA, et al. Use of hepatocellular carcinoma surveillance in patients with cirrhosis: a systematic review and meta-analysis. *Hepatology*. 2021;73(2):713–725. doi:10.1002/hep.31309
6. Kim DY, Toan BN, Tan CK, et al. Utility of combining PIVKA-II and AFP in the surveillance and monitoring of hepatocellular carcinoma in the Asia-Pacific region. *Clin Mol Hepatol*. 2023;29(2):277–292. doi:10.3350/cmh.2022.0212
7. Pang BY, Leng Y, Wang X, et al. A meta-analysis and of clinical values of 11 blood biomarkers, such as AFP, DCP, and GP73 for diagnosis of hepatocellular carcinoma. *Ann Med*. 2023;55(1):42–61. doi:10.1080/07853890.2022.2153163
8. De J, Shen Y, Qin J, et al. A systematic review of des-γ-carboxy prothrombin for the diagnosis of primary hepatocellular carcinoma. *Medicine*. 2016;95(17):e3448. doi:10.1097/MD.0000000000003448
9. Zhou JM, Wang T, Zhang KH. AFP-L3 for the diagnosis of early hepatocellular carcinoma: a meta-analysis. *Medicine*. 2021;100(43):e27673. doi:10.1097/MD.0000000000027673
10. 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
11. Singal A, Tayob N, Mehta A, et al. Doylestown plus and GALAD demonstrate high sensitivity for HCC detection in patients with cirrhosis. *Clin Gastroenterol Hepatol*. 2022;20(4):953–5.e2. doi:10.1016/j.cgh.2021.04.018
12. 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*. 2020;18(3):728–35.e4. doi:10.1016/j.cgh.2019.11.012
13. 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
14. 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):875–86.e6. doi:10.1016/j.cgh.2015.12.042
15. Yang T, Xing H, Wang G, et al. A novel online calculator based on serum biomarkers to detect hepatocellular carcinoma among patients with hepatitis B. *Clin Chem*. 2019;65(12):1543–1553. doi:10.1373/clinchem.2019.308965
16. Lu YC, Su TH, Tseng TC, et al. High PIVKA-II level and ASAP score predict 1-year risk of hepatocellular carcinoma in non-cirrhotic chronic hepatitis B patients. *Am J Cancer Res*. 2023;13(6):2588–2597.
17. Li B, Zhao Y, Cai W, et al. Validation and update of a multivariable prediction model for the identification and management of patients at risk for hepatocellular carcinoma. *Clin Proteomics*. 2021;18(1):21. doi:10.1186/s12014-021-09326-w
18. Wen R, Peng Y, Liang Y, et al. CEUS LI-RADS in combination with the serum biomarker-based asap model improves the diagnostic performance of HCC in high-risk patients. *Ultrasound Med Biol*. 2024;50(11):1739–1744. doi:10.1016/j.ultrasmedbio.2024.08.003
19. Sun LY, Wang NY, Diao YK, et al. Comparison between models for detecting hepatocellular carcinoma in patients with chronic liver diseases of various etiologies: ASAP score versus GALAD score. *Hepatobiliary Pancreat Dis Int*. 2025;24(4):412–422. doi:10.1016/j.hbpd.2023.12.004
20. Le TM, Pham KC. Comparative evaluation of ASAP and GALAD scores for detecting hepatocellular carcinoma in patients with chronic liver diseases. *J Clin Gastroenterol*. 2025. doi:10.1097/MCG.0000000000002257
21. Demirtas CO, Akin S, Yilmaz Karadag D, et al. Enhancing hepatocellular carcinoma surveillance: comparative evaluation of AFP, AFP-L3, DCP and composite models in a biobank-based case-control study. *Cancers*. 2025;17(14):2390. doi:10.3390/cancers17142390
22. Thanapirom K, Suksawatamnuay S, Thaimai P, et al. Comparison of the GALAD, GAAP, and ASAP scores for hepatocellular carcinoma detection in patients with chronic liver diseases. *J Clin Exp Hepatol*. 2025;15(6):102607. doi:10.1016/j.jceh.2025.102607
23. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology*. 2011;53(3):1020–1022. doi:10.1002/hep.24199
24. Singal AG, Llovet JM, Yarchoan M, et al. AASLD practice guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. *Hepatology*. 2023;78(6):1922–1965. doi:10.1097/HEP.0000000000000466
25. Debes JD, Romagnoli PA, Prieto J, et al. Serum biomarkers for the prediction of hepatocellular carcinoma. *Cancers*. 2021;13(7):1681. doi:10.3390/cancers13071681
26. Wang T, Zhang KH. New blood biomarkers for the diagnosis of AFP-negative hepatocellular carcinoma. *Front Oncol*. 2020;10:1316. doi:10.3389/fonc.2020.01316
27. Aijun Niu LD, Xu Jin G. Current status and recommendations for the application of tumor markers across different detection systems. *Chin J Clin Lab Sci*. 2021;39(3):161–164.