

Comparative Analysis of Multi-Biomarker Diagnostic Models for Early Detection of Hepatocellular Carcinoma and Their Potential Implications for Surgical Decision-Making

Ann. Ital. Chir., 2025 96, 10: 1322–1337
<https://doi.org/10.62713/aic.4182>

Xingfen Zhang^{1,†}, Tao Wang^{2,†}, Xiaozhen Xu¹, Guosheng Gao³, Miao Cheng³, Xijie Lai¹, Dafeng Mao⁴

¹Department of Liver Disease, Ningbo No.2 Hospital, 315010 Ningbo, Zhejiang, China

²Guoke Ningbo Life Science and Health Industry Research Institute, 315010 Ningbo, Zhejiang, China

³Department of Clinical Laboratory, Ningbo No.2 Hospital, 315010 Ningbo, Zhejiang, China

⁴Department of Ultrasound, Ningbo Yinzhou No.2 Hospital, 315100 Ningbo, Zhejiang, China

AIM: Hepatocellular carcinoma (HCC) remains a significant global health concern, often diagnosed at advanced stages, limiting the efficacy of surgical interventions. Early and accurate diagnosis is critical for improving surgical outcomes and reducing mortality. Traditional biomarkers, such as alpha-fetoprotein (AFP), des-gamma-carboxyprothrombin (DCP), and the lectin-bound fraction of AFP (AFP-L3), show limited sensitivity and specificity. Advanced diagnostic models, including GALAD, TAGALAD, and GAP_TALAD, offer a promising multi-biomarker approach but lack extensive evaluation in surgical contexts.

METHODS: This retrospective study included a cohort of 267 untreated hepatocellular carcinoma patients and 231 control patients (with hepatitis or cirrhosis). We applied the predefined formulas for the TAGALAD, GAP_TALAD, and other models to the cohort data. The diagnostic performance of each model and individual biomarker for detecting HCC was assessed using receiver operating characteristic (ROC) curve analysis to determine the area under the curve (AUC), sensitivity, and specificity at optimal cut-offs. Additionally, key clinical subgroups, including pathologically confirmed HCC, clinically diagnosed HCC, early-stage HCC (TNM I-II), patients with complete data (no imputation), and hepatitis B virus (HBV)-related disease, were also analyzed.

RESULTS: TAGALAD and GAP_TALAD demonstrated superior performance compared to the GALAD model and traditional biomarkers across all patient subgroups. Notably, TAGALAD achieved the highest diagnostic accuracy, with an AUC of 0.880, sensitivity of 0.760, and specificity of 0.861, followed closely by GAP_TALAD (AUC = 0.874). Both models demonstrated excellent performance in early-stage HCC detection (TAGALAD AUC = 0.860, GAP_TALAD AUC = 0.867), highlighting their potential in identifying candidates for surgical resection or transplant at an early curative stage. In HBV-related HCC, TAGALAD (AUC = 0.874) and GAP_TALAD (AUC = 0.857) showed superior diagnostic accuracy compared to GALAD (AUC = 0.731) and single biomarkers (AUC = 0.598–0.799).

CONCLUSIONS: The TAGALAD and GAP_TALAD models offer a robust and reliable framework that supports early diagnosis of HCC. Their superior accuracy indicates a more reliable foundation for identifying candidates for curative surgical interventions, suggesting the potential to refine patient selection. Future research should focus on multi-center validation and the integration of novel biomarkers to further optimize these models for surgical decision-making and personalized treatment strategies.

Keywords: hepatocellular carcinoma; biomarker; diagnosis; alpha-fetoprotein; des-gamma-carboxyprothrombin; alpha-fetoprotein variant L3

Introduction

Hepatocellular carcinoma (HCC) poses a significant global health challenge due to its high incidence and mortality rates. It ranks as the sixth most common cancer and the third leading cause of cancer-related deaths in 2020 [1]. The high mortality associated with HCC is largely attributed to

the asymptomatic early stages, often diagnosed at advanced stages when treatment options are limited and prognosis is poor. As a primary liver malignancy, HCC predominantly develops in chronic liver disease and cirrhosis, with hepatitis B virus (HBV) and hepatitis C virus (HCV) infections being the primary etiological factors. The epidemiology of HCC exhibits significant geographical variation. In regions like China, where HBV is endemic, the annual incidence of HCC exceeds 25 cases per 100,000 people, accounting for over 50% of the global HCC burden [2,3]. In contrast, Western countries such as the United States and parts of Europe report lower incidence rates, typically ranging from 6–10 cases per 100,000 individuals, with HCV, alcohol-related liver disease, and metabolic disorders being the major risk factors [4,5].

Submitted: 25 May 2025 Revised: 21 July 2025 Accepted: 29 July 2025 Published: 10 October 2025

Correspondence to: Dafeng Mao, Department of Ultrasound, Ningbo Yinzhou No.2 Hospital, 315100 Ningbo, Zhejiang, China (e-mail: 190536996@qq.com).

[†]These authors contributed equally.

The pathological progression of HCC is a multistep process characterized by the malignant transformation of hepatocytes, driven by genetic alterations, chronic inflammation, and environmental factors [6]. Key contributors to this process include chronic HBV and HCV infections, alcohol consumption, obesity, and aflatoxin exposure. However, the primary etiological factors vary significantly by geographic region. For instance, HBV-related carcinogenesis, often exacerbated by aflatoxin exposure, is particularly prevalent in China, while metabolic-associated fatty liver disease (MAFLD) is emerging as a significant risk factor in Western countries [6]. The global burden of HCC is expected to rise further due to the increasing prevalence of obesity and diabetes, underscoring the urgent need for effective preventive and diagnostic strategies.

Traditional serum biomarkers, such as alpha-fetoprotein (AFP), des-gamma-carboxyprothrombin (DCP), and the lectin-bound fraction of AFP (AFP-L3), have been widely employed for diagnosing and monitoring HCC. However, their clinical efficacy has been questioned due to limited sensitivity and specificity. For instance, AFP alone has an approximate sensitivity of 60% for detecting early-stage HCC, making it unreliable for early diagnosis [7,8]. Additionally, AFP's diagnostic specificity is compromised since its levels can be elevated in non-cancerous conditions such as hepatitis, cirrhosis, and pregnancy [9]. Although DCP and AFP-L3 offer slightly improved specificity, they remain insufficient as standalone biomarkers. A study in Japan reported a sensitivity of 70% for DCP, which is still suboptimal for early diagnosis [10]. Similarly, another study found that AFP-L3 offers a sensitivity of 65% in differentiating HCC from liver cirrhosis, indicating its limitations for definitive diagnosis [11,12]. Given these observations, there is an urgent need for more comprehensive approaches that integrate multiple biomarkers and patient-specific factors to improve diagnostic accuracy across diverse patient populations.

To address limitations of traditional biomarkers, recent advancements have focused on multi-biomarker diagnostic models that enhance diagnostic accuracy by integrating serum biomarkers with clinical parameters through multivariate logistic regression. One of the most significant models is the GALAD model, which incorporates gender, age, AFP, AFP-L3, and DCP and has proven to be a promising tool [11,13]. In a multi-center study, the GALAD model demonstrated a substantial improvement in diagnosing early-stage HCC, achieving an area under the receiver operating characteristic curve (AUROC) of over 0.90 [6]. Building on this integrated approach, other models have emerged to further improve the application of multi-biomarker diagnostics. For example, the ASAP model, which combines age, sex, AFP, and specific antibodies, showed a sensitivity of 85% and specificity of 88% in a Japanese cohort. Similarly, the GAAP model, which extends the GALAD framework by including Protein Induced

by Vitamin K Absence (PIVKA)-II, achieved an AUROC of 0.88 for early-stage HCC detection in a multi-ethnic validation study. The development of these models highlights ongoing efforts to optimize diagnostic accuracy and underscores their significance as key benchmarks in our comparative analysis.

Despite these innovations, multi-biomarker models encounter several challenges. A primary concern is the variability in their performance across diverse populations, which limits their universal applicability. For instance, while the GALAD model performed exceptionally well in Western populations, its sensitivity was slightly lower in Asian cohorts, likely due to differences in genetic and etiological factors [7]. Additionally, the reliance on multiple assays can increase costs, posing barriers to widespread adoption, particularly in low-resource settings [8]. These challenges highlight the critical need for additional validation and comparative analysis of existing models within specific demographic groups.

Therefore, this study aims to comprehensively evaluate and compare the diagnostic performance of several novel multi-biomarker models, including TAGALAD and GAP_TALAD, against the established GALAD model and traditional individual markers (AFP, DCP, AFP-L3) within our patient cohort. By identifying the most robust diagnostic framework, this study intends to enhance clinical decision-making through more accurate and early detection of HCC, ultimately providing a stronger foundation for stratifying high-risk patients and facilitating timely and potentially curative interventions [14].

Materials and Methods

Study Subjects and Specimen Requirements

This retrospective study investigated diagnostic biomarkers and methodologies for detecting liver cancer. Patient data were collected from Ningbo No. 2 Hospital between January 2021 and December 2024. Study participants were recruited following a predetermined inclusion-exclusion criteria.

Inclusion criteria for patient selection were as follows: (1) hospitalized patients who underwent simultaneous testing for the three HCC biomarkers (AFP, AFP-L3, and DCP); (2) age 18 years or older; (3) complete primary clinical information (including imaging examinations) in the Hospital Information System (HIS) and primary laboratory test data in the Laboratory Information System (LIS); (4) definitive diagnosis based on established clinical criteria, including imaging (magnetic resonance imaging/computed tomography (MRI/CT)), serum AFP levels, and histopathological confirmation where available for HCC patients; for controls, confirmation of hepatitis or cirrhosis without evidence of HCC via medical records; and (5) ethical approval from the Institutional Review Board (IRB) Ningbo No. 2 Hospital (No. PJ-NBEY-KY-2024-087-01) with a waiver of informed consent for this retrospective analysis, as it involved

de-identified data and posed minimal risk to participants.

Patients were excluded if they met the following criteria:

(1) patients tested for only one or two of the three HCC biomarkers; (2) age under 18 years; (3) current warfarin therapy; (4) HCC biomarker testing performed after tumor-specific treatment in cancer patients; (5) unclear discharge diagnosis; or (6) incomplete primary clinical information (including imaging examinations) and laboratory test data, or patients unable to cooperate.

To minimize potential variations due to circadian rhythms, serum samples were collected from all participants between 7:00 AM and 9:00 AM after an overnight fast of at least 8 hours. Blood samples were processed within two hours of collection, with serum separated by centrifugation at $2000 \times g$ for 10 minutes at 4°C . The serum samples were then aliquoted and stored immediately at -80°C until analysis to prevent protein degradation. The patient selection process, including screening, exclusions, and final cohort composition, is illustrated in **Supplementary Fig. 1**.

Diagnosis of Liver Cancer

Liver cancer was confirmed through imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT) scans, alongside serum AFP level assessments. Histopathological analysis was performed following surgical resection or biopsy to confirm the diagnosis when required. Advanced imaging approaches enabled precise localization and detailed characterization of liver tumors.

Serum Biomarker Detection

Serum samples were collected from venous blood within 24 hours. Serum samples were either analyzed immediately, or they were stored at -80°C until use. The detection of key biomarkers was conducted as follows:

DCP: Measured using a magnetic microparticle chemiluminescent immunoassay on the C3000 immunoassay analyzer (Beijing Hotgen Biotech Co., Ltd., Beijing, China). AFP: Measured using a direct chemiluminescent double-antibody sandwich method on the ADVIA Centaur XP fully automated immunoassay analyzer (Siemens Healthineers, Malvern, PA, USA). AFP-L3: Measured using an immunofluorescence reaction method on the μTAS Wako i30 fully automated electrophoresis fluorescence immunoassay analyzer (Fujifilm, Tokyo, Japan).

Application of Diagnostic Models

This study evaluated the diagnostic performance of several previously established multi-biomarker models. These models, including GALAD and its successors, are primarily based on multivariable logistic regression formulas developed and published in prior research. We applied these predefined formulas to our cohort data to generate diagnostic scores and compare the models' performance. The primary models evaluated were:

(1) GALAD: Integrates gender, age, AFP, AFP-L3, and DCP.

(2) ASAP: Incorporates age, sex, AFP, and specific antibodies to enhance diagnostic accuracy.

(3) GAAP: Extends the GALAD framework by including PIVKA-II.

(4) TAGALAD: A refined model incorporating age, gender, log (AFP), AFP-L3, log (DCP), log (Total Bilirubin (TBIL)), and albumin (ALB).

(5) GAP_TALAD: A model incorporating age, gender, platelet count (PLT), TBIL, ALB, AFP-L3, log (AFP), and log (DCP).

Furthermore, other models, including C_GALAD_II, GALAD_C, and LAD, were also evaluated based on their respective published formulas. This comparative analysis aimed to identify the most robust model for HCC detection within our study population.

Statistical Analysis

Statistical analysis was conducted using R software (version 4.3.3; R Foundation for Statistical Computing, Vienna, Austria). Missing data rates were analyzed, and the five indicators with the highest missing rates were found to be the large platelet ratio (62.45%), Cancer Antigen 125 (CA125, 44.58%), ferritin (42.57%), carbohydrate antigen 19-9 (CA19-9, 41.77%), and Carcinoembryonic Antigen (CEA, 41.77%). Among these, the large platelet ratio, with a missing rate exceeding 50%, was excluded from subsequent analyses. Missing data for the remaining variables were imputed using the 'mice' package with the predictive mean matching (PMM) method. To validate the imputation process, we compared the density plots of the imputed variables before and after imputation. The plots showed good consistency in data distribution, suggesting that the imputation process did not introduce significant bias (**Supplementary Fig. 2**).

The statistical analysis of variables was based on their type and distribution. Continuous variables were evaluated for normality using the Shapiro-Wilk test. Normally distributed continuous variables were presented as mean $\pm$ standard deviation and compared using an independent samples *t*-test. Non-normally distributed continuous variables were presented as medians and interquartile ranges (p25, p75) and analyzed using the Mann-Whitney U test. Furthermore, categorical variables were expressed as counts and percentages (n, %) and analyzed using the chi-square (χ^2) test. Statistical significance was considered at a *p*-value of less than 0.05.

Variables with significant *p*-values (<0.05) in univariate analysis were evaluated for skewness. Those with absolute skewness values greater than 2 underwent a log10 transformation to normalize their distributions. Following this initial cohort characterization, the primary objective was to evaluate and compare the performance of specific, predefined diagnostic models previously established in the lit-

erature (i.e., GALAD, ASAP, GAAP, and others), as well as their recent extensions (TAGALAD, GAP_TALAD). Consequently, the variables included in the subsequent diagnostic accuracy analyses were selected not through an exploratory analysis of all significant markers in Table 1, but based on their inclusion in the formulas of these established models. This approach ensures a direct and valid comparison with prior research.

To evaluate the diagnostic performance of biomarkers and models, sensitivity, specificity, and the AUROC were calculated. The optimal cut-off values for each model and biomarker were determined using the Youden index, which maximizes the sum of sensitivity and specificity. DeLong's test was used to compare AUROC values between models, providing statistical validation of performance differences. These methods ensured a rigorous and comprehensive assessment of the diagnostic models' accuracy.

Results

Baseline Data and Patient Characteristics

A total of 267 untreated HCC patients were included in the disease group, while 231 patients with hepatitis and liver cirrhosis formed the control group. These groups were used to construct diagnostic models for HCC.

The baseline characteristics of the two groups are summarized in Table 1. Patients in the disease group were significantly older than those in the control group. The disease group also showed significantly higher levels of the three liver cancer biomarkers (AFP-L3, DCP, and AFP). Additionally, the prevalence of cirrhosis, hypertension, and hepatitis virus infections, particularly chronic hepatitis B (CHB) was substantially higher in the disease group. Furthermore, the disease group also showed significantly higher smoking rates.

Hematological and biochemical indicators were substantially increased in the disease group compared to the control group. These included fibrinogen, PLT, white blood cell (WBC) count, neutrophil percentage (NEU%), red blood cell (RBC) count, hemoglobin (HGB), Apolipoprotein A1 (ApoA1), lipoprotein (a) (Lp (a)), Sialic Acid (SA), blood urea nitrogen (BUN), glucose (GLU), total protein (TP), ALB, and cholinesterase (CHE). These parameters reflect the metabolic and inflammatory changes associated with HCC.

Conversely, several indicators were significantly lower in the disease group compared to the control group. These included conditions such as alcoholic liver disease, drug-induced liver disease, autoimmune liver disease, and liver failure. Biochemical markers, such as TBIL, direct bilirubin (DBIL), total bile acids (TBA), CA125, ferritin, mean platelet volume (MPV), lymphocyte percentage (LYM%), red cell distribution width (RDW), Apolipoprotein E (ApoE), angiotensin converting enzyme (ACE), and phosphorus, were found to be reduced in the disease group.

Additionally, levels of CK-MB, a cardiac enzyme, were also significantly lower in the disease group.

These differences in baseline characteristics between the disease and control groups highlight the distinct metabolic, inflammatory, and clinical profiles of patients with HCC, suggesting that these parameters could potentially be incorporated into diagnostic models.

Diagnostic Accuracy of the Models (Comparison Between the Models)

The skewness function was applied to variables with a univariate analysis p -value < 0.05 to assess their distribution. Variables with an absolute skewness value greater than 2 were considered severely skewed and underwent a log10 transformation. Moreover, all continuous variables were standardized using the scale function to ensure uniformity in subsequent analyses. Based on the predefined formulas for each model, the diagnostic performance of AFP-L3, DCP, AFP, and various models (GALAD, ASAP, GAAP, C_GALAD_II, GALAD_C, GAP_TALAD, LAD, C_GALAD, and TAGALAD) was evaluated using receiver operating characteristic (ROC) curve analysis.

The findings of the ROC curve analysis are presented in Fig. 1 and Table 2. Among the single biomarkers, AFP-L3, DCP, and AFP showed relatively low diagnostic accuracy, with area under the curve (AUC) values of 0.662 (95% CI: 0.616–0.708), 0.787 (95% CI: 0.746–0.827), and 0.654 (95% CI: 0.606–0.701), respectively. These results highlight the limited sensitivity and specificity of individual biomarkers in diagnosing HCC.

In contrast, multi-biomarker models demonstrated significantly improved diagnostic performance. The TAGALAD model achieved the highest AUC of 0.880 (95% CI: 0.851–0.909), followed closely by the GAP_TALAD model, which yielded an AUC of 0.874 (95% CI: 0.843–0.905). Both models outperformed other diagnostic systems, including the GALAD model (AUC = 0.779, 95% CI: 0.739–0.819), which has been widely validated in previous studies. Other models, such as ASAP (AUC = 0.813), GAAP (AUC = 0.807), and GALAD_C (AUC = 0.831), also showed better diagnostic performance than individual biomarkers, though they did not reach the accuracy levels observed with TAGALAD and GAP_TALAD.

The sensitivity and specificity of TAGALAD and GAP_TALAD were substantially higher than those of other models. TAGALAD achieved a sensitivity of 0.760 (95% CI: 0.663–0.813) and a specificity of 0.861 (95% CI: 0.758–0.909), while GAP_TALAD yielded a sensitivity of 0.820 (95% CI: 0.738–0.865) and a specificity of 0.823 (95% CI: 0.658–0.874). These findings suggest that TAGALAD and GAP_TALAD offer the most reliable diagnostic performance for distinguishing HCC from non-HCC conditions. Conversely, the GALAD model, while outperforming individual biomarkers, demonstrated moderate sensitivity (0.667) and specificity (0.805), suggesting room for improvement.

Table 1. Comparison of baseline indicators between the disease group and the control group.

Characteristic	Control group (hepatitis and liver cirrhosis, n = 231)	Disease group (HCC, n = 267)	p-value
Gender			0.468
Male	52 (22.511)	53 (19.850)	
Female	179 (77.489)	214 (80.150)	
Age, years, mean (SD)	56.143 (12.367)	61.577 (10.969)	<0.001
AFP-L3(%)	1.100 (0.500, 8.300)	7.500 (0.500, 32.100)	<0.001
DCP (mAU/mL)	19.000 (14.720, 32.715)	156.000 (27.000, 3243.000)	<0.001
AFP (ng/mL)	6.000 (2.300, 39.450)	30.400 (4.050, 303.700)	<0.001
Comorbidity			
Liver cirrhosis	162 (70.130)	212 (79.401)	0.017
Fatty liver	21 (9.091)	13 (4.869)	0.062
Alcoholic liver disease	37 (16.017)	19 (7.116)	0.002
Drug-induced liver disease	8 (3.463)	1 (0.375)	0.014
Autoimmune liver disease	17 (7.359)	5 (1.873)	0.003
Hypertension	45 (19.481)	91 (34.082)	<0.001
Type 2 diabetes	41 (17.749)	46 (17.228)	0.879
Hepatic encephalopathy	35 (15.152)	41 (15.356)	0.950
Esophageal and gastric varices	40 (17.316)	42 (15.730)	0.634
Hepatic failure	35 (15.152)	10 (3.745)	<0.001
Hepatorenal syndrome	6 (2.597)	3 (1.124)	0.469
Chronic hepatitis B	143 (61.905)	219 (82.022)	<0.001
Smoke	51 (22.078)	89 (33.333)	0.005
Alcohol	52 (22.511)	76 (28.464)	0.129
BMI (kg/m ²)	23.430 (21.000, 25.745)	23.000 (20.940, 24.910)	0.196
Family history of liver cancer	13 (5.628)	15 (5.618)	0.996
DD (ng/mL)	200.000 (137.000, 652.000)	186.000 (137.000, 523.000)	0.382
PT (s)	13.200 (11.750, 15.100)	12.300 (11.400, 13.450)	<0.001
INR	1.160 (1.040, 1.335)	1.070 (1.010, 1.180)	<0.001
FIB (g/L)	2.520 (1.950, 3.250)	3.140 (2.520, 3.900)	<0.001
TT (s)	24.300 (19.450, 26.850)	22.400 (20.300, 25.050)	0.010
APTT (s)	33.300 (30.050, 37.100)	31.700 (29.100, 34.400)	<0.001
CA125 (U/mL)	16.000 (7.600, 154.600)	12.300 (6.850, 29.800)	0.003
CA19-9 (U/mL)	19.030 (9.590, 30.040)	15.850 (8.450, 26.290)	0.053
CEA (ng/mL)	1.890 (1.120, 2.800)	2.000 (1.165, 2.920)	0.588
FER (ng/mL)	323.400 (96.400, 645.850)	195.700 (95.500, 364.100)	0.001
PLT (10 ⁹ /L)	106.000 (67.000, 165.000)	132.000 (91.000, 179.500)	<0.001
MPV (fL)	10.100 (9.100, 10.900)	9.800 (8.900, 10.750)	0.047
PCT (%)	0.100 (0.065, 0.160)	0.130 (0.090, 0.180)	<0.001
PDW (fL)	16.500 (15.700, 17.150)	16.500 (15.800, 17.000)	0.462
WBC (10 ⁹ /L)	4.500 (3.300, 5.600)	5.100 (3.700, 6.300)	0.001
NEU%	61.006 ± 12.695	63.901 ± 11.732	0.008
LYM%	27.100 ± 11.052	25.001 ± 10.049	0.027
MONO%	9.000 (6.850, 11.150)	7.900 (6.400, 9.750)	0.004
EOS%	1.700 (1.000, 3.000)	1.500 (0.750, 2.800)	0.092
BASO%	0.400 (0.300, 0.600)	0.400 (0.300, 0.600)	0.557
RBC (10 ¹² /L)	3.930 (3.500, 4.480)	4.270 (3.700, 4.675)	0.002
HGB (g/L)	126.000 (109.500, 140.000)	132.000 (117.000, 145.000)	0.002
MCV (fL)	95.100 (90.250, 99.950)	94.100 (90.350, 97.850)	0.259
RDW (fL)	14.700 (13.700, 16.350)	13.600 (13.100, 14.700)	<0.001
TC (mmol/L)	3.820 (3.165, 4.670)	4.030 (3.405, 4.715)	0.077
TG (mmol/L)	1.040 (0.760, 1.605)	0.980 (0.730, 1.310)	0.124
HDL (mmol/L)	1.090 (0.845, 1.370)	1.130 (0.945, 1.380)	0.112
LDL-C (mmol/L)	2.090 (1.680, 2.755)	2.340 (1.835, 2.850)	0.054

Table 1. Continued.

Characteristic	Control group (hepatitis and liver cirrhosis, n = 231)	Disease group (HCC, n = 267)	p-value
ApoA1 (g/L)	1.110 (0.830, 1.320)	1.160 (0.960, 1.330)	0.046
ApoB (g/L)	0.690 (0.560, 0.930)	0.730 (0.590, 0.875)	0.469
ApoE (mg/dL)	5.900 (4.900, 7.150)	5.400 (4.500, 6.350)	0.002
Lp (a) (mg/L)	36.200 (15.400, 81.450)	52.000 (25.450, 106.000)	<0.001
SA (mg/dL)	49.200 (42.750, 57.550)	54.900 (47.450, 63.300)	<0.001
hs-CRP (mg/L)	2.500 (1.070, 7.870)	2.720 (0.915, 14.630)	0.530
BUN (mmol/L)	4.770 (3.675, 5.755)	5.150 (4.035, 6.425)	0.003
UA (μmol/L)	320.200 (250.800, 391.650)	320.200 (260.700, 372.700)	0.924
Cr (μmol/L)	64.500 (56.150, 71.800)	66.000 (56.400, 78.200)	0.104
ACE (U/L)	128.000 (80.500, 178.000)	99.000 (72.000, 148.500)	<0.001
GLU (mmol/L)	5.110 (4.320, 6.860)	5.640 (4.865, 6.860)	<0.001
K (mmol/L)	3.820 (3.535, 4.105)	3.870 (3.600, 4.115)	0.153
Na ⁺ (mmol/L)	139.000 (137.350, 141.350)	139.900 (137.750, 141.700)	0.099
Cl ⁻ (mmol/L)	105.200 (102.700, 107.450)	105.300 (102.800, 107.400)	0.824
CO ₂ CP (mmol/L)	24.075 ± 2.626	24.097 ± 2.727	0.927
P (mmol/L)	1.100 (0.980, 1.215)	1.040 (0.920, 1.140)	<0.001
TBIL (μmol/L)	22.000 (13.000, 42.250)	15.300 (10.150, 22.500)	<0.001
DBIL (μmol/L)	10.200 (4.900, 26.450)	6.100 (3.800, 10.800)	<0.001
TP (g/L)	65.800 (59.700, 71.200)	69.300 (64.850, 73.350)	<0.001
ALB (g/L)	34.300 (30.200, 39.250)	38.900 (33.600, 42.450)	<0.001
AST (U/L)	47.000 (32.500, 96.500)	43.000 (31.000, 77.000)	0.086
ALT (U/L)	34.000 (21.000, 85.500)	31.000 (21.000, 50.000)	0.077
ALP (U/L)	118.000 (95.000, 156.500)	116.000 (88.500, 171.500)	0.899
GGT (U/L)	70.000 (33.500, 160.000)	65.000 (34.000, 159.500)	0.990
LDH (U/L)	209.000 (175.000, 271.000)	217.000 (184.000, 270.500)	0.193
CHE (U/L)	4140.000 (3102.000, 6010.000)	5396.000 (3681.500, 6949.000)	<0.001
ADA (U/L)	15.700 (11.450, 20.100)	14.400 (10.800, 19.100)	0.053
CK (U/L)	91.000 (69.000, 128.000)	98.000 (73.000, 133.000)	0.228
CK-MB (U/L)	20.000 (13.000, 34.000)	17.000 (12.000, 28.000)	0.021
AFU (U/L)	28.700 (21.750, 38.550)	27.600 (21.550, 35.500)	0.442
TBA (μmol/L)	28.600 (10.900, 99.550)	13.500 (6.000, 36.450)	<0.001

Note: Data are presented as mean ± SD or median (p25, p75). Abbreviations: HCC, hepatocellular carcinoma; AFP-L3, lectin-bound fraction of AFP; DCP, des-gamma-carboxyprothrombin; AFP, alpha-fetoprotein; BMI, body mass index; DD, D-dimer; PT, Prothrombin Time; INR, International Normalized Ratio; FIB, Fibrinogen; TT, Thrombin Time; APTT, Activated Partial Thromboplastin Time; CA125, Cancer Antigen 125; CA19-9, Cancer Antigen 19-9; CEA, Carcinoembryonic Antigen; FER, Ferritin; PLT, platelet count; MPV, mean platelet volume; PCT, Plateletcrit; PDW, Platelet Distribution Width; WBC, white blood cell; NEU%, neutrophil percentage; LYM%, lymphocyte percentage; MONO%, Monocyte Percentage; EOS%, Eosinophil Percentage; BASO%, Basophil Percentage; RBC, red blood cell; HGB, hemoglobin; MCV, Mean Corpuscular Volume; RDW, red cell distribution width; TC, Total Cholesterol; TG, Triglycerides; HDL, High-Density Lipoprotein; LDL-C, Low-Density Lipoprotein Cholesterol; ApoA1, Apolipoprotein A1; ApoB, Apolipoprotein B; ApoE, Apolipoprotein E; Lp (a), lipoprotein (a); SA, Sialic Acid; hs-CRP, High-Sensitivity C-Reactive Protein; BUN, blood urea nitrogen; UA, Uric Acid; Cr, Creatinine; ACE, angiotensin converting enzyme; GLU, glucose; K, Potassium; Na⁺, Sodium; Cl⁻, Chloride; CO₂CP, Carbon Dioxide Combining Power; P, Phosphorus; TBIL, Total Bilirubin; DBIL, direct bilirubin; TP, total protein; ALB, albumin; AST, Aspartate Aminotransferase; ALT, Alanine Aminotransferase; ALP, Alkaline Phosphatase; GGT, Gamma-Glutamyl Transferase; LDH, Lactate Dehydrogenase; CHE, cholinesterase; ADA, Adenosine Deaminase; CK, Creatine Kinase; CK-MB, Creatine Kinase-MB; AFU, Alpha-L-Fucosidase; TBA, total bile acids.

Overall, the results underscore the diagnostic superiority of TAGALAD and GAP_TALAD, particularly in terms of AUC values, sensitivity, and specificity. These findings suggest that incorporating additional parameters and advanced computational approaches into these models provides significant advantages over traditional diagnostic ap-

proaches. A detailed comparison of these models is shown in Fig. 1 and Table 2.

Sensitivity Analysis of the Models

To assess the robustness of our primary findings, a series of sensitivity analyses was conducted across several key clin-

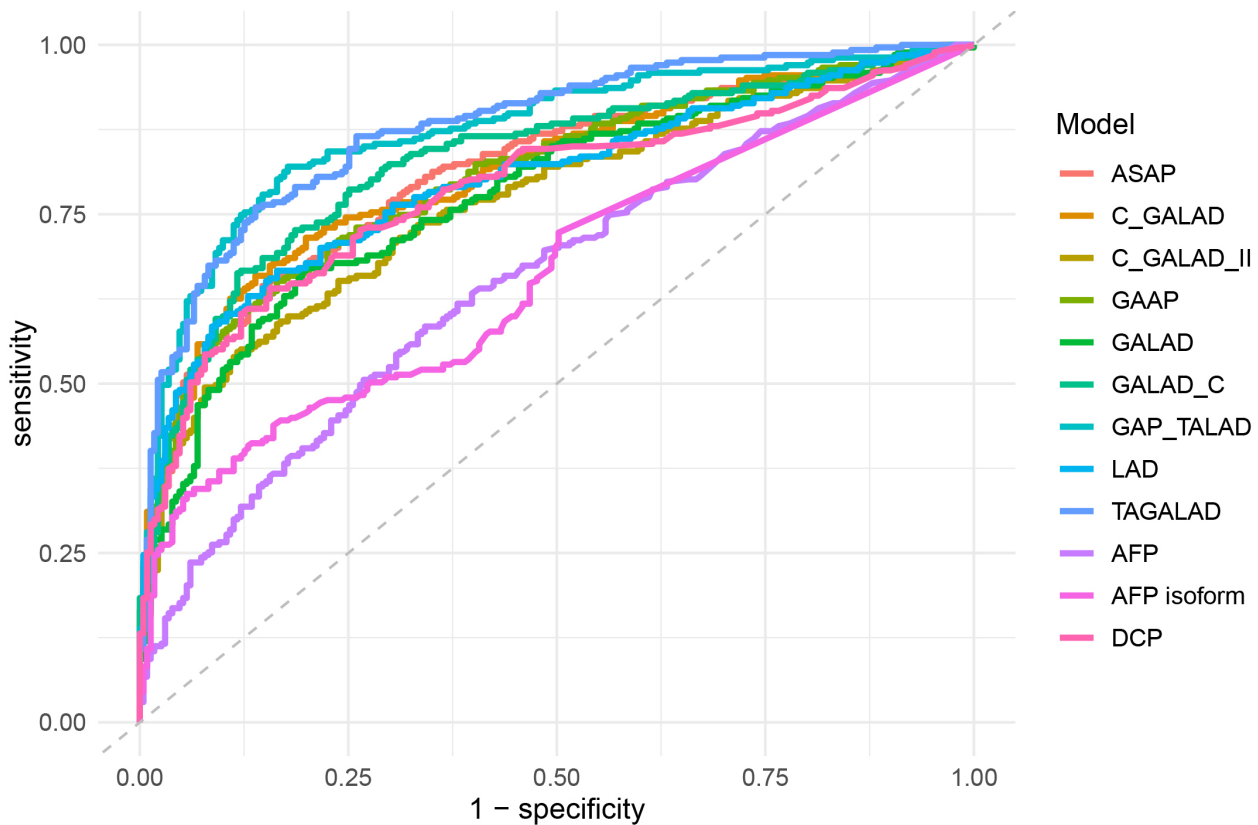


Fig. 1. Receiver operating characteristic (ROC) curves of the models and other individual indicators.

Table 2. ROC curve analysis of the models and other individual indicators.

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.662 (0.616–0.708)	0.337 (0.248–0.393)	0.944 (0.866–0.978)	20.850
DCP	0.787 (0.746–0.827)	0.640 (0.551–0.697)	0.844 (0.727–0.905)	48.350
AFP	0.654 (0.606–0.701)	0.584 (0.491–0.659)	0.658 (0.558–0.732)	13.850
GALAD	0.779 (0.739–0.819)	0.667 (0.566–0.723)	0.805 (0.654–0.861)	2.615
ASAP	0.813 (0.776–0.851)	0.644 (0.551–0.708)	0.853 (0.749–0.905)	1.059
GAAP	0.807 (0.770–0.845)	0.652 (0.569–0.712)	0.835 (0.736–0.900)	7.825
C_GALAD_II	0.769 (0.728–0.810)	0.547 (0.457–0.610)	0.883 (0.788–0.931)	1.952
GALAD_C	0.831 (0.795–0.867)	0.663 (0.532–0.723)	0.883 (0.797–0.922)	0.552
GAP_TALAD	0.874 (0.843–0.905)	0.820 (0.738–0.865)	0.823 (0.658–0.874)	0.773
LAD	0.801 (0.762–0.839)	0.667 (0.588–0.719)	0.835 (0.714–0.896)	0.491
C_GALAD	0.813 (0.776–0.850)	0.659 (0.577–0.723)	0.861 (0.775–0.918)	2.318
TAGALAD	0.880 (0.851–0.909)	0.760 (0.663–0.813)	0.861 (0.758–0.909)	0.525

Note: Cut-off values were determined using the Youden index. AUC, area under the curve.

ical and demographic subgroups: (1) patients with pathologically confirmed HCC, (2) patients with clinically diagnosed HCC, (3) patients with early-stage HCC (TNM I+II), (4) patients with complete data and no imputation, and (5) patients with HBV-related disease.

For the analyses of pathologically confirmed, clinically diagnosed, and early-stage HCC, each subgroup was compared with the entire control group ($n = 231$). The specific control groups for the complete-data and HBV-related analyses are outlined in their respective sections below. In all analyses, the TAGALAD and GAP_TALAD models

consistently outperformed the other models and individual biomarkers in diagnostic accuracy. The detailed ROC analysis results for each of these subgroups are presented in Tables 3,4,5,6,7 and Figs. 2,3,4,5,6, respectively.

ROC Curve Analysis of 89 Pathologically Confirmed HCC Patients

The sensitivity analysis was conducted on 89 pathologically confirmed HCC patients. The ROC curve analysis is presented in Fig. 2 and Table 3. Both the TAGALAD and GAP_TALAD models showed excellent diagnostic perfor-

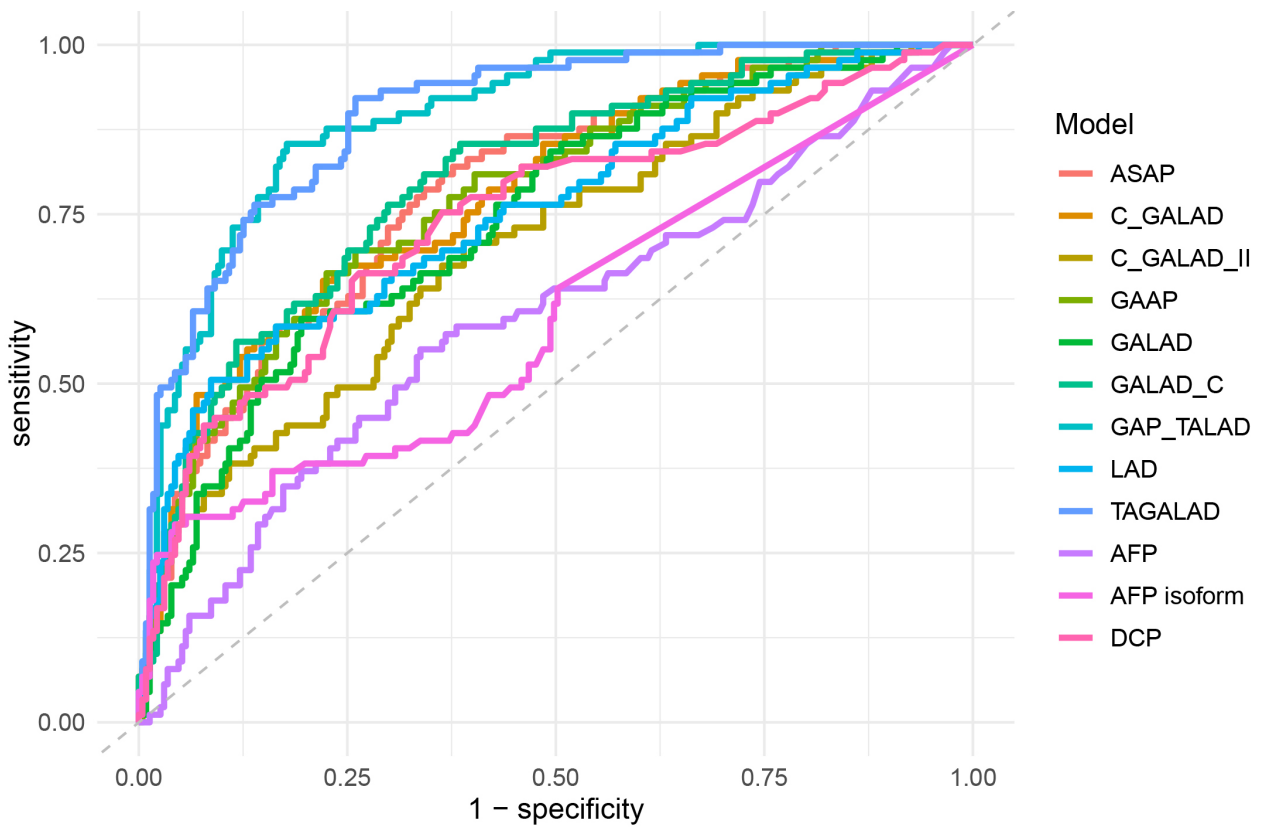


Fig. 2. ROC curves of the models and other individual indicators (subgroups diagnosed through pathological methods).

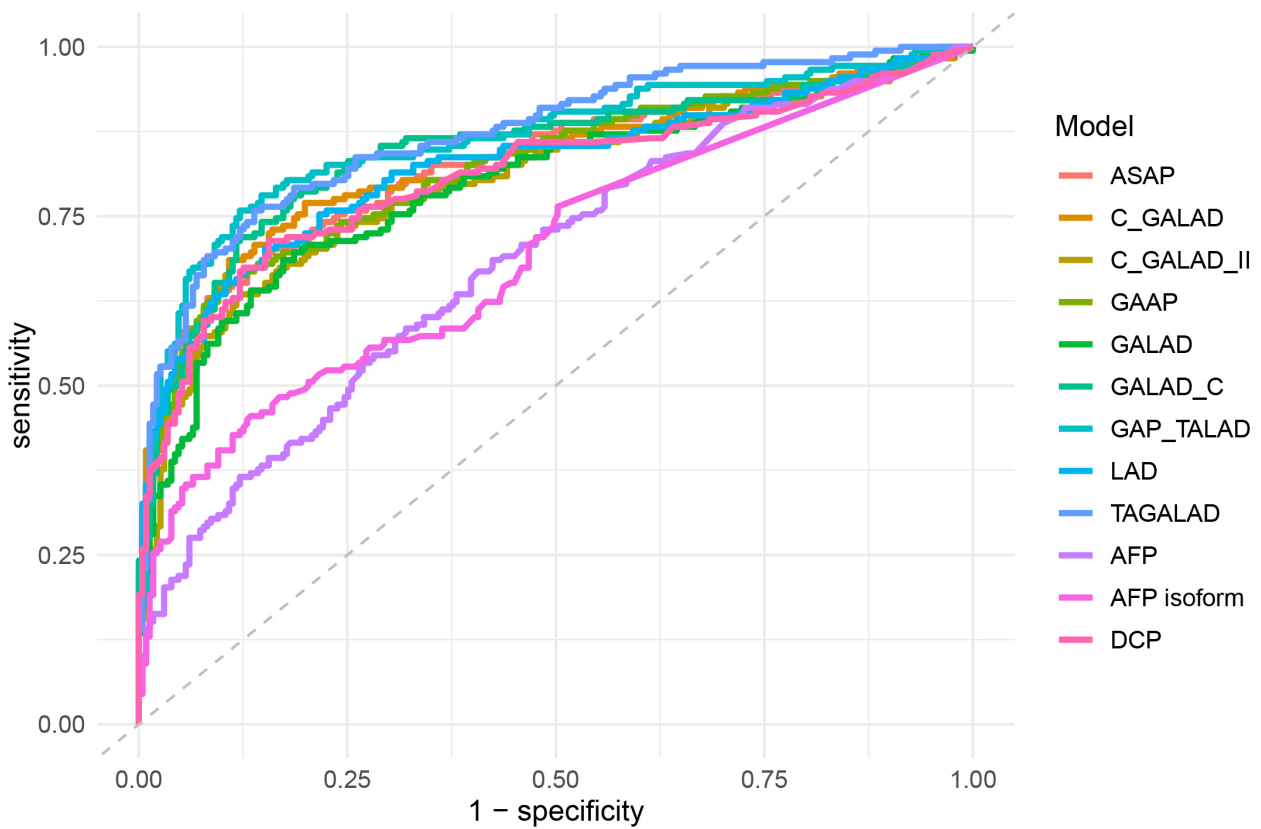


Fig. 3. ROC curves of the models and other individual indicators (clinically diagnosed subgroup).

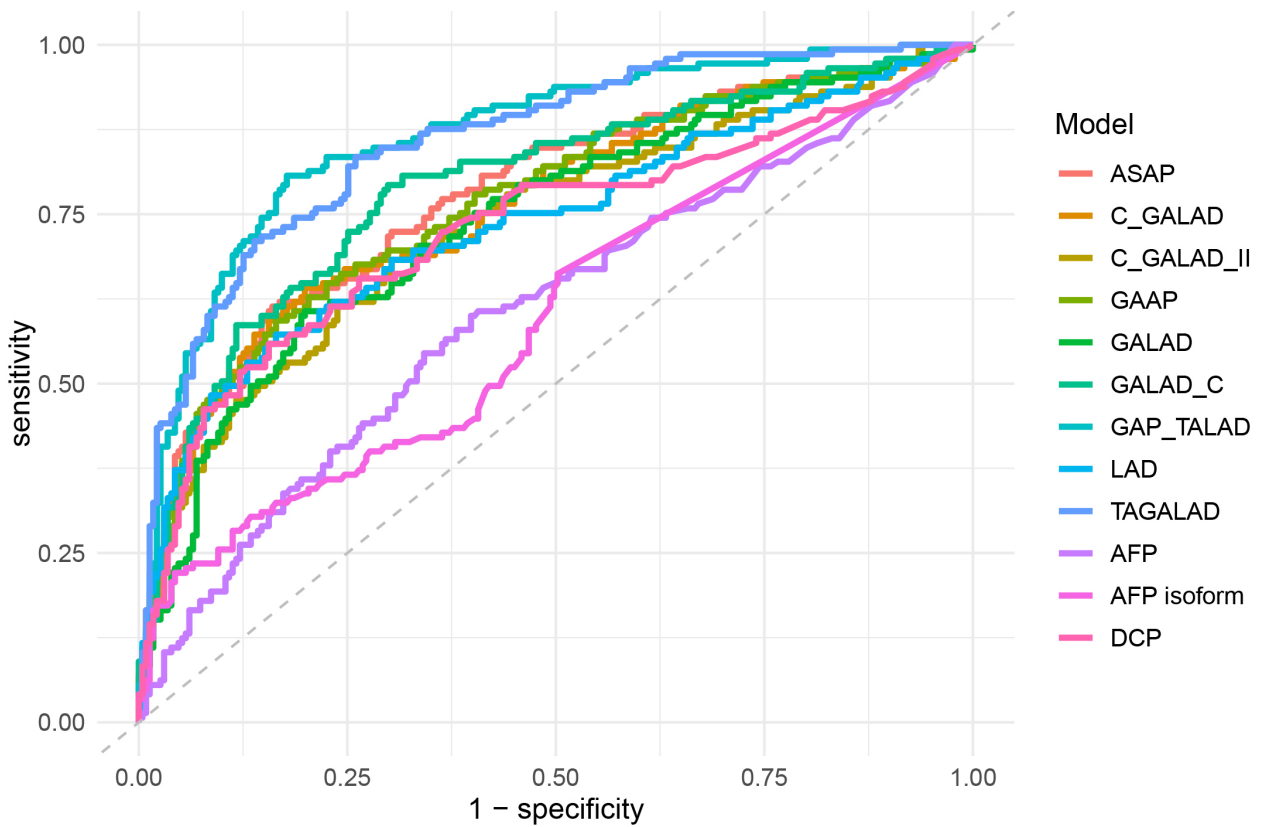


Fig. 4. ROC curves of the models and other individual indicators (early HCC subgroup).

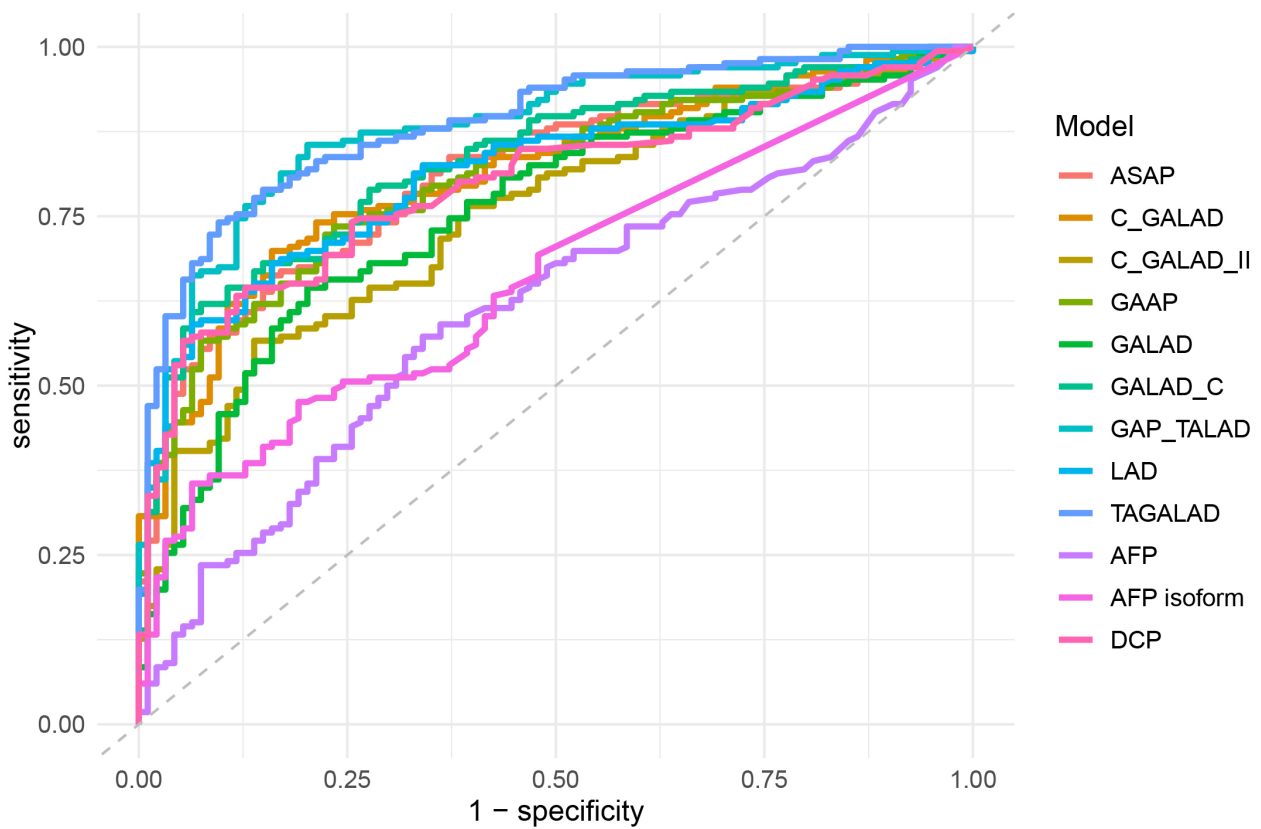


Fig. 5. ROC curves of the models and other individual indicators (no missing data subgroups).

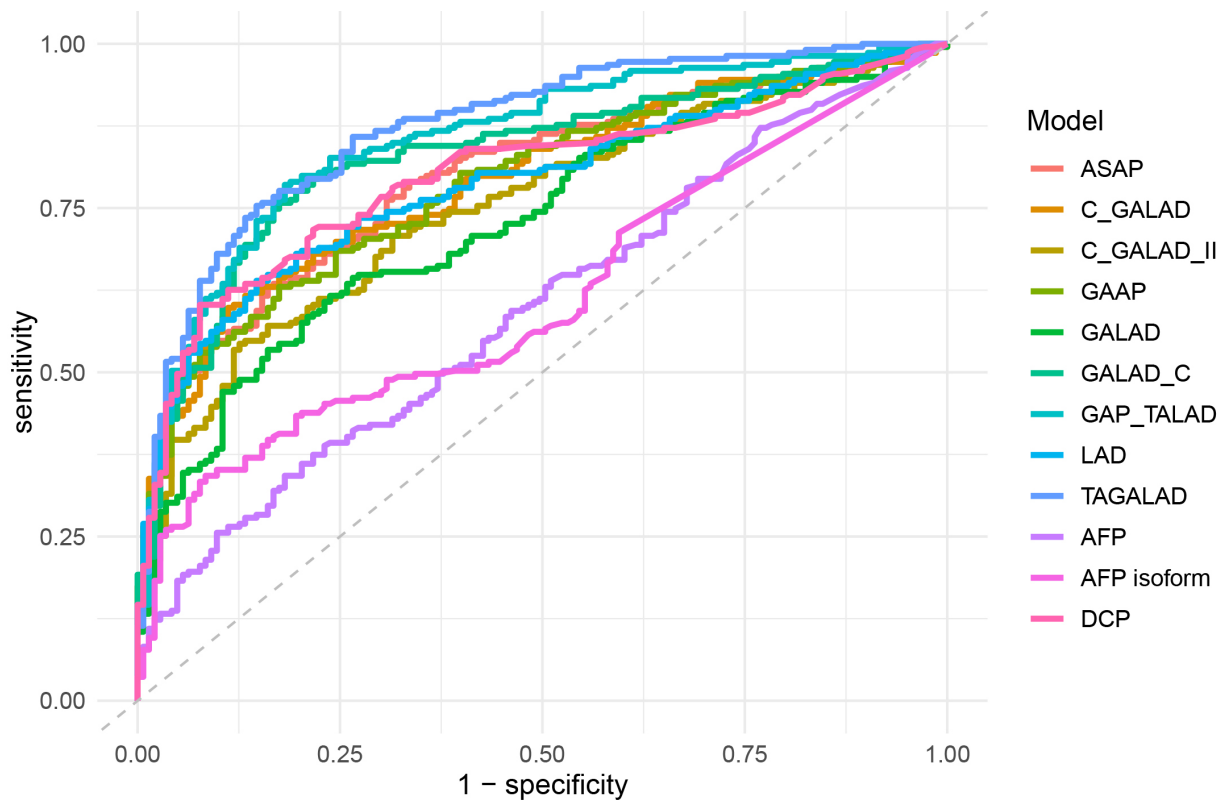


Fig. 6. ROC curves of the multi-biomarker diagnostic models and other individual indicators (chronic hepatitis B (CHB)-related HCC subgroup).

Table 3. ROC curve analysis results of the models and other individual indicators (subgroups diagnosed through pathological methods).

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.599 (0.528–0.670)	0.292 (0.180–0.382)	0.957 (0.736–0.991)	23.250
DCP	0.741 (0.677–0.805)	0.663 (0.506–0.753)	0.736 (0.602–0.805)	30.500
AFP	0.595 (0.523–0.667)	0.551 (0.393–0.652)	0.662 (0.446–0.749)	14.700
GALAD	0.744 (0.684–0.803)	0.596 (0.427–0.685)	0.805 (0.602–0.874)	2.616
ASAP	0.786 (0.731–0.840)	0.809 (0.697–0.899)	0.636 (0.463–0.732)	–0.056
GAAP	0.777 (0.721–0.834)	0.663 (0.517–0.764)	0.775 (0.619–0.853)	6.901
C_GALAD_II	0.704 (0.640–0.768)	0.708 (0.573–0.798)	0.610 (0.415–0.701)	0.006
GALAD_C	0.801 (0.748–0.855)	0.854 (0.742–0.921)	0.615 (0.394–0.714)	0.349
GAP_TALAD	0.894 (0.858–0.931)	0.854 (0.708–0.921)	0.823 (0.628–0.879)	0.773
LAD	0.755 (0.693–0.816)	0.584 (0.449–0.674)	0.835 (0.649–0.931)	0.491
C_GALAD	0.780 (0.724–0.837)	0.652 (0.517–0.753)	0.779 (0.593–0.879)	1.678
TAGALAD	0.897 (0.861–0.933)	0.921 (0.764–0.966)	0.740 (0.558–0.792)	0.397

Note: Cut-off values were determined using the Youden index.

mance and demonstrated the highest discriminative power among all models and individual biomarkers. Specifically, the ROC-AUC values for TAGALAD and GAP_TALAD were 0.897 (95% CI: 0.861–0.933) and 0.894 (95% CI: 0.858–0.931), respectively. These values were significantly higher than those of the GALAD model (AUC = 0.744, 95% CI: 0.684–0.803) and individual biomarkers, including AFP-L3 (AUC = 0.599, 95% CI: 0.528–0.670), DCP (AUC = 0.741, 95% CI: 0.677–0.805), and AFP (AUC = 0.595, 95% CI: 0.523–0.667).

The superior diagnostic accuracy of TAGALAD and GAP_TALAD in this subgroup highlights their robustness and reliability in distinguishing HCC cases, even within a cohort of patients definitively diagnosed through pathological methods. This finding underscores the potential of these models as valuable tools in clinical settings where precise diagnostic approaches are crucial.

Table 4. ROC curve analysis results of the models and other individual indicators (clinically diagnosed subgroup).

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.693 (0.642–0.744)	0.455 (0.360–0.537)	0.866 (0.740–0.926)	11.950
DCP	0.809 (0.764–0.855)	0.713 (0.612–0.770)	0.844 (0.680–0.905)	48.350
AFP	0.683 (0.631–0.735)	0.534 (0.410–0.624)	0.732 (0.628–0.797)	29.900
GALAD	0.797 (0.751–0.842)	0.685 (0.584–0.758)	0.827 (0.675–0.892)	2.917
ASAP	0.827 (0.785–0.870)	0.629 (0.522–0.697)	0.918 (0.827–0.965)	1.882
GAAP	0.822 (0.780–0.865)	0.652 (0.562–0.730)	0.892 (0.758–0.948)	9.167
C_GALAD_II	0.802 (0.756–0.847)	0.635 (0.534–0.708)	0.879 (0.775–0.931)	1.880
GALAD_C	0.845 (0.804–0.887)	0.719 (0.596–0.787)	0.879 (0.788–0.922)	0.541
GAP_TALAD	0.864 (0.826–0.902)	0.758 (0.663–0.826)	0.879 (0.766–0.926)	1.363
LAD	0.824 (0.780–0.867)	0.702 (0.612–0.759)	0.848 (0.740–0.918)	0.507
C_GALAD	0.829 (0.787–0.872)	0.685 (0.562–0.753)	0.892 (0.797–0.935)	2.896
TAGALAD	0.872 (0.837–0.907)	0.758 (0.668–0.815)	0.861 (0.736–0.922)	0.525

Note: Cut-off values were determined using the Youden index.

Table 5. ROC curve analysis results of the models and other individual indicators (early liver cancer subgroup).

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.597 (0.540–0.655)	0.221 (0.124–0.290)	0.957 (0.857–0.987)	23.250
DCP	0.728 (0.672–0.784)	0.559 (0.448–0.634)	0.844 (0.721–0.909)	48.350
AFP	0.606 (0.547–0.666)	0.545 (0.407–0.641)	0.658 (0.532–0.732)	13.850
GALAD	0.741 (0.689–0.794)	0.607 (0.476–0.683)	0.801 (0.641–0.861)	2.520
ASAP	0.777 (0.728–0.827)	0.614 (0.476–0.683)	0.840 (0.688–0.892)	0.935
GAAP	0.770 (0.719–0.820)	0.593 (0.483–0.683)	0.835 (0.727–0.901)	7.835
C_GALAD_II	0.737 (0.684–0.790)	0.683 (0.565–0.759)	0.697 (0.571–0.775)	0.515
GALAD_C	0.793 (0.745–0.842)	0.793 (0.676–0.855)	0.701 (0.498–0.771)	0.390
GAP_TALAD	0.867 (0.829–0.905)	0.807 (0.683–0.862)	0.823 (0.662–0.874)	0.773
LAD	0.737 (0.682–0.791)	0.572 (0.469–0.655)	0.835 (0.710–0.909)	0.492
C_GALAD	0.763 (0.712–0.814)	0.641 (0.538–0.717)	0.801 (0.589–0.879)	1.816
TAGALAD	0.860 (0.822–0.898)	0.834 (0.710–0.890)	0.740 (0.584–0.801)	0.398

Note: Cut-off values were determined using the Youden index.

Table 6. Analysis results of ROC curves of the models and other individual indicators (no missing data subgroups).

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.659 (0.596–0.723)	0.355 (0.211–0.428)	0.936 (0.798–0.979)	21.900
DCP	0.798 (0.745–0.852)	0.645 (0.536–0.717)	0.872 (0.702–0.957)	49.000
AFP	0.612 (0.542–0.682)	0.572 (0.404–0.663)	0.660 (0.489–0.755)	11.950
GALAD	0.756 (0.697–0.816)	0.645 (0.494–0.723)	0.798 (0.617–0.883)	2.722
ASAP	0.813 (0.761–0.864)	0.663 (0.536–0.741)	0.840 (0.670–0.915)	1.056
GAAP	0.807 (0.755–0.859)	0.735 (0.602–0.801)	0.766 (0.595–0.851)	6.725
C_GALAD_II	0.751 (0.692–0.810)	0.566 (0.386–0.639)	0.862 (0.681–0.915)	1.853
GALAD_C	0.831 (0.782–0.880)	0.620 (0.470–0.699)	0.926 (0.787–0.979)	0.714
GAP_TALAD	0.880 (0.838–0.922)	0.855 (0.735–0.904)	0.798 (0.532–0.883)	0.771
LAD	0.814 (0.763–0.865)	0.590 (0.422–0.669)	0.936 (0.819–0.979)	0.742
C_GALAD	0.812 (0.760–0.863)	0.699 (0.548–0.777)	0.840 (0.691–0.915)	2.167
TAGALAD	0.888 (0.849–0.928)	0.741 (0.614–0.825)	0.904 (0.787–0.968)	0.674

Note: Cut-off values were determined using the Youden index.

ROC Curve Analysis of 178 Clinically Diagnosed HCC Patients

A ROC curve analysis was conducted on 178 clinically diagnosed HCC patients to evaluate the diagnostic performance of various models and individual biomarkers. The results, shown in Fig. 3 and Table 4, demonstrated

that the TAGALAD and GAP_TALAD models showed superior diagnostic accuracy in this subgroup. The ROC-AUC values for TAGALAD and GAP_TALAD were 0.872 and 0.864, respectively, significantly outperforming the GALAD model and individual biomarkers, including AFP-L3, DCP, and AFP.

Table 7. ROC curve analysis results of the models and other individual indicators (CHB liver cancer subgroup).

Model	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Cut-off
AFP-L3	0.614 (0.558–0.670)	0.342 (0.242–0.406)	0.916 (0.797–0.965)	20.850
DCP	0.799 (0.754–0.845)	0.603 (0.447–0.667)	0.923 (0.797–0.965)	62.000
AFP	0.598 (0.539–0.657)	0.342 (0.237–0.425)	0.818 (0.706–0.888)	159.150
GALAD	0.731 (0.679–0.782)	0.616 (0.511–0.694)	0.762 (0.580–0.839)	2.917
ASAP	0.800 (0.755–0.844)	0.630 (0.521–0.703)	0.839 (0.720–0.909)	0.912
GAAP	0.789 (0.744–0.835)	0.630 (0.530–0.694)	0.825 (0.692–0.895)	7.825
C_GALAD_II	0.752 (0.702–0.801)	0.543 (0.402–0.612)	0.874 (0.741–0.923)	1.911
GALAD_C	0.831 (0.789–0.874)	0.758 (0.644–0.831)	0.825 (0.727–0.888)	0.428
GAP_TALAD	0.857 (0.818–0.896)	0.785 (0.662–0.845)	0.818 (0.685–0.881)	0.904
LAD	0.789 (0.743–0.835)	0.639 (0.534–0.717)	0.853 (0.734–0.916)	0.501
C_GALAD	0.791 (0.745–0.836)	0.598 (0.447–0.671)	0.888 (0.783–0.937)	3.016
TAGALAD	0.874 (0.838–0.910)	0.753 (0.644–0.817)	0.853 (0.741–0.909)	0.517

Note: Cut-off values were determined using the Youden index.

These findings further emphasize the robustness of TAGALAD and GAP_TALAD in clinical settings, where accurate diagnosis is critical. The superior performance of these models, compared to GALAD and single biomarkers, underscores their enhanced capability to integrate multiple factors, providing a more comprehensive assessment for distinguishing HCC from non-HCC conditions. This suggests that TAGALAD and GAP_TALAD offer substantial clinical utility in improving diagnostic precision in real-world scenarios involving clinically diagnosed HCC patients.

ROC Curve Analysis of 145 Early-Stage HCC (TNM Stage I+II) Patients

A ROC curve analysis was performed on 145 early-stage HCC patients (TNM stages I and II) to assess the diagnostic accuracy of various models and individual biomarkers. The results, presented in Fig. 4 and Table 5, demonstrated that both TAGALAD and GAP_TALAD models showed excellent diagnostic performance in this subgroup. The ROC-AUC values for TAGALAD and GAP_TALAD were 0.860 (95% CI: 0.822–0.898) and 0.867 (95% CI: 0.829–0.905), respectively, significantly higher than those of the GALAD model (AUC = 0.741, 95% CI: 0.689–0.794) and individual biomarkers, including AFP-L3 (AUC = 0.597, 95% CI: 0.540–0.655), DCP (AUC = 0.728, 95% CI: 0.672–0.784), and AFP (AUC = 0.606, 95% CI: 0.547–0.666).

The superior AUC values of TAGALAD and GAP_TALAD highlight their capability to accurately identify early-stage HCC patients, where prompt diagnosis is critical for effective intervention and improved prognoses. In addition to their strong diagnostic accuracy, the calibration analysis revealed that TAGALAD demonstrated optimal calibration performance, with a Brier score of 0.066 and a Hosmer-Lemeshow *p*-value of 0.532, indicating excellent agreement between predicted and observed probabilities. GAP_TALAD also showed robust calibration, with a Brier score of 0.070 and a Hosmer-Lemeshow *p*-value of 0.390, though slightly less accurate in extreme risk ranges.

In contrast, the GALAD model and individual biomarkers exhibited inferior diagnostic and calibration performance, with higher prediction errors and poorer calibration metrics. These findings underscore the clinical utility of TAGALAD and GAP_TALAD in detecting early-stage HCC, providing reliable and accurate predictions that can guide early diagnosis and management decisions.

ROC Curve Analysis of 167 HCC Patients With Complete Data and 95 Hepatitis + Cirrhosis Patients With Complete Data

A ROC curve analysis was conducted on 167 HCC patients and 95 control patients (with hepatitis + liver cirrhosis), all with complete data, to evaluate the diagnostic performance of various models and individual biomarkers. The results, shown in Fig. 5 and Table 6, revealed that TAGALAD and GAP_TALAD achieved the highest diagnostic accuracy among all models and biomarkers. Specifically, the ROC-AUC values for TAGALAD and GAP_TALAD were 0.888 and 0.880, respectively, significantly higher than those of the GALAD model and individual biomarkers, including AFP-L3, DCP, and AFP.

The exceptional diagnostic performance of TAGALAD and GAP_TALAD in this subgroup underscores their robustness and reliability, even in the absence of missing data. These models use a comprehensive array of clinical and biochemical parameters, achieving superior accuracy compared to GALAD, which demonstrated lower AUC values and suboptimal calibration performance. Similarly, individual biomarkers such as AFP-L3, DCP, and AFP demonstrated limited diagnostic utility, further emphasizing the advantages of multi-biomarker models like TAGALAD and GAP_TALAD.

These findings highlight the superior diagnostic capabilities of TAGALAD and GAP_TALAD when complete datasets are available. Their ability to consistently outperform traditional models and individual biomarkers reinforces their potential as reliable tools for accurately diagnosing HCC in diverse clinical settings.

ROC Curve Analysis of 219 HBV-Related HCC Patients and 143 HBV-Related Hepatitis + Cirrhosis Patients

A ROC curve analysis was performed on 219 HBV-related HCC patients and 143 HBV-related control patients (hepatitis + liver cirrhosis) to evaluate the diagnostic performance of various models and individual biomarkers. As shown in Fig. 6 and Table 7, the analysis revealed that TAGALAD and GAP_TALAD demonstrated the highest diagnostic accuracy among all models and individual biomarkers. Specifically, the ROC-AUC values for TAGALAD and GAP_TALAD were 0.874 and 0.857, respectively, both significantly outperforming the GALAD model and single biomarkers (AFP-L3, DCP, AFP).

TAGALAD achieved an excellent AUC of 0.874, with a sensitivity of 0.753 and a specificity of 0.853, highlighting its robust diagnostic performance in this HBV-related subgroup. GAP_TALAD followed closely with an AUC of 0.857, a sensitivity of 0.785, and a specificity of 0.818, further confirming its reliability for diagnosing HBV-related HCC. These results underscore the superior capability of TAGALAD and GAP_TALAD to integrate multiple biomarkers and clinical parameters, providing a more comprehensive assessment compared to traditional models.

In contrast, the GALAD model, while outperforming individual biomarkers, showed a lower AUC of 0.731, with moderate sensitivity (0.616) and specificity (0.762). Single biomarkers, including AFP-L3, DCP, and AFP, showed the weakest diagnostic performance, with AUC values of 0.614, 0.799, and 0.598, respectively. AFP and AFP-L3 were particularly limited by their low sensitivity, reducing their reliability in distinguishing HBV-related HCC from non-HCC conditions.

These findings highlight the diagnostic superiority of TAGALAD and GAP_TALAD in HBV-related HCC, where early and accurate diagnosis is critical. The integration of novel variables and advanced modeling in these frameworks significantly enhances their clinical utility, distinguishing them from GALAD model and traditional single biomarkers.

Discussion

Early diagnosis of HCC is critical for improving survival rates, as most patients are diagnosed at advanced stages when curative treatment options are limited. The current diagnostic approach relies on serum biomarkers such as AFP, DCP, and AFP-L3, along with imaging methods, including ultrasound, CT, and MRI. However, these methods have limitations in sensitivity and specificity, particularly in detecting early-stage HCC or differentiating it from benign liver conditions like hepatitis and cirrhosis [15]. These limitations are evident in our sensitivity analyses, where the diagnostic performance of a single marker like AFP-L3—including its optimal cut-off, sensitivity, and specificity—varied substantially across different patient subgroups (e.g.,

early-stage vs. clinically diagnosed HCC), as shown in Tables 2,3,4,5. Such variation is a well-known consequence of spectrum bias in diagnostic research, reflecting that the efficacy of a test is intrinsically linked to the case-mix and clinical characteristics of the study population. To overcome these challenges, multi-biomarker models such as TAGALAD and GAP_TALAD have emerged as promising tools for early and accurate HCC diagnosis, leveraging advanced statistical modeling to integrate clinical and biochemical parameters [16,17].

It is essential to clarify the framework of ‘impact’ as presented in this study. Our investigation focuses on the principle that accurate, early diagnosis is a crucial prerequisite and key determinant for surgical candidacy in HCC management. The term ‘impact on surgical decision-making’, thus, refers to the foundational effect of improving the diagnostic accuracy at this critical juncture. By demonstrating that advanced models like TAGALAD and GAP_TALAD can more reliably identify patients with early-stage disease, our research addresses the most significant bottleneck that currently limits the number of patients eligible for curative interventions. While this study does not measure downstream metrics like resection rates or survival—areas that remain crucial for future longitudinal research—improving this initial, rate-limiting diagnostic step constitutes a fundamental and direct impact on the entire surgical management process.

The superior performance of TAGALAD and GAP_TALAD likely arises from their advanced integration of multiple variables, offering a more comprehensive view of the host-tumor interaction. While the GALAD model serves as a foundational strategy, it is limited to tumor-derived glycoproteins (AFP, AFP-L3, DCP), which may not fully capture the biological complexity of the disease [18]. In contrast, TAGALAD incorporates albumin and bilirubin, which are not tumor-derived but indicate the liver’s deteriorating synthetic and excretory functions under tumor-induced stress. This inclusion provides an indirect measure of the tumor’s systemic metabolic effect, a critical aspect of its pathophysiology [19].

A high neutrophil-to-lymphocyte ratio (NLR) reflects a state of increased systemic inflammation and suppressed anti-tumor immunity, with its prognostic value as a key factor in the progression of HBV-related HCC now well-established [20]. This mechanistic distinction—from relying purely on tumor-derived markers to incorporating the host’s systemic response—likely accounts for the enhanced robustness of these models. This added complexity has significant clinical potential, as such models could be critical in refining current surveillance guidelines for high-risk populations [21]. Furthermore, they could serve as an objective method for risk stratification in clinical trials. However, to translate this potential into proven clinical benefit, a formal analysis of their utility is a valuable next step. Methods like decision curve analysis (DCA) are essential to precisely

quantify the net benefit of these models in specific clinical decision-making contexts and to guide their responsible implementation [22].

Several factors may influence the performance of diagnostic models, including sample size, patient heterogeneity, and the technologies used for biomarker assessment. In this study, patient cohorts were derived from a single center, which may limit generalizability. Additionally, the use of high-sensitivity immunoassays for biomarker detection likely contributed to the enhanced performance of TAGALAD and GAP_TALAD, but such methods may not be universally accessible [23]. Patient heterogeneity, such as differences in HCC etiology (e.g., HBV, HCV, non-alcoholic fatty liver disease [NAFLD]), poses another challenge, as biomarker levels and disease progression can vary significantly across populations [24]. Addressing these factors in future studies will be crucial for validating the clinical applicability of these models.

The findings of this study align with the broader trend of integrating multi-omics data into diagnostic models. For instance, emerging research has shown that incorporating genomic, proteomic, and metabolomic data can further enhance the diagnostic accuracy of HCC prediction models [25]. Future studies should explore the potential of combining such data with TAGALAD and GAP_TALAD to create even more robust diagnostic tools. Furthermore, longitudinal studies are warranted to evaluate the utility of these models in HCC screening programs and their ability to monitor disease progression or recurrence.

Beyond their analytical superiority, the robust performance of TAGALAD and GAP_TALAD holds significant clinical implications for the management of HCC. By enabling earlier and more accurate diagnosis, these multi-biomarker models are poised to directly impact surgical decision-making, expanding the cohort of patients identified at a curative stage and improving their eligibility for interventions such as surgical resection or liver transplantation. Furthermore, their enhanced precision can improve current surveillance programs for high-risk populations, enabling more effective patient stratification. This could further reduce unnecessary invasive procedures while accelerating diagnosis for those in need. Ultimately, integrating these models into clinical pathways holds the potential to significantly improve patient survival rates and optimize resource utilization within healthcare systems.

Limitations and Future Directions

Despite the promising results, this study has several limitations that should be acknowledged. Firstly, the retrospective, single-center design is a significant constraint. Our cohort is derived from a Chinese population, where HBV infection is the primary etiological factor for HCC, as reflected in our data. Consequently, while the high performance of our models is robust in this context, it may not be directly generalizable to other populations, particularly

those in Western countries, where HCC is more often driven by HCV or NAFLD. This etiological difference is crucial, as biomarker performance can vary significantly across different disease contexts. We acknowledge that external validation is essential. Therefore, our future work will prioritize the external validation of these models through a prospective, multi-center study. This future research will be intentionally designed to recruit diverse international cohorts, including patients from Western countries, to specifically assess the models' performance in populations with HCV- and NAFLD-predominant HCC. Such geographic and etiological validation is essential before these models can be recommended for widespread clinical adoption.

Secondly, while our overall sample size was adequate, the numbers within certain subgroups were too small for a more detailed analysis. This limitation restricted the assessment of the models' performance in less common HCC etiologies or specific patient strata, representing an area for future investigation with larger cohorts.

Additionally, the study intentionally focused on widely available and cost-effective biomarkers, excluding novel markers such as circulating tumor DNA (ctDNA), microRNAs, or extracellular vesicles [26]. These emerging markers offer the potential for highly-specific, non-invasive diagnosis by detecting tumor-specific genetic and epigenetic alterations [27]. For instance, ctDNA analysis can identify somatic mutations and aberrant methylation patterns unique to the tumor, providing a liquid biopsy that may complement or even surpass traditional serological tests in terms of specificity [28]. However, their clinical integration faces several challenges, including higher costs, the need for specialized infrastructure and bioinformatics expertise, and a lack of standardized protocols. Therefore, this study was designed to optimize the utility of biomarkers already integrated into existing diagnostic pathways. Looking forward, future research should focus on integrating these emerging omics-based markers into models like TAGALAD and GAP_TALAD. This approach could create a multi-tiered diagnostic strategy, where our models serve as robust initial screening tools, and more advanced tests like ctDNA analysis are employed for indeterminate cases or to monitor treatment response, thereby maximizing both diagnostic accuracy and resource utilization.

Conclusions

In summary, the findings highlight the superior diagnostic performance of TAGALAD and GAP_TALAD compared to GALAD and single biomarkers in detecting HCC across various patient subgroups. By integrating multiple biomarkers and clinical variables, these models offer a more comprehensive approach to HCC diagnosis. These models have the potential to transform HCC management by enabling early detection, improving diagnostic precision, and supporting personalized treatment strategies. However, further validation through larger, multi-center studies and

the exploration of novel biomarkers is necessary to elucidate their clinical potential.

Availability of Data and Materials

The data analyzed are available from the corresponding author upon reasonable request.

Author Contributions

XFZ and TW equally contributed to this research. XFZ, TW, XZX: investigation, methodology, data collection and analysis. XFZ, writing—original draft. GSG, MC, XJL: methodology, validation, formal analysis, visualization. DFM: conceptualization, supervision, writing—reviewing. All authors have been involved in revising it critically for important intellectual content. All authors gave final approval of the version to be published. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

The present study followed the Declaration of Helsinki. This study was approved by the Human Research Ethics Committee of Ningbo No. 2 Hospital (No. PJ-NBEY-KY-2024-087-01). The informed consent was waived.

Acknowledgment

Not applicable.

Funding

This research was supported by Zhejiang Province Clinical Key Specialty Construction Project (Grant No. 2024023), Ningbo Medical & Health Brand Discipline (Grant No. PPXK2024-04), Ningbo Health Technology Project (Grant No. 2022Y29), the Research Foundation of Ningbo No. 2 Hospital (Grant No. 2021HMKY10, 2023HMKY04), HwaMei Research Foundation of Ningbo No. 2 Hospital (Grant No. 2024HMKYA55).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at <https://doi.org/10.62713/ai.c.4182>.

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a Cancer Journal for Clinicians*. 2021; 71: 209–249. <https://doi.org/10.3322/caac.21660>.
- [2] Yang JD, Heimbach JK. New advances in the diagnosis and management of hepatocellular carcinoma. *BMJ (Clinical Research Ed.)*. 2020; 371: m3544. <https://doi.org/10.1136/bmj.m3544>.
- [3] Zheng RS, Zhang SW, Sun KX, Chen R, Wang SM, Li L, et al. Cancer statistics in China, 2016. *Zhonghua Zhong Liu Za Zhi [Chinese Journal of Oncology]*. 2023; 45: 212–220. <https://doi.org/10.3760/cma.j.cn112152-20220922-00647>. (In Chinese)
- [4] Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. *Nature Reviews. Clinical Oncology*. 2023; 20: 864–884. <https://doi.org/10.1038/s41571-023-00825-3>.
- [5] Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nature Reviews. Disease Primers*. 2021; 7: 6. <https://doi.org/10.1038/s41572-020-00240-3>.
- [6] Llovet JM, Pinyol R, Kelley RK, El-Khoueiry A, Reeves HL, Wang XW, et al. Molecular pathogenesis and systemic therapies for hepatocellular carcinoma. *Nature Cancer*. 2022; 3: 386–401. <https://doi.org/10.1038/s43018-022-00357-2>.
- [7] Li D, Mallory T, Satomura S. AFP-L3: a new generation of tumor marker for hepatocellular carcinoma. *Clinica Chimica Acta; International Journal of Clinical Chemistry*. 2001; 313: 15–19. [https://doi.org/10.1016/s0009-8981\(01\)00644-1](https://doi.org/10.1016/s0009-8981(01)00644-1).
- [8] Ganesan P, Kulik LM. Hepatocellular Carcinoma: New Developments. *Clinics in Liver Disease*. 2023; 27: 85–102. <https://doi.org/10.1016/j.cld.2022.08.004>.
- [9] Yu J, Park R, Kim R. Promising Novel Biomarkers for Hepatocellular Carcinoma: Diagnostic and Prognostic Insights. *Journal of Hepatocellular Carcinoma*. 2023; 10: 1105–1127. <https://doi.org/10.2147/JHC.S341195>.
- [10] Cheng JS, Gholami P, Kappus MT, Kwok RM, Nguyen P, Saab S. Evaluation of Diagnostic Performance of Alpha-FetoProtein (AFP), AFP-L3 and Des-Gamma-Carboxy Prothrombin (DCP) for HCV and Non-HCV Related Hepatocellular Carcinoma (HCC) After Long-Term Follow-Up. *American Journal of Gastroenterology*. 2012; 107: S145–S146.
- [11] Nguyen HB, Le XTT, Nguyen HH, Vo TT, Le MK, Nguyen NT, et al. Diagnostic Value of hTERT mRNA and in Combination With AFP, AFP-L3%, Des-γ-carboxyprothrombin for Screening of Hepatocellular Carcinoma in Liver Cirrhosis Patients HBV or HCV-Related. *Cancer Informatics*. 2022; 21: 11769351221100730. <https://doi.org/10.1177/11769351221100730>.
- [12] Sherman M. Hepatocellular carcinoma: epidemiology, surveillance, and diagnosis. *Seminars in Liver Disease*. 2010; 30: 3–16. <https://doi.org/10.1055/s-0030-1247128>.
- [13] Balcar L, Mrekva A, Scheiner B, Pomej K, Meischl T, Mandorfer M, et al. Management of varices but not anticoagulation is associated with improved outcome in patients with HCC and macrovascular tumour invasion. *Cancer Imaging: the Official Publication of the International Cancer Imaging Society*. 2024; 24: 9. <https://doi.org/10.1186/s40644-024-00657-z>.
- [14] Johnson PJ, Pirrie SJ, Cox TF, Berhane S, Teng M, Palmer D, et al. The detection of hepatocellular carcinoma using a prospectively developed and validated model based on serological biomarkers. *Cancer Epidemiology, Biomarkers & Prevention: a Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*. 2014; 23: 144–153. <https://doi.org/10.1158/1055-9965.EPI-13-0870>.
- [15] El-Serag HB, Kanwal F. Epidemiology of hepatocellular carcinoma in the United States: where are we? Where do we go? *Hepatology (Baltimore, Md.)*. 2014; 60: 1767–1775. <https://doi.org/10.1002/hep.27222>.
- [16] Luo P, Yin P, Hua R, Tan Y, Li Z, Qiu G, et al. A Large-scale, multi-center serum metabolite biomarker identification study for the early detection of hepatocellular carcinoma. *Hepatology (Baltimore, Md.)*. 2018; 67: 662–675. <https://doi.org/10.1002/hep.29561>.
- [17] Best J, Bechmann LP, Sowa JP, Sydor S, Dechêne A, Pflanz K, et al. GALAD Score Detects Early Hepatocellular Carcinoma in an International Cohort of Patients With Nonalcoholic Steatohepatitis. *Clin-*

- ical Gastroenterology and Hepatology: the Official Clinical Practice Journal of the American Gastroenterological Association. 2020; 18: 728–735.e4. <https://doi.org/10.1016/j.cgh.2019.11.012>.
- [18] Marasco G, Dajti E, Serenari M, Alemanni LV, Ravaioli F, Ravaioli M, *et al.* Sarcopenia Predicts Major Complications after Resection for Primary Hepatocellular Carcinoma in Compensated Cirrhosis. *Cancers*. 2022; 14: 1935. <https://doi.org/10.3390/cancers14081935>.
- [19] Xia J, Wang C, Li B. Hepatocellular carcinoma cells induce $\gamma\delta$ T cells through metabolic reprogramming into tumor-progressive subpopulation. *Frontiers in Oncology*. 2024; 14: 1451650. <https://doi.org/10.3389/fonc.2024.1451650>.
- [20] Xu C, Wu F, Du L, Dong Y, Lin S. Significant association between high neutrophil-lymphocyte ratio and poor prognosis in patients with hepatocellular carcinoma: a systematic review and meta-analysis. *Frontiers in Immunology*. 2023; 14: 1211399. <https://doi.org/10.3389/fimmu.2023.1211399>.
- [21] Braillon A. Monitoring of Hepatocellular Carcinoma ... Surveillance: Reaching the Horizon. *Clinical Gastroenterology and Hepatology: the Official Clinical Practice Journal of the American Gastroenterological Association*. 2024; 22: 1340. <https://doi.org/10.1016/j.cgh.2023.09.039>.
- [22] Vickers AJ, Van Claster B, Wynants L, Steyerberg EW. Decision curve analysis: confidence intervals and hypothesis testing for net benefit. *Diagnostic and Prognostic Research*. 2023; 7: 11. <https://doi.org/10.1186/s41512-023-00148-y>.
- [23] European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *Journal of Hepatology*. 2018; 69: 182–236. <https://doi.org/10.1016/j.jhep.2018.03.019>.
- [24] Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, *et al.* Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *The New England Journal of Medicine*. 2020; 382: 1894–1905. <https://doi.org/10.1056/NEJMoa1915745>.
- [25] Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. *Nature Reviews. Gastroenterology & Hepatology*. 2020; 17: 139–152. <https://doi.org/10.1038/s41575-019-0229-4>.
- [26] Wong CM, Tsang FHC, Ng IOL. Non-coding RNAs in hepatocellular carcinoma: molecular functions and pathological implications. *Nature Reviews. Gastroenterology & Hepatology*. 2018; 15: 137–151. <https://doi.org/10.1038/nrgastro.2017.169>.
- [27] Loosen SH, Roderburg C, Kauertz KL, Pombeiro I, Leyh C, Benz F, *et al.* Elevated levels of circulating osteopontin are associated with a poor survival after resection of cholangiocarcinoma. *Journal of Hepatology*. 2017; 67: 749–757. <https://doi.org/10.1016/j.jhep.2017.06.020>.
- [28] Oversoe SK, Clement MS, Weber B, Grønbaek H, Hamilton-Dutoit SJ, Sorensen BS, *et al.* Combining tissue and circulating tumor DNA increases the detection rate of a CTNNB1 mutation in hepatocellular carcinoma. *BMC Cancer*. 2021; 21: 376. <https://doi.org/10.1186/s12885-021-08103-0>.

© 2025 The Author(s).

