OPEN

Evaluating the GALAD Score for Detection of Hepatocellular Carcinoma in Patients With Cirrhosis

Thong Duy Vo, MD, PhD,† Sang Hoai Mai, MD, MSc,‡ and Huong Tu Lam, MD, MSc†*

Introduction: Early diagnosis of hepatocellular carcinoma (HCC) is crucial but challenging, and late detection limits its treatment and prognosis. We aimed to evaluate the GALAD score as a novel yet highly accurate and promising diagnostic tool for HCC.

Methods: A prospective and retrospective cohort study was conducted in 196 adult patients with cirrhosis, including 102 with HCC and 94 without. The diagnostic performance of the GALAD score for HCC detection was compared with that of single biomarkers.

Results: In patients with cirrhosis with HCC, the GALAD score was 2.5 (95% CI: −2.43 to 11.09), which was significantly higher than the GALAD score of −2.46 (95% CI: −6.15 to 2.04) in patients with cirrhosis without HCC ($P<0.001$). Patients with multiple tumors had a significantly higher GALAD score than those with a single tumor ($P=0.0081$). There was a moderate correlation between the GALAD score and tumor size in patients with cirrhosis ($r=0.44$; $P<0.001$). The GALAD score had an area under the receiver operating characteristic curve of 0.91, higher than that of all single biomarkers used to diagnose HCC (all $P<0.001$). The optimal cutoff for diagnosing HCC using the GALAD score was −0.518, achieving a sensitivity of 87.25%, specificity of 82.98%, positive predictive value of 84.62%, and negative predictive value of 84.78%. At this cutoff, the GALAD score demonstrated superior sensitivity compared with single or combined biomarkers.

Conclusions: The GALAD score shows promise in detecting HCC in patients with cirrhosis. The GALAD score could be applied in clinical practice to diagnose HCC in patients with cirrhosis, and calculating the GALAD score in clinical settings may help predict tumor size and quantity before imaging results become available.

Received for publication May 20, 2024; accepted October 7, 2024.
From the †Department of Internal Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City; *Department of Gastroenterology, University Medical Center HCMC; and ‡Department of Gastroenterology, Kien Giang General Hospital, Rach Gia City, Vietnam.
The study protocol was reviewed and approved by the Ethics Committee of the University of Medicine and Pharmacy at Ho Chi Minh City (approval number: 697/HDDD-DHYD).
S.H.M. and T.D.V.: developed the concept of the study, participated in the study design, and collected, analyzed, and interpreted the data. S.H.M., H.T.L., and T.D.V.: wrote the manuscript. T.D.V.: supervised the study and revised the manuscript.
The authors declare that they have nothing to disclose.
Address correspondence to: Thong Duy Vo, MD, PhD, Department of Internal Medicine, Faculty of Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam. Department of Gastroenterology, University Medical Center HCMC, Ho Chi Minh City, Vietnam. (e-mail: thong.vd@umc.edu.vn, duythong@ump.edu.vn).

DOI: 10.1097/MCG.0000000000002097

Key Words: GALAD score, hepatocellular carcinoma, early detection

(*J Clin Gastroenterol* 2025;59:915–919)

Primary liver cancer was the sixth most commonly diagnosed cancer and the third leading cause of cancer-related deaths worldwide in 2020, with approximately 906,000 new cases and 830,000 deaths. Hepatocellular carcinoma (HCC) is the most common primary malignancy of the liver, accounting for 75% to 85% of cases.[1] However, the early diagnosis of HCC is crucial but challenging, and late detection limits its treatment options and prognosis.[2] Therefore, the goal of HCC surveillance programs is to detect HCC in its early stages when treatment is feasible. Despite western guidelines favoring ultrasound alone, studies have revealed its limited sensitivity (~63%) in detecting HCC, which is affected by imaging quality, sonographer expertize, and patient factors such as obesity. To improve surveillance, the American Association for the Study of Liver Diseases recommends HCC surveillance using a combination of liver ultrasonography and alpha-fetoprotein (AFP). However, combining AFP with ultrasound is considered insufficient for HCC surveillance because this combination only increases the detection rate by ~6% to 8%.[3,4]

The GALAD score, developed in 2014 and based on sex, age, and 3 biomarkers [AFP, AFP-L3, and PIVKA-II, also known as Des-gamma-carboxy prothrombin], is a promising tool that has shown good diagnostic ability for HCC through various studies.[5] A recent case study in patients with metabolic-associated steatohepatitis[6] reported that the GALAD score identified patients with HCC of any stage with an area under the receiver operating characteristic curve (AUROC) of 0.96, which was significantly greater than the values for serum levels of AFP (AUROC, 0.88), AFP-L3 (AUROC, 0.86), and PIVKA-II (AUROC, 0.87). The study also concluded that the GALAD score could help detect patients with early-stage HCC and might facilitate the surveillance of patients with metabolic-associated steatohepatitis, who often present with obesity and in whom HCC is difficult to identify on ultrasound. Another study by Best et al[7] compared AFP, AFP-L3, and PIVKA-II individually with GALAD in early HCC diagnosis in patients with chronic liver disease and found that GALAD had the highest sensitivity and specificity at 93.3% and 85.6%, respectively, with an AUROC of 0.92. Another study by Yang et al,[8] comparing GALAD with ultrasound in early HCC diagnosis revealed the AUROC of GALAD to be 0.95 (95% CI: 0.93-0.97), higher than that of ultrasound (0.82 with $P<0.01$). At a cutoff point of −0.76, GALAD exhibited a sensitivity of 91% and a specificity of

85% in HCC diagnosis. The AUROC of GALAD in early-stage HCC diagnosis was 0.92 (95% CI: 0.88-0.96) at a cutoff point of −1.18, with a sensitivity of 92% and a specificity of 79%. The combination of GALAD and ultrasound (GALADUS) yielded superior results with an AUROC of 0.98 (95% CI: 0.96-0.99) at a cutoff point of −0.18, demonstrating a sensitivity of 95% and specificity of 91%.

With the potential benefits of the GALAD score, we aimed to validate the GALAD score in Vietnamese patients with cirrhosis, to contribute more evidence to support its use in HCC surveillance within this specific population.

METHODS

A prospective and retrospective cohort study was conducted in adult patients with cirrhosis at the Liver Clinic and Hepatobiliary Tumor-Liver Transplantation Clinic, University Medical Center HCMC, from January 2018 to May 2021.

Study Population

The inclusion criteria were patients aged 18 years or older who had the following:[9,10]

- A diagnosis of cirrhosis confirmed by liver biopsy or noninvasive fibrosis assessment methods such as (1) APRI $\geq$ 2, (2) Fibroscan $\geq$ 12.5 to 14 kPa, or (3) MRE $\geq$ 5.9 kPa
- Dynamic contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI)
- Measured levels of AFP, AFP-L3, and PIVKA-II

Patients were excluded if they had been diagnosed with any other malignant disease besides HCC or if they had previously been diagnosed and treated for HCC.

The diagnosis of HCC was confirmed by histopathology or dynamic contrast-enhanced CT/MRI (using gadolinium-based contrast agents or Primovist, a hepatocyte-specific contrast agent). Typical imaging characteristics of HCC require the presence of both arterial phase enhancement and "washout" at portal venous and/or delayed phases.[11,12] If the initial imaging modality did not yield typical results, a secondary imaging modality was evaluated.

The formulas used for the calculations were as follows:[7]

$\text{GALAD} = -10.08 + 0.09 \times \text{age} + 1.67 \times \text{sex} + 2.34 \log_{10}(\text{AFP}) + 0.04 \times \text{AFP-L3} + 1.33 \times \log_{10}(\text{PIVKA-II})$. Sex was set to 1 for females and 0 for males.

Study Protocol

We screened patients meeting the time and location criteria on a computer system for diagnosed or confirmed cirrhosis. Subsequently, we proceeded with further screening to select patients with abdominal CT and/or MRI results for diagnosing focal liver lesions, accompanied by quantitative testing of AFP, AFP-L3, and PIVKA-II performed within 3 months of the CT or MRI scan. We then compiled the necessary data from the population that met the inclusion criteria and lacked the exclusion criteria.

Statistical Analysis

Data analysis was conducted using Stata 16.0 software. Continuous variables with a normal distribution were expressed as mean ± SD, while those without a normal distribution were presented as median and interquartile range (IQR). Categorical variables were presented as counts and percentages. Comparison of mean scores for normally distributed quantitative variables was performed using the Student *t* test and ANOVA. The nonparametric Wilcoxon test was used to analyze non-normally distributed quantitative variables. The Shapiro-Wilk test was used to assess whether the quantitative variables adhered to a normal distribution. The diagnostic performance assessment of biomarkers or the GALAD score involved the use of receiver operating characteristic curves and the calculation of the AUROC. The optimal cutoff values, sensitivity, and specificity were determined based on the highest Youden index. Statistical significance was defined as a *P*-value < 0.05.

RESULTS

Patient Characteristics

In total, 196 patients were recruited, including 102 with HCC and 94 without. Characteristics of the study population are summarized in Table 1. The HCC group showed a higher median AFP [15748.29 ng/mL (IQR: 2.2 to 3748) vs. 23.26 ng/mL (IQR: 1.1 to 111.1), $P < 0.001$], higher mean AFP-L3 [20.15% (SD: 0.5 to 71.9) vs. 3.48% (SD: 0.5 to 17.4), $P < 0.001$], higher mean PIVKA-II [223.49 (SD: 0.21 to 2088.97) vs. 35.88 (SD: 0.17 to 2.25), $P < 0.001$], and higher GALAD score [2.5 (SD: −2.43 to 11.09) vs. −2.46 (SD: −6.15 to 2.04), $P < 0.001$] compared with the non-HCC group.

The Correlation Between the GALAD Score and Tumor Characteristics

Regarding tumor size, our study established a correlation between tumor size and GALAD score. Tumor size refers to the size of a single lesion, not the cumulative size of

TABLE 1. Demographics and Characteristics of non-HCC and HCC Groups

Parameter	Non-HCC (n = 94)	HCC (n = 102)	*P*
Age, mean (SD), y	59.5 (12.6)	61.9 (11.1)	0.17
Male, n (%)	51 (54.3)	67 (65.7)	0.11
Etiology, %			
Hepatitis B virus	43	49	> 0.05
Hepatitis C virus	22	27	—
Nonalcoholic steatohepatitis	3	2	—
Alcohol	5	5	—
Other	20	18	—
Child-Pugh class, %			
Stage A	85	76	> 0.05
Stage B	12	18	—
Stage C	2	6	—
Thrombocytes, k/μL	156.2 (73.2)	155.2 (82.8)	0.8
eGFR (mL/min/1.73 m^2)	80 (18.35)	80.05 (17.15)	0.9
AFP, median (IQR), ng/mL	23.26 (1.1-111.1)	15748.29 (2.2-3748)	< 0.001
AFP-L3, %	3.48 (0.5-17.4)	20.15 (0.5-71.9)	< 0.001
PIVKA II (mAU/mL)	35.88 (0.17-2.25)	223.49 (0.21-2088.97)	< 0.001
GALAD	−2.46 (−6.15 to 2.04)	2.5 (−2.43 to 11.09)	< 0.001

AFP indicates alpha-fetoprotein; AFP-L3, lens culinaris agglutinin A-reactive fraction of alpha-fetoprotein; eGFR, estimated glomerular filtration rate; HCC, hepatocellular carcinoma; IQR, interquartile range; PIVKA-II, protein induced by vitamin K absence or antagonist II.

multiple lesions. Specifically, we observed that among patients with tumor sizes >3 cm and up to 5 cm, as well as exceeding 5 cm, the GALAD score tended to be higher than that of patients with tumor sizes equal to or <2 cm. By constructing a linear regression model between the GALAD score and tumor size, we obtained an r-value of 0.444 with $P<0.001$, indicating a moderate correlation between the GALAD score and tumor size using the following equation: tumor size = 3.537 + 0.444 × GALAD score. This finding may partially facilitate the clinical prediction of tumor size based on the GALAD score.

Our study demonstrated a statistically significant correlation between the number of tumors and the GALAD score ($P=0.0081$). In terms of tumor location, our study revealed no correlation with the GALAD score. The correlation between the GALAD score and tumor characteristics is shown in Table 2.

Performance of Single Biomarkers for HCC Detection

AFP, at a cutoff of 10 ng/mL, demonstrated a moderate sensitivity of 71.57% and specificity of 78.72%, with a relatively balanced positive predictive value (PPV) and negative predictive value (NPV), and an AUROC of 0.75 (95% CI: 1.58-2.89). However, at a higher cutoff of 20 ng/mL, sensitivity decreased to 52.94% while specificity increased to 88.3%, leading to a higher PPV but a lower NPV, with an AUROC of 0.71 (95% CI: 1.4-2.88). The AFP-L3 test, recommended at a cutoff of 10%, showed a lower sensitivity of 48.04% compared with AFP but higher specificity of 92.55%, resulting in a higher PPV and similar NPV, with an AUROC of 0.70 (95% CI: 1.58-3.3). PIVKA-II, at a cutoff of 7.5 ng/mL, exhibited the lowest sensitivity of 45.1% but notably the highest specificity of 96.81%, leading to a higher PPV and comparable NPV, with an AUROC of 0.71 (95% CI: 2-4.3). A comparison of the performance of single biomarkers for HCC diagnosis is shown in Figure 1.

Performance of the GALAD Score for HCC Detection

In our study, the AUROC of the GALAD score was 0.91 (95% CI: 0.86-0.94). The diagnostic threshold selected with the highest Youden index was chosen as GALAD ≥ −0.518, with a sensitivity of 87.25% and specificity of 82.98% (Table 3).

TABLE 2. Correlation Between the GALAD Index and Tumor Characteristics

Tumor characteristics	The GALAD score	ANOVA test
Tumor size		
≤ 2 cm	1.24 ± 2.92	F = 10.1 $P<0.001$
> 2 and ≤ 3 cm	0.33 ± 2.36	
> 3 and ≤ 5 cm	3.17 ± 3.26	
> 5 cm	4.81 ± 3.85	
No. tumors		
1	1.84 ± 2.9	F = 7.29 $P=0.0081$
≥ 2	3.81 ± 4.43	
Tumor location		
Left lobe	1.86 ± 3.49	F = 1.88 $P=0.16$
Right lobe	2.19 ± 3.16	

Tumor size represents the size of a single lesion.

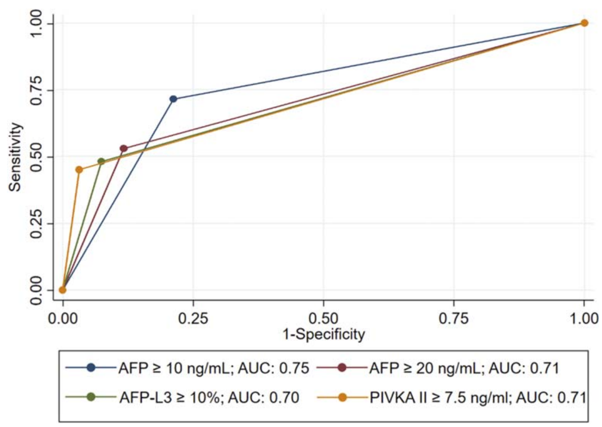

FIGURE 1. Comparison of the performance of single biomarkers for the diagnosis of hepatocellular carcinoma.

Comparison of the Performance of the GALAD Score and Single Biomarkers in HCC Detection

The accuracy of HCC diagnosis using single biomarkers and the GALAD score is shown in Table 4. When comparing the AUROC between the single biomarkers and the GALAD score, the highest AUROC for the GALAD score (0.91) was statistically significant compared with that for all single biomarkers (all $P<0.001$).

DISCUSSION

Patient Characteristics

This study focused on patients with cirrhosis, as they constitute a high-risk population requiring HCC surveillance. Consequently, serum tumor biomarkers offer a potentially more convenient, cost-effective, objective, and reproducible method for HCC screening. Among these, AFP, AFP-L3, and PIVKA-II have been clinically used as serum biomarkers for the early detection and diagnosis of HCC. Similar to previous research, our study demonstrated that the levels of individual biomarkers and the GALAD score were significantly higher in the HCC group than in the non-HCC group. This finding is consistent with the prior utilization of these biomarkers to distinguish between HCC and non-HCC conditions in patients with cirrhosis.

TABLE 3. Sensitivity, Specificity, and Youden Index of the GALAD Score at Different Cutoffs

Cutoff value	Sensitivity	Specificity	Youden index
≥ −0.790	88.24	79.79	0.68
≥ −0.780	87.25	79.79	0.67
≥ −0.742	87.25	80.85	0.68
≥ −0.616	87.25	81.91	0.69
≥ −0.518	87.25	82.98	0.70
≥ −0.480	86.27	82.98	0.69
≥ −0.464	85.29	82.98	0.68
≥ −0.409	84.31	82.98	0.67
≥ −0.397	82.35	82.98	0.65
≥ −0.328	81.37	82.98	0.64

TABLE 4. Accuracy of HCC Diagnosis Using the GALAD Score, in Comparison to That Using Single Biomarkers

	Cutoff value	AUC	*P*-value AUC (vs. GALAD score)	Sens, %	Spec, %	PPV, %	NPV, %
GALAD	−0.518	0.91	—	86.27	82.98	84.62	84.78
AFP	10 ng/mL	0.75	10 ng/mL	71.57	78.72	78.49	71.84
AFP	20 ng/mL	0.71	20 ng/mL	52.94	88.3	88	63.36
AFP-L3	10%	0.70	10%	48.04	92.55	87.50	62.14
PIVKA II	*7.5* ng/mL	0.71	*7.5* ng/mL	45.10	96.81	93.88	61.90

AFP indicates alpha-fetoprotein; AFP-L3, lens culinaris agglutinin A-reactive fraction of alpha-fetoprotein; HCC, hepatocellular carcinoma; NPV, negative predictive value; PIVKA-II, protein induced by vitamin K absence or antagonist II; PPV, positive predictive value; sens, sensitivity; spec, specificity.

The Correlation Between the GALAD Score and Tumor Characteristics

Previous studies have shown inconsistent results regarding the correlation between AFP and AFP-L3 and tumor size. Some studies have reported a linear correlation between PIVKA-II and tumor size. For instance, Yamamoto et al[13] reported a strong correlation between PIVKA-II and tumor size (r-value 0.66 and $P<0.001$). Notably, while domestic and international studies have extensively explored the correlation between AFP, AFP-L3, and PIVKA-II levels and tumor size, there has been limited investigation into the correlation between the GALAD score and tumor size. In our study, we found a significant correlation between the GALAD score and tumor size ($P<0.001$). This suggests that the GALAD score may partially predict tumor size clinically.

In our study, we did not find any correlation between the GALAD score and tumor location. However, we observed a correlation between the GALAD score and the number of tumors. Specifically, the multiple tumor group exhibited a higher GALAD score than the single tumor group ($P=0.0081$). This suggests that, clinically, the GALAD score may offer guidance on the number of tumors in patients with cirrhosis with HCC, assisting clinicians in treatment decisions and patient prognosis. To the best of our knowledge, no domestic or international studies have investigated this correlation. Nevertheless, some studies have failed to find a correlation between AFP, AFP-L3, and PIVKA-II and tumor location and number.

Performance of the GALAD Score for HCC Detection

In our study, we identified a cutoff point for the GALAD score at −0.518, which yielded the highest Youden score. At this cutoff point, the sensitivity was 86.27%, specificity was 82.98%, PPV was 84.62%, and NPV was 84.78%. Compared with single biomarkers, the GALAD score demonstrated superiority, as recommended by some guidelines.

For instance, at a cutoff point of AFP ≥ 10 ng/mL, the sensitivity was 71.57%, and the specificity was 78.72%. AFP-L3 and PIVKA-II are known for their low sensitivity and high specificity in the diagnosis of HCC. At a cutoff point of 10% for AFP-L3, the sensitivity and specificity were 48.04% and 92.55%, respectively. For PIVKA-II, at a cutoff point of 7.5 ng/mL, the sensitivity was 45.1%, whereas the specificity was relatively high at 96.81%. Our study results were consistent with those of Berhane et al[14] in a UK population, where a cutoff point of GALAD at −0.63 showed a sensitivity of 91.6% and specificity of 89.75%. Similarly, in the German population, a GALAD cutoff point of GALAD at −0.68 yielded a sensitivity of 88.4% and specificity of 88.2%. However, in the Japanese population, a lower cutoff point of −1.95 showed a sensitivity of 81.4% and a specificity of 89.1%. Furthermore, Best et al[7] reported a sensitivity of 85.6% and specificity of 93.3% at a GALAD cutoff point of −0.63. These findings underscore the consistency of GALAD score performance across different populations and its superiority in diagnosing HCC.

In our study, the AUROC of the GALAD score was 0.91 with a 95% CI of 0.86-0.94, higher than individual biomarker AUROC values at recommended thresholds, such as AFP at a cutoff of 10 ng/mL (AUROC: 0.75), AFP at a cutoff of 20 ng/mL (AUROC: 0.71), AFP-L3 at a cutoff of 10% (AUROC: 0.70), and PIVKA-II at a cutoff of 7.5 ng/mL (AUROC: 0.71), with all *P*-values < 0.001. These results are similar to those reported by Berhane et al[14]. In the UK population, the GALAD score had an AUROC of 0.97 with a 95% CI of 0.96-0.98; in the Japanese population, an AUROC of 0.93 with a 95% CI of 0.92-0.94; and in the German population, an AUROC of 0.94 with a 95% CI of 0.93-0.96. Similarly, in a study by Best et al,[7] the AUROC for GALAD was 0.92, with a 95% CI of 0.89-0.96.

Our study had a relatively small sample size and focused only on patients with cirrhosis. Therefore, the GALAD score should be further investigated in larger studies with more advanced study designs, longer durations, and more tightly controlled patient subgroups, including cirrhotic and noncirrhotic populations. This will help assess the value of this score for the clinical diagnosis and prognosis of HCC more comprehensively.

In conclusion, the GALAD score combines sex (G), age (A), and the biomarkers AFP-L3 (L), AFP (A), and Des-gamma-carboxy prothrombin or PIVKA-II (D), thus partly eliminating the shortcomings of individual biomarkers. The GALAD score demonstrates promise in detecting HCC among patients with cirrhosis, exhibiting superior accuracy compared with single biomarkers. Its application in clinical settings could aid in diagnosing HCC in patients with cirrhosis, and calculating the GALAD score may assist in predicting tumor size and quantity before imaging results are available.

ACKNOWLEDGMENTS

The authors thank the Faculty of Medicine, University of Medicine and Pharmacy at Ho Chi Minh City and University Medical Center HCMC, Vietnam.

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 *Cancer J Clin*. 2021;71:209–249.
2. Villanueva A. Hepatocellular carcinoma. *N Engl J Med*. 2019; 380:1450–1462.
3. Singal AG, Llovet JM, Yarchoan M, et al. AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. *Hepatology*. 2023;78:1922–1965.

4. EASL Clinical. Practice guidelines: management of hepatocellular carcinoma. *J Hepatol*. 2018;69:182–236.
5. 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;23:144–153.
6. 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;18:728–735.e4.
7. 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;54:1296–1305.
8. Yang JD, Addissie BD, Mara KC, et al. GALAD score for hepatocellular carcinoma detection in comparison with liver ultrasound and proposal of GALADUS score. *Cancer Epidemiol Biomarkers Prev*. 2019;28:531–538.
9. Rich N, Singal AG. Hepatocellular carcinoma tumour markers: current role and expectations. *Best Pract Res Clin Gastroenterol*. 2014;28:843–853.
10. Kojiro M, Roskams T. Early hepatocellular carcinoma and dysplastic nodules. *Semin Liver Dis*. 2005;25:133–142.
11. Bruix J, Sherman M. Management of hepatocellular carcinoma: an update. *Hepatology*. 2011;53:1020–1022.
12. Omata M, Cheng AL, Kokudo N, et al. Asia-pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatol Int*. 2017;11:317–370.
13. Yamamoto K, Imamura H, Matsuyama Y, et al. AFP, AFP-L3, DCP, and GP73 as markers for monitoring treatment response and recurrence and as surrogate markers of clinicopathological variables of HCC. *J Gastroenterol*. 2010;45:1272–1282.
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;14:875–886.e6.

