Once LC is established, there are no practical treatments for preventing liver cirrhosis (LC) diagnosis in patients with hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). More non-invasive tools are needed to diagnose LC but the predictive abilities of current models are still inconclusive. This study aimed to develop and validate a novel and non-invasive artificial neural network (ANN) model for diagnosing LC in patients with HBV-related HCC using routine laboratory serological indicators.

**Methods:** A total of 1152 HBV-related HCC patients who underwent hepatectomy were included and randomly divided into the training set (n = 864, 75%) and validation set (n = 288, 25%). The ANN model was constructed from the training set using multivariate Logistic regression analysis and then verified in the validation set.

**Results:** The morbidity of LC in the training and validation sets was 41.2% and 46.8%, respectively. Multivariate analysis showed that age, platelet count, prothrombin time and total bilirubin were independent risk factors for LC (P < 0.05). The area under the ROC curve (AUC) analyses revealed that the ANN model had higher predictive accuracy than the Logistic model (ANN: 0.757 vs Logistic: 0.721; P < 0.001), and other scoring systems (ANN: 0.757 vs CP: 0.532, MELD: 0.594, ALBI: 0.575, APRI: 0.621, FIB-4: 0.644, AAR: 0.491, and GPR: 0.604; P < 0.05 for all) in diagnosing LC. Similar results were obtained in the validation set.

**Conclusion:** The ANN model has better diagnostic capabilities than other commonly used models and scoring systems in assessing LC risk in patients with HBV-related HCC.

**Keywords:** chronic hepatitis B, hepatocellular carcinoma, liver cirrhosis, serological indicators, non-invasive assessment, artificial neural network

---

**Background**

In the highly endemic Asia-Pacific region, especially China, 70% to 90% of patients with hepatocellular carcinoma (HCC) are associated with hepatitis B virus (HBV) infection.\(^1\)\(^2\) Liver cirrhosis (LC) is a common end-stage liver disease and is the main contributor to morbidity and mortality in HBV-related HCC patients.\(^3\)\(^4\) Once LC is established, there are no practical treatments for preventing and curing end-stage complications other than follow-up and monitoring.\(^5\) Liver transplantation is the only effective therapy for LC, nevertheless this procedure is very expensive and access to donor transplants remains limited.\(^6\) Therefore, early and accurate evaluation of the disease is of great clinical significance.\(^4\)

Liver biopsy (LB) still the “gold standard” for LC diagnosis.\(^4\)\(^7\) However, LB is an expensive and invasive method that is hard to apply to routine treatments.
Meanwhile, observer discrepancy and sampling errors related to LB contribute to unreliable diagnoses. Furthermore, LB does not facilitate dynamic observation of LC severity. Thus, non-invasive, inexpensive, and convenient methods for the diagnosis of LC are urgently needed.

Current research on non-invasive diagnosis of LC has focused on the use of transient elastography (TE) or scoring systems. TE is a highly reproducible method of indirect LC assessment via liver stiffness measurements. However, this test is of relatively high cost and is a difficult technology to widely popularize. Additionally, the clinical applications of TE are limited due to susceptibility of the technique to variability introduced by many factors, such as elevated aminotransferase and bilirubin levels, obesity, narrow intercostal space, and ascites. Conventional scoring systems including Child-Pugh (CP), model for end-stage liver disease (MELD), albumin-bilirubin (ALBI), aspartate aminotransferase to platelet ratio (APRI), fibrosis index based on the 4 factor (FIB-4), aspartate aminotransferase to alanine aminotransferase ratio (AAR), and gamma-glutamyl transpeptidase to platelet ratio (GPR) scores have been common models used to estimate the degree of LC in chronic liver disease patients. The CP and MELD scores are the most commonly used models for evaluating the prognosis of patients with chronic liver disease. However, the predictive accuracy of the CP score in LC assessment is inevitably affected by subjective factors like ascites and hepatic encephalopathy while the MELD score is optimized for patients with end-stage liver diseases and lacks consideration for the presence of HCC. The ALBI score was developed to assess liver reserve function and predict the incidence and severity of postoperative hepatic dysfunction or failure in HCC patients. Studies have shown that the ALBI score can also assess the prognosis for patients with LC associated with chronic hepatitis B (CHB). The ALBI score has demonstrated simplicity and accuracy in its prognostic abilities, but further exploration is needed. APRI and FIB-4 scores have been recommended in evaluating LC in chronic hepatitis C (CHC) patients within regions with limited resources, however, these models remain controversial for the assessment of LC in patients with CHB. Recently, AAR and GPR scores have been reported in predictions of LC in CHB patients. While these scores may be more accurate than others in assessing LC in CHB patients, it has been difficult to demonstrate satisfactory results in clinical practice. Additionally, more importantly, the interaction between these risk factors in these scoring systems is complex and non-linear, which makes it difficult for traditional linear discriminant analysis to differentiate the significance of each risk variable. An artificial neural network (ANN), which utilizes computer technology to model biological nervous systems functionally and structurally, has been proven to be superior to traditional discriminant analyses.

The purpose of this research was to build a novel non-invasive ANN model for LC diagnosis in HBV-related HCC patients using routine clinical serological indicators. Then, the diagnostic accuracy of the ANN model was compared with the Logistic model and other commonly used scoring systems.

Methods

Patients

This study was conducted retrospectively in patients that had been diagnosed with HBV-related HCC and were initially treated with hepatectomy. These patients treated at the Guangxi Medical University Cancer Hospital between October 11, 2013 and December 31, 2017. The study was conducted in compliance with the Helsinki Declaration and approved by the institutional Ethics Committee of our hospitals, and all patients provided written informed consent. The inclusion criteria were as follows: (1) pre-operative CP grade A/B, (2) pre-operative hepatitis B surface antigen positive, and (3) post-operative pathology confirmed HCC. The exclusion criteria were as follows: (1) pre-operative CP grade C, (2) had concomitant hepatitis C virus, alcoholic liver disease, nonalcoholic fatty liver disease or others, (3) had received any pre-operative anticancer treatments, (4) had cardiopulmonary, renal, or cerebral dysfunction, (5) had cholangiocarcinoma and/or other simultaneous malignancies, and (6) had incomplete clinical information.

Collection of Data on Serological Indicators

Patients remained fasting 12 hrs before examination, and 10 mL venous blood was taken as the research sample in the next morning. The following routine laboratory serological indicators were collected in this study: hepatitis B surface antigen (HBsAg), anti-hepatitis B virus (Anti-HBc), anti-hepatitis B core (Anti-HBc), hepatitis B virus DNA (HBV-DNA), white blood cell (WBC), neutrophils (N),
lymphocyte (L), hemoglobin (HB), platelet (PLT), prothrombin time (PT), fasting blood glucose (FBG), 2h-postprandial blood glucose (2h-PBG), total bilirubin (TBil), prealbumin (PA), albumin (ALB), alanine aminotransferase (ALT), aspartate-aminotransferase (AST), gamma-glutaminate transferase (GGT), alkaline phosphatase (ALP), blood urea nitrogen (BUN), creatinine (CR), hyaluronic acid (HA), laminin (LN), procollagen III N-terminal peptide (PIIINP), collagen IV (IVC), α-fetoprotein (AFP).

Non-Invasive Scoring Systems and Calculation Formulae

The CP score was calculated using T-Bil, ALB, PT, and encephalopathy and ascites;\textsuperscript{11} MELD score = \(3.8 \times \log_2[TBil \ (mg/dL)] + 11.2 \times \log_2(INR) + 9.6 \times \log_2[CR \ (mg/dL)] + 6.4;\textsuperscript{12} \) ALBI score = \(0.660 \times \log_2[TBil \ (\mu mol/l)] - 0.085 \times ALB \ (g/L);\textsuperscript{13} \) APRI score = \([AST \ (U/L)/ULN]/[PLT \ count \ (10^9/L) \times 100;\textsuperscript{14} \) FIB-4 score = \([age \ (years) \times AST \ (U/L)]/[PLT \ count \ (10^9/L) \times \sqrt{ALT \ (U/L)}];\textsuperscript{15} \) AAR score = \([AST \ (U/L)/ALT \ (U/L)];\textsuperscript{16} \) GPR score = \([GGT \ (U/L)/ULN]/[PLT \ count \ (10^9/L) \times 100.\textsuperscript{17} \) All the serological indicators of above formulas were tested and analyzed in the first morning after admission.

Diagnostic Criteria

CHB was defined as the persistent presence of HBV surface antigen with detectable levels of HBV-DNA for more than 6 months.\textsuperscript{29} Diagnoses of HCC and LC were based on histological examination of pathological tissue.\textsuperscript{30,31} Specimens were collected from all the patients after hepatectomy, and pathologic slides were prepared and graded independently by two to three pathologists from our center. According to the Laennec liver fibrosis scoring system,\textsuperscript{31} liver samples were histologically diagnosed without cirrhosis or with cirrhosis.

Hepatectomy and Follow-Up

The operation was performed by experienced surgeons. Before surgery, the type and extent of hepatectomy were determined according to tumor size, number, location and liver functional reserve. The indications and details of the surgical procedures have been described by previous researchers.\textsuperscript{32}

Patients were followed up with 1 month post-operation, then at 2–3 months intervals for the first year, and every 3–6 months for the next few years. Routine re-examinations included blood tests, AFP, abdominal ultrasound scan, CT, or MRI.

Development of the Logistic Model

Variables with \(P \text{ value} < 0.05\) in the univariate analysis were incorporated into the multivariate Logistic regression model to identify independent risk parameters for LC. The Logistic model used the sum of the relevant risks that affect the hazard function to diagnose LC in patients with HBV-related HCC.

Development of the ANN Model

The ANN model was developed by the SPSS 25.0 Neural Network module. The ANN model used in this research was standard feed-forward, back-propagation neural networks with three layers: input nodes, hidden layers and output node. The multilayer perceptron (MLP) network is a new design tool for layered feed-forward network. The input layer is composed of source nodes and the output layer is made up of neurons. These two layers connect the network to the outside world. Apart from these two layers, the MLPs typically have one or more layers of neurons, called hidden neurons, because they can not be accessed directly. The hidden neurons extract the vital characteristics contained in the input data. The MLPs are usually trained forward and backward through back-propagation algorithms. The back-propagation learning algorithms are easy to implement, and their linear complexity on network synaptic weights makes them have satisfactory computationally efficient. To achieve optimal learning efficiency, neurons are usually activated by non-symmetric functions and anti-symmetric functions.\textsuperscript{27,28}

In this research, four selected risk factors (age, PLT count, PT and T-Bil) were served as the input layers neurons, and one factor (presence or absence LC) was served as the output layer neuron. Patients were randomly classified into a derivation group, a test group and a holdout group at the ratio of 7:2:1. The derivation group was used to train an ANN model, the test group was used to verify these results, and the holdout group was used to assess the predictive capacities of the model.

Statistical Analyses

Continuous data were reported as the median (IQR 25–75) and compared by Mann–Whitney \(U\)-test, and categorical data were expressed as frequency and proportion and compared by chi-squared test. The diagnostic accuracy of these models for LC was evaluated via the area under ROC curve (AUC). Discrepancy among the AUCs was compared by \(Z\)-test. SPSS software version 25.0 was
Results

Patients Characteristics

Finally, 1152 HBV-related HCC patients who met the inclusion and exclusion criteria were recruited in this research. Those whole patients were then randomly classified into a training set (n = 864, 75%) and a validation set (n = 288, 25%). The baseline characteristics of these patients are displayed in Table 1. Histopathologically identified LC was found in 523 (45.4%), 396 (45.8%), and 127 (44.1%) patients in the whole set, training set and validation set, respectively.

Identification of Independent Risk Factors

LC-related factors for the univariable analysis (P < 0.05) included: age, positive anti-HBC, WBC, L, PLT count, PT, T-Bil, PA, ALB and BUN (Table 2). For the multivariable analysis, age (OR 1.017 (95% CI, 1.002 to 1.032); P = 0.023), PLT count (OR 0.994 (95% CI, 0.992 to 0.997); P < 0.001), PT (OR 1.582 (95% CI, 1.363 to 1.837); P < 0.001) and T-Bil (OR 1.025 (95% CI, 1.003 to 1.046); P = 0.023) were independently associated with LC in the training set (Table 2).

Development of the ANN Model

The ANN model was developed on the basis of these four risk factors mentioned above (Table 2). When an ANN model was established (Figure 1), we used these four risk factors as the input nodes. The ROC analysis showed that the AUC of the ANN model was greatly higher than the Logistic model (ANN: 0.757 vs Logistic: 0.721; P < 0.001) and these four risk factors individually (Figure 2A and Supplement Table 1, P < 0.001 for all). Analysis of the importance of these four risk factors indicated that T-Bil (100%) was the most vital factor in the ANN model; the following factors were PLT count (91.1%), PT (69.5%), and age (20.1%) (Figure 2B). The prediction probability plot revealed that the ANN model can accurately identify patients without LC (Figure 2C).

Comparison of Diagnostic Accuracy for LC Between the ANN Model and Others Conventional Scoring Systems

In the training set, the ROC analysis showed that the AUC of the ANN model was higher than all the conventional scoring systems (ANN: 0.757 vs CP: 0.532, MELD: 0.594, ALBI: 0.575, APRI: 0.621, FIB-4: 0.644, AAR: 0.491, and GPR: 0.604; P < 0.05 for all) (Figure 3A and Table 3). Similar results were obtained in the validation set (Figure 3B and Table 3).

Discussion

This retrospective study successfully established an ANN model for LC diagnosis in patients with HBV-related HCC. It has obtained satisfactory discriminatory capacities in both training set and the validation set, and has greater diagnostic accuracy than Logistic model and other commonly used scoring systems.

LC is a common end-stage liver disease that affects numerous patients with HBV-related HCC worldwide.\(^1\)–\(^3\) Timely detection and accurate risk diagnosis of LC are critical as the disease becomes more difficult to cure as it progresses.\(^4\) Considering the deficiencies of LB and TE technologies,\(^7\)–\(^9\) the development of non-invasive models for estimation of LC severity using routine serum-based indicators has been a priority for researchers and clinicians.\(^4\) Although several serum marker-based models exist, such as CP, MELD, ALBI, APRI, FIB-4, AAR, and GPR scores, these models have significant limitations that limit their clinical utilization. Therefore, investigation of clinical models for LC diagnosis in HBV-related HCC patients using routine laboratory serological indicators is critical.

Routine blood examinations are standard throughout the course of patient treatment for HCC as it provides significant insight into treatment effectiveness and evaluation of liver disease severity.\(^33\) In this research, we aimed to build a novel non-invasive, in-expensive, and user-friendly tool for LC diagnosis based on serological indicators tested for during routine blood examinations. Ultimately, the multivariate analysis identified age, PLT count, PT and T-Bil were independent risk predictors of LC (Table 2). All these risk factors are common and readily available in clinical practice. Generally, the development of LC is a gradual process. It may take years or even decades to change from mild fibrosis to severe cirrhosis. Studies have shown that the duration of HBV infection is related to the occurrence and development of LC.\(^34\) Though it is hard to obtain accurate HBV infection duration in real world because of the long-term asymptomatic stage, we still believe that age is an alternative indicator of HBV infection duration in China, where the possibility of vertical infection...
Table 1: Patient Characteristics of the Whole Set, Training Set and Validation Set

| Variables       | Whole Set (n=1152) | Training Set (n=864) | Validation Set (n=288) | P value |
|-----------------|--------------------|----------------------|------------------------|---------|
| Age, years      | 49 (41, 57)        | 49 (41, 57)          | 48 (41, 55)            | 0.776   |
| Sex             |                    |                      |                        |         |
| Male            | 998 (86.6)         | 751 (86.9)           | 247 (85.8)             | 0.617   |
| Female          | 154 (13.4)         | 113 (13.1)           | 41 (14.2)              |         |
| Weight, kg      | 60 (54, 68)        | 60 (55, 69)          | 60 (54, 66)            | 0.171   |
| Height, m       | 165 (161, 170)     | 166 (162, 170)       | 165 (160, 170)         | 0.336   |
| BMI, kg/m²      | 22.1 (20.0, 24.4)  | 22.2 (20.1, 24.5)    | 22.0 (20.0, 24.0)      | 0.331   |
| HBeAg           |                    |                      |                        | 0.249   |
| Positive        | 333 (28.9)         | 242 (28.0)           | 91 (31.6)              |         |
| Negative        | 818 (71.1)         | 621 (72.0)           | 197 (68.4)             |         |
| Anti-HBe        |                    |                      |                        | 0.790   |
| Positive        | 317 (27.5)         | 236 (27.3)           | 81 (28.1)              |         |
| Negative        | 835 (72.5)         | 628 (72.7)           | 207 (71.9)             |         |
| Anti-HBC        |                    |                      |                        | 0.129   |
| Positive        | 274 (23.8)         | 196 (22.7)           | 78 (27.1)              |         |
| Negative        | 878 (76.2)         | 668 (77.3)           | 210 (72.9)             |         |
| HBV-DNA, IU/mL  |                    |                      |                        | 0.753   |
| > 2000          | 705 (61.2)         | 531 (61.5)           | 174 (60.4)             |         |
| ≤ 2000          | 447 (38.8)         | 333 (38.5)           | 114 (39.6)             |         |
| WBC, 10⁹/L      | 6.2 (5.0, 7.5)     | 6.2 (5.1, 7.6)       | 6.1 (4.8, 7.4)         | 0.914   |
| N, %            | 59.7 (53.1, 66.3)  | 59.7 (52.7, 66.0)    | 59.7 (53.6, 67.1)      | 0.305   |
| L, %            | 28.0 (21.9, 34.4)  | 28.0 (21.9, 34.5)    | 27.7 (21.2, 34.0)      | 0.515   |
| NLR             | 2.1 (1.6, 3.0)     | 2.1 (1.5, 3.0)       | 2.1 (1.6, 3.1)         | 0.358   |
| HB, g/L         | 138.0 (126.0, 150.0) | 138.0 (127.0, 149.0) | 139.0 (125.0, 151.0)   | 0.560   |
| PLT count, 10⁹/L| 203.6 (158.0, 262.0) | 202.7 (157.8, 261.2) | 206.0 (163.4, 266.6)   | 0.440   |
| PT, s           | 12.9 (12.2, 13.8)  | 12.9 (12.2, 13.8)    | 12.9 (12.3, 13.9)      | 0.823   |
| T-Bil, μmol/L   | 12.5 (9.4, 16.4)   | 12.5 (9.4, 16.5)     | 12.4 (9.5, 16.4)       | 0.794   |
| PA, mg/L        | 179.0 (131.0, 227.0) | 180.0 (132.0, 226.3) | 177.0 (130.0, 232.5)   | 0.660   |
| ALB, g/L        | 40.3 (37.3, 43.2)  | 40.2 (37.3, 43.0)    | 40.7 (37.2, 44.1)      | 0.740   |
| ALT, U/L        | 35.0 (24.0, 51.0)  | 35.0 (24.0, 50.0)    | 35.0 (24.0, 54.0)      | 0.771   |
| AST, U/L        | 39.0 (30.0, 58.0)  | 39.0 (30.0, 58.0)    | 38.0 (29.0, 58.0)      | 0.578   |
| GGT, U/L        | 66.0 (38.0, 122.0) | 65.5 (38.0, 117.0)   | 70.0 (56.0, 104.5)     | 0.928   |
| ALP, U/L        | 73.0 (56.0, 96.0)  | 73.0 (56.0, 95.0)    | 77.0 (56.0, 104.5)     | 0.217   |
| BUN, mmol/L     | 4.9 (4.1, 6.1)     | 4.9 (4.1, 6.0)       | 4.9 (4.1, 6.2)         | 0.634   |
| CR, umol/L      | 77.0 (68.0, 89.0)  | 77.0 (68.0, 88.0)    | 78.0 (68.0, 90.0)      | 0.489   |
| FBG, mmol/L     | 4.6 (4.2, 5.1)     | 4.6 (4.2, 5.1)       | 4.6 (4.2, 5.1)         | 0.249   |
| 2h-FBG, mmol/L  | 6.8 (6.0, 7.7)     | 6.8 (6.0, 7.7)       | 6.8 (5.9, 7.9)         | 0.916   |
| HA, ng/mL       | 41.2 (26.8, 68.1)  | 41.2 (26.1, 67.6)    | 41.2 (29.2, 73.4)      | 0.629   |
| LN, ng/mL       | 30.2 (21.1, 45.6)  | 30.2 (21.6, 46.4)    | 30.2 (20.1, 43.2)      | 0.232   |
| PIIINP, ng/mL   | 2.1 (1.3, 3.7)     | 2.1 (1.4, 3.8)       | 2.1 (1.1, 3.4)         | 0.519   |
| IVC, ng/mL      | 55.0 (41.6, 70.8)  | 55.0 (41.7, 69.5)    | 55.0 (41.3, 73.6)      | 0.830   |
| AFR, ng/mL      |                    |                      |                        | 0.273   |
| > 400           | 508 (44.1)         | 389 (45.0)           | 119 (41.3)             |         |
| ≤ 400           | 644 (55.9)         | 475 (55.0)           | 169 (58.7)             |         |
| CP score        | 5 (5, 5)           | 5 (5, 5)             | 5 (5, 5)               | 0.203   |
| MELD score      | 4 (2, 6)           | 4 (2, 6)             | 5 (2, 7)               | 0.573   |
| ALBI score      | −2.7 (−3.0, −2.4)  | −2.7 (−3.0, −2.4)    | −2.7 (−3.0, −2.4)      | 0.855   |
| APRI score      | 0.5 (0.3, 0.8)     | 0.5 (0.3, 0.8)       | 0.5 (0.3, 0.8)         | 0.896   |

(Continued)
Table 1 (Continued).

| Variables            | Whole Set (n=1152) | Training Set (n=864) | Validation Set (n=288) | P value |
|----------------------|--------------------|----------------------|------------------------|---------|
| FIB-4 score          | 1.7 (1.1, 2.5)     | 1.7 (1.1, 2.5)       | 1.6 (1.1, 2.6)         | 0.913   |
| AAR score            | 1.1 (0.8, 1.5)     | 1.1 (0.9, 1.5)       | 1.1 (0.8, 1.6)         | 0.919   |
| GPR score            | 0.6 (0.4, 1.3)     | 0.6 (0.4, 1.3)       | 0.7 (0.3, 1.3)         | 0.937   |
| LC, pathological     | Yes                | 523 (45.4)           | 396 (45.8)             | 127 (44.1) | 0.608   |
|                      | No                 | 629 (54.6)           | 468 (54.2)             | 161 (55.9) |

Note: Data are median (25th–75th interquartile range) unless otherwise indicated.

Abbreviations: HCC, hepatocellular carcinoma; LC, liver cirrhosis; BMI, body mass index; HBeAg, hepatitis Be antigen; Anti-HBe, anti-hepatitis B virus; Anti-HBC, anti-hepatitis B core; HBV-DNA, hepatitis B virus DNA; WBC, white blood cell; N, neutrophils; L, lymphocyte; HB, hemoglobin; PLT, platelet; PT, prothrombin time; FBG, fasting blood glucose; 2h-PBG, 2h-postprandial blood glucose; T-Bil, total bilirubin; PA, prealbumin; ALB, albumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; BUN, blood urea nitrogen; CR, creatinine; HA, hyluronic acid; LN, laminin; PIIINP, procollagen III N-terminal peptide; IVC, collagen IV; AFP, α-fetoprotein; CP, Child-Pugh; MELD, model for end-stage liver disease; ALBI, albumin-bilirubin; APRI, aspartate aminotransferase to platelet ratio index; FIB-4, fibrosis index based on the 4 factor; AAR, aspartate aminotransferase to alanine aminotransferase ratio; GPR, gamma-glutamyl transpeptidase to platelet ratio.

Table 2 Univariate and Multivariate Analyses on Serological Variables in Diagnosing LC in Patients with HBV-Related HCC in the Training Set

| Variables            | Univariate Logistic Regression | Multivariate Logistic Regression |
|----------------------|--------------------------------|---------------------------------|
|                      | β                               | Odds Ratio                      | P value       | β                        | Odds Ratio | P value       |
| Age, years           | 0.028                           | 1.029 (1.015, 1.042)            | < 0.001       | 0.017                    | 1.017 (1.002, 1.032) | 0.023       |
| Sex, male            | −0.090                          | 0.914 (0.615, 1.358)            | 0.655         |                          |            |               |
| Weight, kg           | 0.004                           | 1.004 (0.991, 1.017)            | 0.565         |                          |            |               |
| Height, m            | 0.005                           | 1.005 (0.985, 1.026)            | 0.618         |                          |            |               |
| BMI, kg/m²           | 0.005                           | 1.005 (0.964, 1.048)            | 0.809         |                          |            |               |
| Positive HBeAg       | 0.190                           | 1.209 (0.898, 1.629)            | 0.211         |                          |            |               |
| Positive anti-HBC    | 0.113                           | 1.120 (0.830, 1.512)            | 0.459         |                          |            |               |
| HBV-DNA, ≥2000 IU/mL | −0.027                          | 0.973 (0.739, 1.281)            | 0.847         |                          |            |               |
| WBC, 10⁹/L           | −0.195                          | 0.823 (0.768, 0.882)            | < 0.001       | −0.056                   | 0.945 (0.869, 1.028) | 0.191       |
| N, %                 | −0.010                          | 0.990 (0.977, 1.002)            | 0.099         |                          |            |               |
| L, %                 | 0.017                           | 1.017 (1.003, 1.032)            | 0.022         | 0.008                    | 1.008 (0.990, 1.026) | 0.408       |
| HB, g/L              | −0.001                          | 0.999 (0.992, 1.006)            | 0.761         |                          |            |               |
| PLT count, 10⁹/L     | −0.008                          | 0.992 (0.990, 0.994)            | < 0.001       | −0.006                   | 0.994 (0.992, 0.997) | < 0.001     |
| PT, s                | 0.565                           | 1.759 (1.537, 2.013)            | < 0.001       | 0.459                    | 1.582 (1.363, 1.837) | < 0.001     |
| T-Bil, μmol/L        | 0.048                           | 1.049 (1.026, 1.072)            | < 0.001       | 0.024                    | 1.025 (1.003, 1.046) | 0.023       |
| PA, mg/L             | −0.003                          | 0.997 (0.995, 0.999)            | 0.001         | 0.001                    | 1.000 (0.997, 1.002) | 0.928       |
| ALB, g/L             | −0.036                          | 0.964 (0.938, 0.991)            | 0.010         | −0.021                   | 0.980 (0.952, 1.008) | 0.157       |
| ALT, U/L             | 0.001                           | 1.000 (0.998, 1.003)            | 0.753         |                          |            |               |
| AST, U/L             | 0.001                           | 1.001 (0.998, 1.004)            | 0.350         |                          |            |               |
| GGT, U/L             | 0.001                           | 1.000 (0.999, 1.002)            | 0.705         |                          |            |               |
| ALP, U/L             | 0.002                           | 1.002 (0.999, 1.005)            | 0.144         |                          |            |               |
| BUN, mmol/L          | 0.003                           | 1.003 (1.000, 1.006)            | 0.026         | 0.001                    | 1.001 (0.998, 1.004) | 0.452       |
| CR, umol/L           | −0.001                          | 0.999 (0.994, 1.003)            | 0.638         |                          |            |               |
| FBG                  | −0.028                          | 0.973 (0.904, 1.047)            | 0.462         |                          |            |               |
| 2h-FBG               | −0.013                          | 0.987 (0.949, 1.027)            | 0.515         |                          |            |               |
| HA, mg/L             | 0.001                           | 1.001 (1.000, 1.002)            | 0.242         |                          |            |               |
| LN, ng/mL            | 0.001                           | 1.000 (0.997, 1.003)            | 0.959         |                          |            |               |
| PIIINP, ng/mL        | 0.001                           | 1.001 (0.995, 1.007)            | 0.717         |                          |            |               |
| IVC, ng/mL           | 0.003                           | 1.003 (1.000, 1.006)            | 0.065         |                          |            |               |
| AFP, ≥400 ng/mL      | −0.062                          | 0.940 (0.718, 1.230)            | 0.651         |                          |            |               |

Abbreviations: HCC, hepatocellular carcinoma; LC, liver cirrhosis; BMI, body mass index; HBeAg, hepatitis Be antigen; Anti-HBe, anti-hepatitis B virus; Anti-HBC, anti-hepatitis B core; HBV-DNA, hepatitis B virus DNA; WBC, white blood cell; N, neutrophils; L, lymphocyte; HB, hemoglobin; PLT, platelet; PT, prothrombin time; FBG, fasting blood glucose; 2h-PBG, 2h-postprandial blood glucose; T-Bil, total bilirubin; PA, prealbumin; ALB, albumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; BUN, blood urea nitrogen; CR, creatinine; HA, hyluronic acid; LN, laminin; PIIINP, procollagen III N-terminal peptide; IVC, collagen IV; AFP, α-fetoprotein.
or transmission is high during childhood. Previous study has also revealed that higher age was associated with higher likelihood of significant fibrosis and cirrhosis in CHB patients.\textsuperscript{35} Similar results were revealed in the research by Hedenstierna et al\textsuperscript{36} which found that age was the main risk factor for persisting LC. In addition, previous studies have shown that the decreased PLT counts, prolonged PT and elevated T-Bil were known surrogate markers and independent predictors of LC.\textsuperscript{37–42} With the aggravating of LC and portal hypertension, the destruction and sequestration of PLTs resulted in decreased PLT counts and splenomegaly.\textsuperscript{39} Besides, progression of LC was also linked to reduced production of thrombopoietin by hepatic cells, which also results in decreased PLT counts.\textsuperscript{40} Similarly, PT and T-Bil levels increased with LC progression, leading to portal hypertension and hepatic cell death.\textsuperscript{41,42} These studies revealed that multiple risk factors were related to the development and progression of LC. Since the relationships between these risk factors are complex, multi-dimensional, and non-linear, the significance of these factors and their interrelationships are hard to reflect only with simple linear models. When exploring the relationship between multiple risk factors and events, the ANN model undoubtedly has a potential to be more successful. As expected, in our study, the ANN model established by the four risk factors mentioned above, performed better than the Logistic model (AUC, ANN: 0.757 vs Logistic: 0.721; $P < 0.001$) and these four risk factors individually in diagnosing LC risk of patients with HBV-related HCC (Figure 2A and Supplement Table 1). All of these factors are easily available, making the ANN model easier to generalize than other studies that use complex radiologic factors in their models.\textsuperscript{43,44}

Furthermore, after comparing discriminatory abilities of the ANN model to other scoring systems for LC diagnosis, we found that the ANN model obtained higher AUCs than did the other commonly used scoring systems (Figure 3 and Table 3). Previous studies suggested that CP, MELD and ALBI scores are most useful in evaluating liver reserve function in predicting post-operative liver failure,\textsuperscript{11–13} but that clinical applications of these models

---

**Figure 1** Schematic representation of the ANN model developed to diagnose LC for patients with HBV-related HCC. The blue lines represent synaptic weight $< 0$, the grey lines represent synaptic weight $> 0$.

**Abbreviations:** ANN, artificial neural network; LC, liver cirrhosis; HBV, hepatitis B virus; HCC, hepatocellular carcinoma.

**Figure 2** (A) ROC curves for the ANN model, Logistic model and four risk factors to diagnose LC; (B) The relative importance of the four risk factors to the ANN model; (C) Prediction probability histograms for the ANN model in patients with HBV-related HCC.

**Abbreviations:** ANN, artificial neural network; LC, liver cirrhosis; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; ROC, receiver operating characteristic.
in diagnosing LC remain controversial.\textsuperscript{18,19,21-23} APRI and FIB-4 scores overtook LB to become the most effective model for CHC-induced LC diagnosis over the past decade.\textsuperscript{14,15} Recently, Wu et al\textsuperscript{26} and Xiao et al\textsuperscript{25} suggested that APRI and FIB-4 scores could identify CHB-related LC with a moderate sensitivity and accuracy. However, Kim et al\textsuperscript{15} proposed that APRI and FIB-4 scores were not suitable for clinical evaluation of LC risk in CHB-related patients. In our study, the accuracy of these two models in evaluating LC risk was also unsatisfactory. Thus, their predictive value for CHB patients remains uncertain. GPR score is a non-invasive method of LC risk determination in CHB patients and they have better predictive accuracy than APRI and FIB-4 scores in West African patients.\textsuperscript{17} Nevertheless, studies have shown that GPR scores were not beneficial in diagnosing LC in Brazilian or Chinese groups.\textsuperscript{45,46} This is consistent with the results of our research on Chinese patients as we found that the predictive accuracy and reliability of LC diagnosis based on GPR scores were not satisfactory compared to the performance of the ANN model. Additionally, AAR score had been used to estimate LC severity in patients

**Table 3** The Performance of the ANN Model and Scoring Systems in Diagnosing LC in the Training Set and Validation Set

| Scoring Systems | Training Set | Validation Set |
|-----------------|--------------|----------------|
|                 | AUC | 95% CI | P value | AUC | 95% CI | P value |
| CP score        | 0.532 | 0.493–0.571 | 0.103 | 0.538 | 0.470–0.605 | 0.269 |
| MELD score      | 0.594 | 0.556–0.631 | < 0.001 | 0.544 | 0.477–0.612 | 0.197 |
| ALBI score      | 0.575 | 0.536–0.613 | < 0.001 | 0.512 | 0.445–0.579 | 0.723 |
| APRI score      | 0.621 | 0.584–0.658 | < 0.001 | 0.603 | 0.537–0.668 | 0.003 |
| FIB-4 score     | 0.644 | 0.608–0.681 | < 0.001 | 0.582 | 0.516–0.648 | 0.017 |
| AAR score       | 0.491 | 0.452–0.530 | 0.638 | 0.470 | 0.403–0.537 | 0.384 |
| GPR score       | 0.604 | 0.567–0.642 | < 0.001 | 0.623 | 0.559–0.688 | < 0.001 |
| ANN model       | 0.757 | 0.726–0.788 | < 0.001 | 0.767 | 0.713–0.820 | < 0.001 |

*Abbreviations:* ANN, artificial neural network; LC, liver cirrhosis; CP, Child-Pugh; MELD, model for end-stage liver disease; ALBI, albumin-bilirubin; APRI, aspartate aminotransferase to platelet ratio index; FIB-4, fibrosis index based on the 4 factor; AAR, aspartate aminotransferase to alanine aminotransferase ratio; GPR, γ-glutamyl transpeptidase to platelet ratio.
with CHB,16 but our research found that there was no relationship between AAR score and incidence of LC. More importantly, these models were also based on the linear relationship of risk indicators, so it is difficult to accurately explain the correlation among them. On the contrary, the ANN models can effectively manage the interactions among these predictors and help improve their diagnostic accuracy. Remarkably, the ANN model developed in this study was considered superior to conventionally used scoring systems due to the ANN model having greater discriminatory performance in diagnosing LC and being resistant to significant influence by inherent heterogeneity in different sets.

The advantages of the ANN model make it maintain satisfactory reliability and stability in diagnosing LC risk. At the same time, the ANN model has a strong applicability in processing complex biological data, but other existing statistical techniques cannot meet this requirement. The ANN model can accurately offer scores from 0 (no probability of LC) to 1 (100% probability of LC) to assess the risk of LC. In the real world, the ANN model can be integrated into an easily accessible website, and clinicians can analyze their own data to assess an individual’s risk of LC before hepatectomy for more rational treatment. For example, patients who are categorized as high probability by their ANN model score should be considered for pre-operative active liver care, bleeding reduction during the operation, limitations of operation duration, and retaining sufficient residual liver volume. Heeding these precautions while strengthening postoperative management has the potential to significantly improve the outcomes of high probability patients.

Finally, other factors such as HBV markers (HBeAg, anti-HBe, anti-HBC and HBV-DNA), inflammatory biomarkers (AST, ALT, GGT and ALP), hepatic fibrosis indices (HA, LN, PIIINP and IVC), and metabolic synthesis indices (PA, ALB) were not predictive parameters for the diagnosis of LC in our study. Our results are inconsistent with previous studies that have concluded that the above serum markers are independent risk factors for predicting LC.437 These differences may be due to differences in research objectives and different pathogenic manifestations of liver disease.

There were several noteworthy limitations of this study. One limitation is that most patients in this study were infected by HBV. Therefore, the predictive value of this ANN model needs further exploration for other etiologies such as HCV. Our study was also limited by a lack of information available about TE, which prevented us from comparing the performances of the ANN model and of TE in estimating LC. Therefore, future studies should include a comparison of the ANN model and TE performances in LC detection in patients with HBV-related HCC.

In conclusion, our study found that age, PLT count, PT and T-Bil were independent risk factors for diagnosing LC. The ANN model we developed based on these four risk variables has improved diagnostic capabilities over conventional Logistic model and commonly used scoring systems. These advancements in LC diagnosis for HBV-related HCC patients are of great importance in guiding clinical treatment decision and improving patient prognosis.

Abbreviations
HCC, hepatocellular carcinoma; LC, liver cirrhosis; CHB, chronic hepatitis B; HBV, hepatitis B virus; PLT, platelet; PT, prothrombin time; Tbil, total bilirubin; LB, liver biopsy; TE, transient elastography; CP, Child-Pugh; MELD, model for end-stage liver disease; ALBI, albumin-bilirubin; APRI, aspartate aminotransferase to platelet ratio index; FIB-4, fibrosis index based on the 4 factor; AAR, aspartate aminotransferase to alanine aminotransferase ratio; GPR, gamma-glutamyl transpeptidase to platelet ratio; AUC, area under the receiver operating characteristic curve.

Funding
The study was supported by the National Science Foundation of China (No. 81803007 and 81660498); China Postdoctoral Science Foundation (NO. 2019M663412); GuangXi Natural Science Foundation (No. 2018GXNSFBA281030, 2018GXNSFBA281091 and 2019JJA140151); Guangxi Key Research and Development Plan (No. GUIKEAB19245002); Guangxi Medical and Health Appropriate Technology Development and Application Project (No. S2017101 and S2018062).

Disclosure
The authors declare that they have no conflicts of interest.

References
1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017;67(1):7–30. doi:10.3322/caac.21387
2. Asia-Pacific Working Party on Prevention of Hepatocellular Carcinoma. Prevention of hepatocellular carcinoma in the Asia-Pacific region: consensus statements. J Gastroenterol Hepatol. 2010;25(4):657–663. doi:10.1111/j.1440-1746.2009.06167.x
3. Yoneda M, Imajo K, Eguchi Y, et al. Noninvasive scoring systems in patients with nonalcoholic fatty liver disease with normal alanine aminotransferase levels. J Gastroenterol. 2013;48(9):1051–1060. doi:10.1007/s00535-012-0704-y
38. Ceylan B, Fincancı M, Yardımcı C, et al. Can mean platelet volume determine the severity of liver fibrosis or inflammation in patients with chronic hepatitis B. *Eur J Gastroenterol Hepatol*. 2013;25(5):606–612. doi:10.1097/MEG.0b013e32835d08da
39. Aster RH. Pooling of platelets in the spleen: role in the pathogenesis of “hypersplenic” thrombocytopenia. *J Clin Invest*. 1966;45(5):645–657. doi:10.1172/JCI105380
40. Giannini E, Borro P, Botta F, et al. Serum thrombopoietin levels are linked to liver function in untreated patients with hepatitis C virus-related chronic hepatitis. *J Hepatol*. 2002;37(5):572–577. doi:10.1016/S0168-8278(02)00274-X
41. Papatheodoridis GV, Papakonstantinou E, Andrioti E, et al. Thrombotic risk factors and extent of liver fibrosis in chronic viral hepatitis. *Gut*. 2003;52(3):404–409. doi:10.1136/gut.52.3.404
42. Chang Y, Ryu S, Zhang Y, et al. A cohort study of serum bilirubin levels and incident non-alcoholic fatty liver disease in middle aged Korean workers. *PLoS One*. 2012;7(5):e37241. doi:10.1371/journal.pone.0037241
43. Da BL, Surana P, Kleiner DE, Heller T, Koh C. The Delta-4 fibrosis score (D4FS): a novel fibrosis score in chronic hepatitis D. *Antiviral Res*. 2020;174:104691. doi:10.1016/j.antiviral.2019.104691
44. Kao WY, Chang IW, Chen CL, et al. Fibroscan-based score to predict significant liver fibrosis in morbidly obese patients with nonalcoholic fatty liver disease. *Obes Surg*. 2020;30(4):1249–1257. doi:10.1007/s11695-019-04192-w
45. Schiavon LL, Narciso-Schiavon JL, Ferraz MLG, Silva AEB, Carvalho-Filho RJ. The y-glutamyl transpeptidase to platelet ratio (GPR) in HBV patients: just adding up? *Gut*. 2017;66(6):1169–1170. doi:10.1136/gutjnl-2016-312658
46. Li Q, Song J, Huang Y, et al. The gamma-glutamyl-transpeptidase to platelet ratio does not show advantages than APRI and Fib-4 in diagnosing significant fibrosis and cirrhosis in patients with chronic hepatitis B: a retrospective cohort study in China. *Medicine*. 2016;95(16):e3372. doi:10.1097/MD.0000000000003372