Serum fibrosis index-based risk score predicts hepatocellular carcinoma in untreated patients with chronic hepatitis B

Lilian Yan Liang¹²³,*, Hye Won Lee⁴,*, Vincent Wai-Sun Wong¹²³, Terry Cheuk-Fung Yip¹²³, Yee-Kit Tse¹²³, Vicki Wing-Ki Hui¹²³, Grace Chung-Yan Lu²³, Henry Lik-Yuen Chan¹²³, and Grace Lai-Hung Wong¹²³

¹Institute of Digestive Disease, ²Department of Medicine and Therapeutics, ³Medical Data Analytic Centre (MDAC), The Chinese University of Hong Kong, Hong Kong SAR; ⁴Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Korea

Graphical Abstract

Abbreviations:
AFP, alpha-fetoprotein; ALT, alanine aminotransferase; APRI, aspartate aminotransferase to platelet ratio index; AUROC, area under the time-dependent receiver operating characteristic curve; CDARS, the Clinical Data Analysis and Reporting System; CHB, chronic hepatitis B; CI, confidence interval; CU-HCC, the Chinese University of Hong Kong-HCC; DM, diabetes mellitus; FIB-4, fibrosis-4; GAG-HCC, the guide with age, gender, HBV DNA, core promoter mutations and cirrhosis-HCC; GGT, gamma-glutaryl transpeptidase; HBeAg, hepatitis e antigen; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; ICD-9-CM, the International Classification of Disease, Ninth Revision, Clinical Modification; IQR, interquartile range; NPV, negative predictive value; REACH-B, the risk estimation for HCC in CHB; WHO, World Health Organization

Corresponding author: Grace Lai-Hung Wong
Department of Medicine and Therapeutics, The Chinese University of Hong Kong, 9/F, Lui Che Woo Clinical Sciences Building, Shatin, N.T., Hong Kong
Tel: +852-2632-3538, Fax: +852-2637-3852
E-mail: wonglahung@cuhk.edu.hk
https://orcid.org/0000-0002-2863-9389

*Liang LY and Lee HW have equal contribution to this manuscript.
INTRODUCTION

Chronic hepatitis B virus (HBV) infects approximately 292 million persons around the world and attributes to about half of the hepatocellular carcinoma (HCC) cases every year. HCC is a cancer with high mortality rate, with most cases detected at advanced stages. HCC surveillance remains indispensable to detect tumor at early stages in order to improve the prognosis of patients. Chronic hepatitis B (CHB) patients with high serum HBV DNA levels, elevated alanine aminotransferase (ALT) levels, advanced fibrosis and necroinflammation are at risk of developing HCC, hence are indicated for antiviral therapy. CHB patients received antiviral treatment have significantly lower risk of HCC and liver-related mortality. On the other hand, patients who do not fulfill treatment criteria may also develop HCC. Thus, some untreated patients may have a significant risk of HCC development and thus benefit from antiviral treatment.

Advanced fibrosis and cirrhosis are the key risk factors of HCC. Liver biopsy examination was commonly used to assess liver fibrosis, yet its invasiveness and potential complications limit its generalizability. Liver stiffness measurement by transient elastography is a non-invasive method to assess liver fibrosis, yet the measurement is limited by the availability of the machine. Instead, various serum test-based formulae have been developed to predict liver fibrosis with a reasonable accuracy. While some of these formulae (for example Enhanced Liver Fibrosis) require the measurement of uncommon and expensive laboratory parameters including aminoterminal propeptide of type III procollagen, hyaluronic acid and tissue inhibitor of matrix metalloproteinase type-1. A few formulae comprised of common laboratory tests have been widely used. The most commonly used formulae include aspartate aminotransferase to platelet ratio index (APRI), fibrosis-4 (FIB-4) index, and Forns index. These serum fibrosis scores that reflected the severity of liver fibrosis could be useful to define the HCC risk of CHB patients, in particulars in those remained untreated. Existing HCC risk scores for untreated CHB patients including the Chinese University of Hong Kong-HCC (CU-HCC) score, the guide with age, gender, HBV DNA, core promoter mu-
tations and cirrhosis-HCC (GAG-HCC) score (without core promoter mutation), and the risk estimation for HCC in CHB (REACH-B) score rely heavily on clinical diagnosis of cirrhosis. The clinical diagnosis can however be subjective and vary among centers as liver biopsy is not performed in most of the patients. Any misdiagnosis of liver cirrhosis will affect much on the classification of patients to different risk groups.

Based on the recommendations of American Association for the Study of Liver Diseases, an annual HCC risk of <0.2% is probably not cost-effective for HCC surveillance. In view of the huge number of CHB patients, a prediction score with high sensitivity to identify patients at very low risk of HCC would be pivotal to appropriate allocation of healthcare resources for HCC surveillance. We may exempt these very-low-risk patients from HCC surveillance and offer regular monitoring at primary care setting instead of secondary or tertiary centers. Appropriate resource allocation facilitates the elimination of HBV infection by 2030, a goal set by the World Health Organization (WHO). In the current study, we aimed to derive and validate a HCC risk score developed based on serum test formulae to exclude treatment-naive CHB patients with low HCC risk.

PATIENTS AND METHODS

Study design and data source

The training cohort was a territory-wide retrospective cohort identified using data from the Clinical Data Analysis and Reporting System (CDARS) under the management of Hospital Authority, Hong Kong. Different territory-wide studies have been conducted using CDARS. The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) coding system was used in CDARS. The use of ICD-9-CM codes in CDARS to identify hepatic events has been found 99% accurate when referenced to clinical, laboratory, imaging and endoscopy results from the electronic medical records. The derived HCC risk score was externally validated in an independent cohort of CHB patients from Severance Hospital, Seoul, Korea.

Subjects

Consecutive adult CHB patients with positive hepatitis B surface antigen for 6 months and/or defined by the ICD-9-CM diagnosis codes between 2000 and 2017 in Hong Kong were identified. Those patients with serum HBV DNA and at least one of the serum fibrosis scores (APRI, FIB-4, and Forns index) measured between January 2000 and June 2018 were included. Follow-up duration were censored at the start date of first antiviral treatment for patients ever received antiviral treatment. HCC surveillance was done every 6 to 12 months by ultrasound and alpha-fetoprotein (AFP) testing under the Hospital Authority. APRI, FIB-4, and Forns index were calculated as described (Supplementary Table 1).

We excluded the following patients: those co-infected with hepatitis C virus and/or hepatitis D virus based on ICD-9-CM diagnosis codes, viral and serological markers and/or use of antiviral therapy for hepatitis C; co-infected with human immunodeficiency virus based on ICD-9-CM diagnosis codes and/or use of antiviral therapy for human immunodeficiency virus; had other co-existing autoimmune and metabolic liver diseases based on ICD-9-CM diagnosis codes; had cancer including HCC and/or liver transplantation before or within the first 6 months after baseline date; follow-up duration less than 6 months; received rituximab and any chemotherapeutic agents (Supplementary Tables 1, 2). We also excluded patients with ICD-9-CM diagnosis codes of acute hepatitis B and/or positive immunoglobulin M to hepatitis B core antigen unless it was more than 6 months apart from a positive hepatitis B surface antigen result, patients with Child-Pugh class B or C or had any hepatic events before baseline. For validation cohort, CHB patients identified in 2000–2017 with available HBV DNA results in Severance Hospital, Seoul, Korea were included. We excluded patients age younger than 18 years; develop HCC before or within first 6 months after baseline; follow-up less than 6 months; without available serum fibrosis scores at baseline; with Child Pugh B or C and/or had any hepatic events before baseline.

Patients were followed until diagnosis of HCC, death, liver transplantation, starting antiviral treatment, last follow-up date (December 31, 2018), or up to 5 years of follow-up, whichever came first. The study protocol was approved by the Joint Chinese University of Hong Kong - New Territories East Cluster Clinical Research Ethics Committee.

Data collection

Data of training cohort were retrieved from CDARS in January 2019. Baseline date was set on the first serum HBV DNA assay. Demographic data including date of birth and gender were captured. At baseline, liver and renal biochemistries, hematological and virologic parameters were collected. Thereafter, serial liver
and renal biochemistries as well as HBV viral markers were collected until the last follow-up date. We also retrieved data on other relevant diagnoses, procedures, concomitant drugs, laboratory parameters. Drug data of any antiviral treatment for CHB including any approved nucleos(t)ide analogues, namely lamivudine, adefovir dipivoxil, entecavir, tenofovir disoproxil fumarate and tenofovir alafenamide, as well as (pegylated)-interferon were also retrieved during the study period; patients’ follow-up was censored at the time of first antiviral treatment.

Definitions

The primary endpoint was HCC. HCC was identified based on diagnosis codes (155.0: HCC and 155.2: carcinoma of liver) or procedure codes for HCC treatment according to ICD-9-CM codes from CDARS. Liver cirrhosis was identified by ICD-9-CM diagnosis codes for cirrhosis and its related complications (Supplementary Table 2). Diabetes mellitus (DM) was defined by exposure to any anti-diabetic agents, and/or hemoglobin A1c ≥6.5% and/or fasting plasma glucose ≥7.0 mmol/L in one measurement, and/or the ICD-9-CM diagnosis codes for DM.

Statistical analysis

Data were analyzed using SPSS version 25.0 (IBM Corp., Armonk, NY, USA), and R software (4.0.0; R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were expressed in mean±standard deviation or median (interquartile range [IQR]), as appropriate, while categorical variables were presented as frequency (percentage). Time to HCC diagnosis was defined as the time from the baseline date to the date of first HCC diagnosis code, or the date of last follow-up if patients were still alive without HCC at the time of analysis. Kaplan-Meier analysis with log-rank test was used to estimate and compare the cumulative risks of development of HCC in different groups of patients. Cox proportional hazards model was performed to determine the relationship of clinical characteristics of HBV DNA level, gender, age, serum bilirubin level, Fib-4 index and serum hepatitis e antigen (HBeAg) status with the development of HCC. Missing values were excluded from the analysis. A simple HCC risk score was derived by significant variables obtained from backward stepwise multivariable analysis with P<0.05. The score was the weighted sum of those variables of which the weights were defined as the

Figure 1. The selection of patients in the training cohort and validation cohort. CHB, chronic hepatitis B; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; AST, aspartate aminotransferase; PLT, platelet; ALT, alanine aminotransferase; GGT, gamma-glutaryl transpeptidase; NA, nucleos(t)ide analogues.
quotient (rounded to nearest integer) of corresponding estimated coefficient from a Cox regression analysis divided by the smallest chi-square coefficient.11

The performance of serum fibrosis scores and the HCC risk score was assessed by area under the time-dependent receiver operating characteristic curve (AUROC).27 The 95% confidence interval (CI) for the comparison of two ROC curves was computed by bootstrap sampling of 1,000 samples. Diagnostic accuracy of the suggested cutoff values was assessed by sensitivity, specificity, positive predictive value, and negative predictive value (NPV) with 95% CI. Youden’s index was used to select the best cutoff value. High sensitivity and NPV were used to further select the cutoff values with comparable Youden’s index to exclude patients with low HCC risk. All statistical tests were two-sided. P <0.05 was taken as statistically significant.

**RESULTS**

**Patients characteristics**

We identified in total 135,395 CHB patients in the public healthcare system in 2000–2017.28 There were 60,539 CHB patients with available HBV DNA results at baseline. After excluding patients fulfilling the exclusion criteria, the final number of treatment-naïve CHB patients included in the training cohort was 15,187. We identified 8,559 untreated CHB patients with available HBV DNA measurement in the validation cohort, 4,286 patients were finally analyzed after exclusion (Fig. 1).

During a mean follow-up of 52 months, 180 patients developed HCC in the training cohort. We compared the characteristics of patients developed HCC and without HCC. Patients who devel-

| Table 1. Clinical characteristics of patients in the training and validation cohorts |
|-------------------|-------------------|-------------------|-------------------|-------------------|
|                   | Training cohort   | Validation cohort |
|                   | HCC (n=180)       | Non-HCC (n=15,007) | HCC (n=47)       | Non-HCC (n=4,239) |
| Male              | 133 (73.9)        | 8,220 (54.8)      | <0.001           | 39 (83.0)         | 2,249 (53.1)      | <0.001 |
| Age (years)       | 61.4±11.4         | 51.8±13.6         | <0.001           | 58.0±11.9         | 48.4±12.6         | <0.001 |
| Age <50 years     | 24 (13.3)         | 6,295 (41.9)      | <0.001           | 10 (21.3)         | 2,228 (51.6)      | <0.001 |
| Albumin (g/L)     | 38.1±6.6          | 41.6±4.9          | <0.001           | 39.5±5.3          | 42.8±4.7          | <0.001 |
| Bilirubin (µmol/L)| 14.1 (9.2–19.1)   | 11.0 (8.0–15.8)   | <0.001           | 13.7 (10.3–18.8)  | 12.0 (10.3–17.1)  | 0.964 |
| ALT (U/L)         | 46.0 (31.0–72.3)  | 370 (22.0–73.0)   | 0.001            | 26.0 (17.0–39.0)  | 24.6 (16.0–42.0)  | 0.678 |
| AST (U/L)         | 33.1 (24.0–52.2)  | 30.0 (23.0–49.0)  | 0.128            | 30.0 (24.0–53.0)  | 24.0 (19.0–36.0)  | 0.001 |
| Platelet (×10^9/L)| 161.3±68.2        | 208.5±68.2        | <0.001           | 155.1±63.2        | 201.2±68.3        | <0.001 |
| GGT (U/L)         | 70.2 (37.0–164.5) | 290 (19.0–58.0)   | <0.001           | 56.5 (24.0–100.5) | 25.0 (16.0–46.0)  | 0.001 |
| Cholesterol (mmol/L) | 4.6±1.2         | 4.8±1.1           | 0.002            | 4.2±1.0          | 4.7±1.0           | 0.002 |
| Positive HBeAg    | 28 (18.8)         | 2,403 (17.7)      | 0.738            | 5 (13.2)         | 786 (23.3)        | 0.140 |
| HBV DNA (log_{10} IU/mL) | 5.0±1.8       | 4.4±2.1          | <0.001           | 3.1±1.7          | 3.5±2.1           | 0.014 |
| HBV DNA >2,000 IU/mL | 138 (76.7)   | 8,711 (58.0)      | <0.001           | 11 (23.4)        | 1,374 (32.4)      | 0.189 |
| Cirrhosis         | 56 (31.1)         | 765 (51.1)        | <0.001           | 37 (78.7)        | 916 (21.6)        | <0.001 |
| APRI              | 0.7 (0.5–1.2)     | 0.4 (0.3–0.7)     | <0.001           | 0.5 (0.4–0.9)    | 0.3 (0.2–0.6)     | <0.001 |
| FIB-4 index       | 2.5 (1.6–4.4)     | 1.3 (0.9–2.1)     | <0.001           | 2.4 (1.6–4.1)    | 1.3 (0.8–2.0)     | <0.001 |
| FIB-4 >3.25       | 54 (38.6)         | 1,611 (12.5)      | <0.001           | 18 (38.3)        | 471 (11.1)        | <0.001 |
| Forns index       | 74±2.0            | 5.6±2.0           | <0.001           | 6.8±2.1          | 4.8±2.1           | <0.001 |
| Advanced fibrosis (%) | 70 (38.9)      | 2,087 (13.9)      | <0.001           | 20 (42.6)        | 537 (12.7)        | <0.001 |
| Follow-up duration (months) | 48.9±34.8 | 68±28.7        | <0.001           | 46.0±32.0        | 49.5±33.6        | 0.477 |

Values are presented as mean±standard deviation, median (interquartile range), or number (%). 
HCC, hepatocellular carcinoma; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutaryl transpeptidase; HBeAg, hepatitis B e antigen; HBV, hepatitis B virus; APRI, aspartate aminotransferase to platelet ratio index; FIB-4, fibrosis-4.
Opiated HCC were older (61.4±11.4 vs. 51.8±13.6 years), and more likely to be male (73.9% vs. 54.8%) than patients without HCC. More patients of non-HCC group were younger than 50 years (41.9%). For liver biochemistries, patients who developed HCC had lower albumin level (38.1±5.6 vs. 41.6±4.9 g/L), platelet count (161.3±68.2 vs. 208.5±68.2×10^9/L), cholesterol level (4.6±1.2 vs. 4.8±1.1 mmol/L) and higher total bilirubin (14.1 [IQR, 9.2–19.1] vs. 11.0 [IQR, 8.0–15.8] µmol/L), ALT (46.0 [IQR, 31.0–72.3] vs. 37.0 [IQR, 22.0–73.0] IU/L), gamma-glutaryl transpeptidase (GGT; 70.2 [IQR, 37.0–164.5] vs. 29.0 [IQR, 19.0–58.0] U/L) than patients without HCC. Patients with HCC had cirrhosis (31.1% vs. 5.1%) and higher HBV DNA level (5.0±1.8 vs. 4.4±2.1 log_{10} IU/mL). Their serum fibrosis scores of APRI (0.7 [IQR, 0.5–1.2] vs. 0.4 [IQR, 0.3–0.7]), FIB-4 index (2.5 [IQR, 1.6–4.4] vs. 1.3 [IQR, 0.9–2.1]), and Forns index (7.4±2.0 vs. 5.6±2.0) were all higher than those of non-HCC patients (Table 1). The AFP levels in patients developed HCC was significantly higher than patients did not develop HCC (6.1 [IQR, 3.3–11.6] vs. 3.1 [IQR, 2.2–4.9] ng/mL, P<0.001). Patients developed HCC had more proportion of DM than patients did not develop HCC (37.2% vs. 18.6%, P<0.001).

In the validation cohort, 47 patients developed HCC during a mean follow-up of 50 months. Compared with the training cohort, patients in the validation cohort were younger (48.5±12.6 vs. 52.0±13.6) and had more patients with cirrhosis (22.2% vs. 5.4%). Patients had similar albumin level, platelet count, APRI and FIB-4 index in both of the training and validation cohorts (Table 1).

### Predictors of HCC

We performed univariate and multivariable analysis for APRI, FIB-4 index separately. Forns index is not generally used in clinical assessment of Hong Kong and there were more than half of missing values in our cohort. So, we did not select Forns index to develop the new score. APRI ≥2 was not associated with HCC development in both univariate and multivariable analysis (Supplementary Table 3). FIB-4 >3.25, male gender, age older than 50 years old, high total bilirubin (bilirubin >18 µmol/L), low platelet (platelet <150×10^9/L), high HBV DNA level (HBV DNA >2,000 IU/mL) and DM were correlated with HCC development by univariate analysis. There were 2,155 patients (14.2%) without FIB-4 index and these patients were excluded from the Cox proportional hazards model. Male gender, age older than 50, high HBV DNA level and high FIB-4 could predict HCC development significantly in multivariable analysis (Table 2).

### Derivation of HCC prediction score

FIB-4 score combined with gender, age and HBV DNA level which were significant in the multivariable analysis comprised the new HCC prediction score named Liang score and ranged from 0 to 22 (Table 3). We listed out several potential cutoff values and calculated the diagnostic accuracy of these values (Table 4). Among these val-

| Table 2. Univariate and multivariable analysis of Cox proportional hazards models on risk factors associated with hepatocellular carcinoma in the training cohort (follow-up duration: 5 years) |
|---|---|---|---|---|---|---|---|
| **Factor** | **No. of patients** | **Univariate** | **Multivariable** |
| **** | **HR** | **95% CI** | **P-value** | **χ² score** | **Adjusted HR** | **95% CI** | **P-value** |
| Male sex | 8,353 (55.0) | 1.96 | 1.31–2.95 | 0.001 | 6.71 | 1.96 | 1.18–3.26 | 0.010 |
| Age (years) | | | | | | | |
| <50 | 6,319 (41.6) | 1 | | | | | |
| 50–60 | 4,833 (31.8) | 5.64 | 2.83–11.23 | <0.001 | 15.89 | 4.93 | 2.25–10.79 | <0.001 |
| ≥60 | 4,035 (26.6) | 10.61 | 5.44–20.71 | <0.001 | 21.32 | 6.51 | 2.94–14.43 | <0.001 |
| Bilirubin >18 µmol/L | 2,537 (16.7) | 2.00 | 1.33–3.01 | 0.001 | | | |
| Platelet <150×10^9/L | 2,654 (17.5) | 3.40 | 2.34–4.95 | <0.001 | | | |
| HBV DNA >2,000 IU/mL | 8,849 (58.3) | 2.07 | 1.36–3.15 | 0.001 | 4.26 | 1.77 | 1.03–3.03 | 0.039 |
| FIB-4 >3.25 | 1,667 (11.0) | 4.46 | 2.89–6.88 | <0.001 | 15.16 | 2.70 | 1.64–4.45 | <0.001 |
| HBeAg (+) | 2,431 (16.0) | 0.86 | 0.49–1.49 | 0.586 | | | |
| DM | 2,865 (18.9) | 3.14 | 2.16–4.58 | <0.001 | | | |

Values are presented as number (%) unless otherwise indicated.

HR, hazard ratio; CI, confidence interval; HBV, hepatitis B virus; FIB-4, fibrosis-4; HBeAg, hepatitis B e antigen; DM, diabetes mellitus.
ues, cutoff value of 10 had the highest Youden's index of 0.406, followed by the cutoff value of 9 with Youden's index of 0.391. As the corresponding sensitivity and NPV of 9 were 83.7% (95% CI, 73.9–90.5%) and 99.8% (95% CI, 99.7–99.9%), which were higher than those of the cutoff value of 10 (sensitivity, 74.4%; specificity, 99.7%), we finally selected 9 as the cutoff value of the Liang score. By using the cutoff value of 9, patients were divided into two groups: low-risk group (7,191 patients) with Liang score ≤9 and high-risk group (5,841 patients) with Liang score >9. The corresponding number of patients developed HCC were 14 (0.19%) and 72 (1.12%) in the low- and high-risk groups at 5 years’ follow-up duration. The cumulative incidence of HCC was significantly different between low- and high-risk patients by Kaplan-Meier analysis (log-rank test, P<0.001). The annual HCC incidences of low-risk group were less than 0.2%, which is the threshold of offering cost-effective HCC surveillance (Fig. 2A).

We compared the performance of Liang score in patients never receive antiviral treatment and patients received antiviral treatment and censored at the start date of the treatment (Supplementary Table 4). The sensitivity (81.82% vs. 84.38%) and NPV (99.91% vs. 99.61%) were similar between these two groups.

**External validation of HCC prediction score**

In the validation cohort, there were 2,686 patients (62.8%) classified as low-risk groups, among which three patients (0.11%) developed HCC. The corresponding sensitivity and NPV of cutoff of 9 were 90.0% (95% CI, 72.3–97.4%) and 99.9% (95% CI, 99.6–99.97%), respectively (Table 5). The cumulative incidence of HCC in low-risk patients was significantly lower than that in high-risk patients (log-rank test, P<0.001), and was less than 0.2% per year (Fig. 2B). Liang score performed well in patients with or without cirrhosis and subgroups of patients with different AFP levels (Supplementary Tables 5, 6).

**Comparison with existing HCC prediction scores**

We compared the performance of Liang score with other HCC prediction scores which were developed in treatment-naïve CHB patients: CU-HCC score, GAG-HCC score and REACH-B score in the validation cohort. We selected conventionally used cutoff values of 5, 100 and 8 for CU-HCC score, GAG-HCC score and REACH-B

### Table 3. Components of the Liang score

<table>
<thead>
<tr>
<th>Factor</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
</tr>
<tr>
<td>Male</td>
<td>+3</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>&lt;50</td>
<td>0</td>
</tr>
<tr>
<td>50–60</td>
<td>+7</td>
</tr>
<tr>
<td>≥60</td>
<td>+10</td>
</tr>
<tr>
<td>HBV DNA</td>
<td></td>
</tr>
<tr>
<td>≤2,000</td>
<td>0</td>
</tr>
<tr>
<td>&gt;2,000</td>
<td>+2</td>
</tr>
<tr>
<td>FIB-4</td>
<td></td>
</tr>
<tr>
<td>≤3.25</td>
<td>0</td>
</tr>
<tr>
<td>&gt;3.25</td>
<td>+7</td>
</tr>
<tr>
<td>Total</td>
<td>0 to 22</td>
</tr>
</tbody>
</table>

HBV, hepatitis B virus; FIB-4, fibrosis-4.

### Table 4. Diagnostic accuracy of Liang score (follow-up duration: 5 years)

<table>
<thead>
<tr>
<th>Cut off value</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>97.7 (91.1–99.6)</td>
<td>95.3 (87.9–98.5)</td>
<td>90.7 (82.0–95.6)</td>
<td>89.5 (80.6–94.8)</td>
<td>83.7 (73.9–90.5)</td>
<td>74.4 (63.7–82.9)</td>
<td>55.8 (44.7–66.4)</td>
</tr>
<tr>
<td>Specificity</td>
<td>19.5 (18.8–20.2)</td>
<td>28.8 (28.0–29.6)</td>
<td>42.0 (41.2–42.9)</td>
<td>48.2 (47.3–49.0)</td>
<td>55.4 (54.6–56.3)</td>
<td>66.2 (65.4–67.0)</td>
<td>80.2 (79.5–80.9)</td>
</tr>
<tr>
<td>PPV</td>
<td>0.8 (0.6–1.0)</td>
<td>0.9 (0.7–1.1)</td>
<td>1.0 (0.8–1.3)</td>
<td>1.1 (0.9–1.4)</td>
<td>1.2 (1.0–1.6)</td>
<td>1.4 (1.1–1.8)</td>
<td>1.8 (1.4–2.5)</td>
</tr>
<tr>
<td>NPV</td>
<td>99.9 (99.7–99.98)</td>
<td>99.9 (99.7–99.97)</td>
<td>99.9 (99.7–99.9)</td>
<td>99.9 (99.7–99.9)</td>
<td>99.8 (99.7–99.9)</td>
<td>99.7 (99.6–99.8)</td>
<td>99.6 (99.5–99.7)</td>
</tr>
<tr>
<td>No. of HCC predicted</td>
<td>84/86</td>
<td>82/86</td>
<td>78/86</td>
<td>77/86</td>
<td>72/86</td>
<td>64/86</td>
<td>48/86</td>
</tr>
<tr>
<td>No. of low-risk patients</td>
<td>2,522/13,032</td>
<td>3,730/13,032</td>
<td>5,451/13,032</td>
<td>5,451/13,032</td>
<td>7,191/13,032</td>
<td>8,592/13,032</td>
<td>10,422/13,032</td>
</tr>
</tbody>
</table>

Values are presented as percentage (95% confidence interval) or number.

PPV, positive predictive value; NPV, negative predictive value; HCC, hepatocellular carcinoma.
The sensitivity and NPV results were 83.3% and 99.8% for CU-HCC score; 70.0% and 99.8% for GAG-HCC score and 65.2% and 99.6% for REACH-B score, which were all lower than those of Liang score (90.0% and 99.9%). The number of low-risk patients defined by CU-HCC score, GAG-HCC score and REACH-B score were 2,403, 3,639, and 1,933 with HCC incidence of 0.21%, 0.25%, and 0.41%, respectively (Table 5).

Time-dependent AUROC was used to compare the predictive ability of Liang score with other HCC prediction scores. The time-dependent AUROC of Liang score was 0.79 (95% CI, 0.70–0.89) which was not significantly different to those of CU-HCC score (0.82; 95% CI, 0.75–0.89; P=0.434) and GAG-HCC score (0.83; 95% CI, 0.74–0.91; P=0.342). The time-dependent AUROC of Liang score was significantly higher than that of REACH-B score (0.61; 95% CI, 0.53–0.70; P<0.001; Supplementary Fig. 1).

Patients who may be exempted from HCC surveillance

Patients with Liang score of 9 or lower may be exempted from HCC surveillance which include the following groups: female patients younger than 50; female patients age between 50 and 60 with low FIB-4 score or male patients younger than 50 with low FIB-4 score (Table 3). The corresponding annual HCC incidences of these groups were all less than 0.2% in the training cohort.

DISCUSSION

We aimed to derive and validate a HCC risk score based on serum fibrosis scores to identify a group of patients with very low risk of HCC in treatment-naive CHB patients. The newly developed
Liang score was composed of gender, age, HBV DNA and FIB-4 index, which was accurate and reproducible to exclude HCC. By using the cutoff value of 9, around 50% of patients were classified as very low risk of HCC (annual incidence <0.2%), so that HCC surveillance would be exempted. As Liang score was composed of common clinical and laboratory parameters, it would be applicable in most clinic settings. This score would contribute to the WHO’s hepatitis elimination advocacy by better resource allocation.

The current international guidelines recommend that adult CHB patients with cirrhosis should undergo regular HCC surveillance. However, the recommendations in non-cirrhotic CHB patients are not as clear. For example, the American guidelines recommend screening in male CHB patients above 40 years old, which refers to majority of the male patients as adults of early 30s or younger are mostly vaccinated in our region. With this novel Liang score, the following types of CHB patients would be classified as high-risk with a cutoff value of 9: 1) female patients older than 60 years; 2) female patients with age between 50 and 60 years and FIB-4 >3.25; 3) male patients older than 50 years; and 4) male patients younger than 50 years with FIB-4 >3.25. The corresponding cumulative HCC incidence of these groups at 5 years’ follow-up were 1.5%, 1.8%, 1.7%, and 2.2% in the training cohort. The annual HCC incidences were all higher than 0.2% that indicates the possible need of HCC surveillance. These few groups of patients account for 45% in the training cohort and 37% in the validation cohort.

A few existing HCC prediction scores such as CU-HCC score and GAG-HCC score include clinical cirrhosis as a key component with strong weight; yet cirrhosis may not be diagnosed accurately by ultrasonography. We substituted clinical cirrhosis with FIB-4 score, a more objective serum fibrosis score, and then built the Liang score, which is highly applicable and reproducible. By using the cutoff value of 9, Liang score had the highest sensitivity (90.0%) and NPV (99.9%) compared with CU-HCC score, GAG-HCC score and REACH-B score which were developed in treatment-naive patients. Therefore, Liang score was accurate to rule out HCC in low-risk patients and it worked excellently among current HCC prediction scores in treatment-naive CHB patients without the necessity of accurate diagnosis of liver cirrhosis.

Our study had some strengths. First, Liang score is composed of common clinical and laboratory parameters such that it is applicable in most clinic settings. Aiming a highly sensitive cutoff achieving very high NPV can identify very-low-risk patients who would not need HCC surveillance in the near future. Second, Liang score is accuracy to exclude low HCC risk patients without the diagnosis of cirrhosis and not affected by operators compared with ultrasound, which is relatively inaccurate to diagnose cirrhosis for HCC surveillance. Third, Liang score was derived in the cohort of Hong Kong patients and externally validated in the Korean cohort. Both of the cohorts had respectable sample size and the results supported that Liang score is accurate and generally applicable in different ethnic groups in Asia.

There were also few limitations of our study. First, Liang score was developed in Hong Kong cohort and validated in Korean cohort, such that its accuracy in Caucasians is uncertain. Further studies based on patients of other ethnic groups are needed. Second, our study was based on treatment-naive CHB patients. Until further data are available, the findings of our study cannot be extrapolated to patients on antiviral treatment. Third, we did not explore the role of the HBV genotypes as it is not a routine test in Hong Kong and we cannot show the size and stage of HCC. The common HBV genotypes in Hong Kong include genotype B and C; whereas it is exclusively genotype C in Korea. Fourth, the HCC incidence of our cohort was only 1.2% in more than 5 years, as it was a treatment-naive cohort. Such low incidence further support the urgent need of a highly sensitive score like Liang score which saves the resources of HCC surveillance in around 50% of such patients. Fifth, there were many missing HBV DNA results in our cohort as the test was available widely in Hong Kong from the year of 2012. Many old patients may did not receive the test. Sixth, we did not collect the presence of non-alcoholic fatty liver disease and other metabolic factors like insulin resistance and obesity which may affect the risk of HCC at long term. Finally, the diagnosis of cirrhosis in our study most based on the ultrasonography, which may be not accurate and influence the reliability of CU-HCC score and GAG-HCC score.

In conclusion, Liang score, composed of common parameters including gender, age, HBV DNA and FIB-4 index, is accurate to exclude HCC in treatment-naive CHB patients. Liang score identifies patients at very low risk of HCC so that HCC surveillance can be exempted. Further studies are needed to validate the Liang score in prospective cohorts.

Authors’ contribution
All authors were responsible for the study concept and design. Lilian Liang, Yee-Kit Tse, Terry Yip and Grace Wong were responsible for the acquisition and analysis of data, had full access to all of the data in the study and take responsibility for the integrity of
the data and the accuracy of the data analysis. All authors were responsible for the interpretation of data, the drafting, and critical revision of the manuscript for important intellectual content.

Acknowledgements
This work was supported by the Health and Medical Research Fund (HMRF) of the Food and Health Bureau (Reference no: 15160551).

Conflicts of Interest
Vincent Wong has served as an advisory committee member for AbbVie, Allergan, Gilead Sciences, Janssen, Perspectum Diagnostics and Pfizer; and a speaker for Bristol-Myers Squibb, EchoSens, Gilead Sciences and Merck.

Terry Yip has served as an advisory committee member and speaker for Gilead Sciences.

Henry Chan is an advisor for AbbVie, Aptorum, Alimmune, Arbutus, Intellia, Janssen, Gilead, GRAIL, Medimmune, Roche, Vir Biotechnology; and a speaker for AbbVie, Gilead and Roche.

Grace Wong has served as an advisory committee member for Gilead Sciences and Janssen, as a speaker for Abbott, Abbvie, Ascleisis, Bristol-Myers Squibb, EchoSens, Furui, Gilead Sciences, Janssen and Roche, and received research grant from Gilead Sciences.

The other authors have no conflicts to disclose.

SUPPLEMENTARY MATERIAL
Supplementary material is available at Clinical and Molecular Hepatology website (http://www.e-cmh.org).

REFERENCES


