Transient elastography-based risk estimation of hepatitis B virus-related occurrence of hepatocellular carcinoma: development and validation of a predictive model

Do Young Kim1,2,4,*
Ki Jun Song2,4,*
Seung Up Kim1,2,4
Eun Jin Yoo1
Jun Yong Park1,2,4
Sang Hoon Ahn1,2,4
Kwang-Hyub Han1,2,4

1Department of Internal Medicine,
2Institute of Gastroenterology,
3Department of Biostatistics, Yonsei University College of Medicine,
4Liver Cirrhosis Clinical Research Center, Seoul, Korea
*These authors contributed equally to this work

Background: The purpose of this study was to develop and validate a novel transient elastography-based predictive model for occurrence of hepatocellular carcinoma (HCC).

Methods: A total of 1,250 patients with chronic hepatitis B and baseline liver stiffness values were recruited between May 2005 and December 2007. The predictive model for HCC occurrence was constructed based on a Cox proportional hazards model. We estimated baseline disease-free probabilities at 3 years. Discrimination and calibration were used to validate the model.

Results: HCC occurred in 56 patients during a median follow-up of 30.7 months. Multivariate analysis revealed that age, male gender, and liver stiffness values were independent predictors of HCC (all \(P<0.05\)), whereas hepatitis B virus DNA \(>20,000\) IU/L showed borderline statistical significance (\(P=0.0659\)). We developed a predictive model for HCC using these four variables, which showed good discrimination capability, with an area under the receiver operating characteristic curve (AUROC) of 0.806 (95% confidence interval 0.738–0.874). We used the bootstrap method to assess discrimination. The AUROC remained largely unchanged between iterations, with an average value of 0.802 (95% confidence interval 0.791–0.812). The predicted risk of occurrence of HCC calibrated well with the observed risk, with a correlation coefficient of 0.905 (\(P<0.001\)).

Conclusion: This novel model accurately estimated the risk of HCC occurrence in patients with chronic hepatitis B.

Keywords: hepatocellular carcinoma, transient elastography, prediction, hepatitis B

Introduction
Hepatocellular carcinoma (HCC) is the fifth most common cancer globally and is a serious public health problem, particularly in association with chronic viral hepatitis. Several factors, such as chronic hepatitis B virus (HBV) or hepatitis C virus infection, cirrhosis, alcohol consumption, and obesity have been identified as risk factors for development of HCC.1–5 Moreover, the observation that over half of HCC cases and deaths are estimated to occur in far Eastern countries, including Korea, Taiwan, and China, indicates that HBV infection has a major influence on worldwide HCC epidemiology.6

Because timely and appropriate treatment and monitoring could reduce the incidence of HCC or detect early cancers in those with a high probability of progressing to HCC, it is important to be able to predict HCC occurrence in patients with chronic hepatitis B (CHB). Both host and viral factors can affect progression of CHB to cirrhosis.
Kim et al

Patients and methods

Predictive model derivation

We prospectively identified 1,250 patients with CHB who visited the outpatient clinic and received a transient elastography examination at our institution between May 2005 and December 2007. We then excluded 140 patients who did not provide written informed consent, had unreliable transient elastography data, or had insufficient clinical data for statistical analysis. Finally, the study cohort consisted of 1,110 patients, and this was used to develop the predictive model for HCC occurrence. The database for our cohort included information on patient demographics, laboratory results, and liver stiffness values at the time of enrollment. A trained medical reviewer collected patient data from medical charts. The study was approved by the independent institutional review board of Severance Hospital, Yonsei University College of Medicine, Seoul, Korea.

Liver stiffness measurement

Liver stiffness measurement was performed according to previously reported methods. Briefly, liver stiffness was measured on the right lobe of the liver through the intercostal space, with the patient lying in the dorsal decubitus position with the right arm in maximal abduction. The operator located a liver portion that was at least 6 cm thick and free of large vascular structures, and pressed the probe button to commence the measurement. Only one experienced technician, blind to the patient clinical data, performed the transient elastography, with results expressed in kilopascals. Only transient elastography examinations with at least 10 validated liver stiffness measurements, a success rate of at least 60%, and an interquartile range to median value ratio ≤0.3 were considered reliable. Any liver stiffness value that did not satisfy these conditions was considered unreliable and was excluded from further analysis. When a reliable examination was conducted, the median value of the successful measurements was selected as a representative liver stiffness value in a given patient.

Diagnosis of cirrhosis and HCC

Liver cirrhosis was clinically diagnosed as follows: platelet count <100,000/µL and ultrasonographic findings suggestive of cirrhosis, including a blunted, nodular liver edge accompanied by splenomegaly (>12 cm); esophageal or gastric varices; or overt complications of liver cirrhosis. HCC was diagnosed histologically or clinically based on the guidelines of the American Association for the Study of Liver Diseases. Briefly, HCC was diagnosed for tumors >2 cm when typical features of HCC were detected (hypervascularity and washout in the portal/venous phase) by one dynamic imaging technique and when the alpha fetoprotein level was >200 ng/mL. For nodules measuring 1–2 cm, it was necessary to detect typical features of HCC using two dynamic imaging techniques. Ultrasonography was repeated after 3 months for tumors <1 cm.

Statistical analyses

All statistical analyses were conducted using SAS version 9.2 software (SAS Institute, Cary, NC, USA). The predictive model of HCC occurrence was constructed based on
a Cox proportional hazards model, which was used to identify independent predictors of HCC occurrence with their respective $\beta$-regression coefficient, $P$-value, hazard ratio, and 95% confidence interval (CI). The median follow-up period was 30.7 (range 24.0–50.9) years, and we estimated baseline disease-free probabilities at 3 years.

\[
1 - P_0 \exp(\beta X)
\]

where $P_0$ is the baseline disease-free probability and $\beta_i$ is the $\beta$-regression coefficient for the $i$th independent covariate. The model was validated by discrimination and calibration. Discrimination was assessed using the receiver operating characteristic curve, area under the receiver operating characteristic curve (AUROC), sensitivity, and specificity. We calculated the predicted risk of HCC occurrence at 3 years based on our predictive model and then estimated the observed Kaplan–Meier risk of HCC to assess calibration. Kaplan–Meier estimates were plotted against the predicted risk in the group to create a calibration chart.

## Results

### Baseline characteristics

Table 1 shows the baseline characteristics of the cohort ($n=1,110$). The mean age was 50 years, and there were 760 males and 350 females. Cirrhosis was clinically diagnosed in 182 (16.3%) patients. The median ALT level and liver stiffness value was 40 IU/L and 7.7 kPa, respectively. HBeAg positivity was observed in 400 (36.0%) patients, and HBV DNA $\geq$20,000 IU/L was observed in 349 (31.4%) patients. Fifty-six patients developed HCC during the study period. The one-year, 2-year, and 3-year cumulative HCC incidence was 0.90%, 1.45%, and 6.05%, respectively.

### Development of a predictive model of HCC occurrence

In addition to the three independent variables (age, male gender, and liver stiffness value), HBV DNA $\geq$20,000 IU/L was also incorporated as a constituent variable to develop the predictive model for HCC occurrence, which has been a generally accepted significant risk factor for HCC.

#### Table 1 Baseline characteristics of the cohort ($n=1,110$)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic data</strong></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>50 (21–81)</td>
</tr>
<tr>
<td>Male gender</td>
<td>760 (68.5)</td>
</tr>
<tr>
<td>Clinically diagnosed cirrhosis</td>
<td>182 (16.3)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>31 (2.8)</td>
</tr>
<tr>
<td>Previous or ongoing antiviral treatment</td>
<td>420 (37.8)</td>
</tr>
<tr>
<td>Body mass index, kg/m$^2$</td>
<td>27 (22–35)</td>
</tr>
<tr>
<td><strong>Laboratory data</strong></td>
<td></td>
</tr>
<tr>
<td>Serum albumin, g/dl</td>
<td>4.5 (2.8–6.3)</td>
</tr>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>0.9 (0.2–3.8)</td>
</tr>
<tr>
<td>Aspartate aminotransferase, IU/L</td>
<td>38 (12–167)</td>
</tr>
<tr>
<td>Alanine aminotransferase, IU/L</td>
<td>40 (8–172)</td>
</tr>
<tr>
<td>Prothrombin time, %</td>
<td>92 (45–100)</td>
</tr>
<tr>
<td>Platelet count, $\times 10^9$/L</td>
<td>160 (45–520)</td>
</tr>
<tr>
<td>HBeAg positivity</td>
<td>400 (36.0)</td>
</tr>
<tr>
<td>High HBV DNA $\geq$20,000 IU/L</td>
<td>349 (31.4)</td>
</tr>
<tr>
<td><strong>Liver stiffness measurement</strong></td>
<td></td>
</tr>
<tr>
<td>Liver stiffness, kPa</td>
<td>7.7 (2.9–70)</td>
</tr>
<tr>
<td>Interquartile range, kPa</td>
<td>1.7 (0.1–9.8)</td>
</tr>
<tr>
<td>Success rate, %</td>
<td>95 (64–100)</td>
</tr>
</tbody>
</table>

**Note:** Variables are expressed as median (range) or n (%).

**Abbreviations:** HBeAg, hepatitis B e-antigen; HBV, hepatitis B virus.

#### Table 2 Univariate and multivariate analysis to identify independent predictors of HCC occurrence

<table>
<thead>
<tr>
<th>Factor</th>
<th>Univariate $P$-value</th>
<th>Multivariate $P$-value</th>
<th>Multivariate Beta</th>
<th>Hazard ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$&lt;0.0001$</td>
<td>0.0531</td>
<td>0.0003</td>
<td>1.054 (1.025–1.085)</td>
</tr>
<tr>
<td>Male gender</td>
<td>0.0140</td>
<td>1.106</td>
<td>0.0044</td>
<td>3.022 (1.411–6.473)</td>
</tr>
<tr>
<td>LSM value</td>
<td>$&lt;0.0001$</td>
<td>0.0486</td>
<td>$&lt;0.0001$</td>
<td>1.050 (1.034–1.065)</td>
</tr>
<tr>
<td>HBV DNA</td>
<td>0.0465</td>
<td>0.5097</td>
<td>0.0659</td>
<td>1.665 (0.967–2.866)</td>
</tr>
<tr>
<td>$\geq$20,000 IU/L</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** HCC, hepatocellular carcinoma; CI, confidence interval; LSM, liver stiffness measurement; HBV, hepatitis B virus.

### Multivariate Cox proportional hazards model

First, the influence of each variable (age, gender, clinically diagnosed cirrhosis, diabetes mellitus, previous or ongoing antiviral treatment, body mass index, alpha fetoprotein, serum albumin, total bilirubin, aspartate aminotransferase, ALT, prothrombin time, platelet count, HBeAg positivity, HBV DNA, and liver stiffness value) was tested in a univariate analysis. Only age, male gender, liver stiffness value, and HBV DNA were statistically significant in univariate analysis (Table 2). A subsequent multivariate analysis was then performed using these four variables. Table 2 shows the $\beta$-regression coefficient estimates with the multivariate Cox proportional hazards model. In our cohort, age, male gender, and liver stiffness values were selected as independent predictors of HCC occurrence (all $P<0.05$), and HBV DNA $\geq$20,000 IU/L showed borderline statistical significance ($P=0.0659$). The other variables was not statistically significant ($P>0.1$).
This newly developed predictive model showed fairly good discrimination capability, with an AUROC of 0.806 (95% CI 0.738–0.874, Figure 1).

Probability = 1 – \( P^A \)

\[
A = \exp(0.05306 \times \text{age} + 1.106 \times \text{male gender} + 0.04858 \times \text{liver stiffness values} + 0.50969 \times \text{HBV DNA} \geq 20,000 \text{ IU/L})
\]

**Discrimination and calibration**

We used the bootstrap method, in which 1,000 random samples drawn with replacement from the original data set and the coefficients were recalculated in each bootstrap sample. The AUROCs remained largely unchanged between iterations, with an average AUROC of 0.802 (95% CI 0.791–0.812). We plotted a calibration chart for predicted and observed risk of HCC occurrence (Figure 2). The predicted risk of HCC occurrence calibrated well with the observed risk, with a correlation coefficient of 0.905 (\( P < 0.001 \)).

**Discussion**

This prospective cohort study was undertaken to develop a clinically useful predictive model for HCC occurrence in patients with CHB using a single tertiary hospital-based cohort from South Korea, and the resulting predictive model was validated internally.

Although extensive studies have identified risk factors associated with HCC occurrence in patients with CHB, few mathematically modeled risk estimations have been proposed.\(^8,10,14,26,27\) The use of hard-to-obtain variables, such as HBeAg and specific HBV mutations (precore or core promoter), as constituent variables or selecting naïve patients for antiviral treatment to develop a predictive model for HCC occurrence had made it difficult to apply universally to all patients infected with HBV.\(^14\) In this regard, our new model has the advantage that four simple, not exhaustive, noninvasively obtainable, and objective variables of age, male gender, HBV DNA, and liver stiffness values were included. Although statistical significance was not identified, even in the univariate analysis, we tried to adjust the influence of antiviral treatment. Finally, our new predictive model for HCC occurrence was accurate and reliable, with an AUROC of 0.806, similar to that of a study by Yang et al in which the AUROC of the risk scores was 0.811,\(^10\) and with a predicted risk correlating well with observed risk (correlation coefficient of 0.905).

The suggested formula for 3-year probability of HCC occurrence, ie, \( 1 – P^A \) (\( A = \exp(0.05306 \times \text{age} + 1.106 \times \text{male gender} + 0.04858 \times \text{liver stiffness values} + 0.50969 \times \text{HBV DNA} \geq 20,000 \text{ IU/L}) \)), is a revision of what we have previously reported, in which age, gender, cirrhosis, HBV/hepatitis C virus infection, ALT, alpha fetoprotein, and alcohol consumption were included as variables.\(^27\) It may be argued that cirrhosis, which has been the single most important risk factor for initiating a surveillance program to detect HCC occurrence early and was selected as a significant predictor of HCC occurrence in univariate analysis of our study, should be included in this model. However, a histologic diagnosis of cirrhosis is not feasible in clinical practice, and no established and sufficiently validated clinical diagnostic criteria for cirrhosis are available, which is the reason why cirrhosis has not been incorporated into several previous risk prediction models for HCC occurrence. Further, increasing evidence has indicated that liver stiffness values, measured objectively by transient elastography, correspond accurately to the severity of liver fibrosis in both HBV and hepatitis C virus-related chronic liver diseases and that baseline liver stiffness values in patients with CHB predict HCC occurrence better than clinical diagnostic criteria of cirrhosis.\(^23,24\) All these results help to explain the superiority of liver stiffness values for diagnosing compensated cirrhosis. Thus, apart from the finding that liver stiffness value, not clinically diagnosed cirrhosis, was selected as one of the independent predictors of HCC occurrence in the multivariate analysis, our model might have been powered by liver stiffness values which are noninvasive and more reproducible than liver biopsy with a potential sampling error and interpretational variability, and more reliable and accurate than clinical diagnostic criteria.
Risk estimation of HCC based on liver stiffness measurement

for cirrhosis without sufficient validations. In addition, when the liver stiffness value was removed from our model, AUROC declined significantly from 0.806 to 0.713 (95% CI 0.689–0.787; \( P = 0.0296 \)), which clearly indicates that the liver stiffness value can enhance the performance of our prediction model for HCC occurrence.

In addition to cirrhosis, ambiguous variables such as ALT, alfa fetoprotein, and alcohol, were removed in our revised model. Although several studies have shown that these factors might be closely associated with HCC occurrence, fluctuations in serum ALT levels and the influence of other conditions on alfa fetoprotein levels have made these markers unreliable for predicting HCC occurrence. Further, information on the amount of alcohol consumed by each patient is still dependent solely on patient interview, without any objective measure of alcohol consumption, and no definition of heavy alcohol consumption exists.

The finding that HBeAg status was not a risk factor for HCC occurrence in our study is in line with previous studies from Hong Kong and Taiwan, but disagrees with another report from Taiwan where a risk score incorporating gender, age, ALT level, HBeAg, and HBV DNA level was developed from a community-based untreated patient cohort with CHB. Therefore, it would be ideal to derive a model for predicting HCC occurrence in patients with CHB using multinational data with a large sample size, considering that a heterogeneous study population even in the same country could generate different risk factors associated with HCC occurrence. Unfortunately, because we used a less sensitive hybridization capture assay for measuring HBV DNA levels, we were not able to lower the cutoff of HBV DNA levels below 20,000 IU/mL during the study period.

Finally, antiviral treatment was not selected as an independent predictor of HCC occurrence. This result can be explained by the current antiviral treatment protocol for patients with active hepatitis to suppress viral replication and HCC occurrence according to most guidelines of the American Association for the Study of Liver Diseases, European Association for the Study of the Liver, and Asia Pacific Association for the Study of the Liver.

Our predictive model for HCC occurrence will be useful for several purposes in clinical practice. First, it will provide an optimal surveillance strategy for patients with CHB, which are not covered by current management recommendations. More intensive follow-up for HCC surveillance will be needed for patients assessed as having a high probability of HCC occurrence, whereas those with a low risk could be monitored with a less vigorous surveillance protocol. Second, because HBV DNA and liver stiffness values can change during the natural course of the disease, or by antiviral

Figure 2 Calibration chart for predicted versus observed risk of occurrence of HCC. The predicted risk of occurrence of HCC calibrated well with the observed risk, with a correlation coefficient of 0.905 (\( P < 0.001 \)).

Abbreviation: HCC, hepatocellular carcinoma.
treatment, and changes in these values also have a significant influence on HCC occurrence, dynamic changes in the risk of HCC occurrence could be theoretically monitored by our new predictive model. If this hypothesis can be validated in future studies, patient compliance in a follow-up protocol and initiation of treatment could be enhanced by showing them the changing risk probability.

Several issues remain unresolved in our study. First, although use of transient elastography as a tool for estimation of liver fibrosis in predicting HCC occurrence is a novel approach, the use of transient elastography is still confined to referral hospitals. Thus, a limitation on the widespread use of our new model is anticipated. However, because use of transient elastography in clinical practice is currently increasing, this new predictive model in our pilot trial could be a cornerstone for future studies. Second, the risk calculation for HCC occurrence in our study was based on the probability of HCC occurrence at 3 years, considering the median follow-up period of our study cohort (30.7 months). Thus, more long-term follow-up seems needed to assess accurately the risk of HCC occurrence. Third, a potentially overestimated HCC incidence rate could be another issue, because our study subjects were recruited from a tertiary hospital. Last, external validation of our model in different countries with different ethnicities should be conducted for widespread practical use.

In conclusion, we have developed a simple transient elastography-based predictive model for HCC occurrence in patients with CHB. Although the performance of this new model was reasonable, based on an internal validation, the implementation of transient elastography in private clinics and external validation of our model are necessary before this new model can be used in the clinic.

Acknowledgment

This study was supported by the Liver Cirrhosis Clinical Research Center, in part by a grant from the Korea Healthcare Technology R and D Project, Ministry of Health and Welfare, Republic of Korea (A102065), the Yonsei Liver Blood Bank, and in part by a grant from Sanofi-aventis Korea. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References