**Hepatoma Research**

**Review**                                                                                    **Open Access**

# Statistical strategies for HCC risk prediction models in patients with chronic hepatitis B

**Terry Cheuk-Fung Yip**[1,2,3], **Vicki Wing-Ki Hui**[1,2,3], **Yee-Kit Tse**[1,2,3], **Grace Lai-Hung Wong**[1,2,3]

[1]Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China.
[2]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China.
[3]Medical Data Analytic Centre (MDAC), The Chinese University of Hong Kong; Hong Kong, China.

**Correspondence to**: Prof. Grace Lai-Hung Wong, Department of Medicine and Therapeutics, 9/F Lui Che Woo Clinical Sciences Building, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong, China. E-mail: wonglaihung@cuhk.edu.hk

## Abstract

Risk prediction modelling for hepatocellular carcinoma (HCC) has been the focus of research in the last decade. The prediction models would help HCC risk stratification, so that patients at high risk of HCC would be able to receive more appropriate management and HCC surveillance. These models were mostly developed in treatment-naïve chronic hepatitis B patients in the early days. In recent years, more prediction models were derived and validated in patients who have received antiviral treatment, which account for the majority of patients who are at increased risk of HCC. Various statistical tests are adopted in developing and validating a risk prediction model - commonly Cox proportional hazards regression, time-dependent receiver operating characteristic (ROC) curve and area under the ROC curve. Even in well-validated models, there may be some pitfalls, *e.g.*, generalizability and clinical applicability. The future direction of prediction model development should be directed towards a more personalised approach. Continuous optimisation of the predictive accuracy of the models would be achieved by involving more serial and dynamic parameters.

**Keywords**: Cirrhosis, hepatitis B virus, hepatocellular carcinoma, liver stiffness measurement, prediction models

## INTRODUCTION

Development and validation of hepatocellular carcinoma (HCC) risk prediction models remain a hot area of liver research. Its importance is not just at the academic level, but also at the practical level. The burning

need of some accurate as well as applicable HCC risk prediction models is intensified by the World Health Organization's goal of eliminating hepatitis B virus (HBV) infection by 2030. This initiative calls for actions to reduce chronic viral hepatitis incidence and mortality to 80% and 65% respectively[1]. As the majority of the mortality from chronic hepatitis B (CHB) is secondary to HCC[2,3], accurate HCC risk prediction is the key component of secondary prevention of HCC[4].

HCC is one of the top killers as it carries a high mortality rate, despite advances in HCC treatments[5]. HCC represents the third most frequent cause of cancer death globally (782,000 deaths in 2018)[2]. Chronic HBV infection is a key risk factor for HCC development, which accounts for approximately 50% of cases worldwide and as high as 70%-80% of cases in regions where HBV is highly endemic[6]. HCC surveillance facilitates early HCC diagnosis and makes curative treatments possible[7]. Regular surveillance with transabdominal ultrasound scanning with or without tumour markers every 6 months in all CHB patients would be a significant burden on healthcare resources[8]. This is especially true in the Asia-Pacific region, as the majority of HCC disease burden (85%) locates in low- and middle-income countries with high prevalence of HBV in the region[9]. Accurate HCC models enable risk-stratification for the huge number of CHB patients, so that healthcare resources can be targeted to patients who are at risk.

There are more than a dozen well-validated HCC predication models; some were developed mainly in untreated CHB patients, whereas some intended for nucleos(t)ide analogues (NA)-treated patients[4,10]. In this review article, we present a focused discussion on the key statistical strategies adopted in the development and validation of HCC prediction models.

## COMMON STATISTICAL TESTS ADOPTED WHEN DEVELOPING AND VALIDATING A RISK PREDICTION MODEL

Although a semi-parametric Cox proportional hazards (PH) regression is widely used for developing a prediction model of a time-to-event outcome, the sample size requirements and follow-up durations for derivation and validation datasets of risk prediction models must be carefully considered. Of note, the effective sample size is defined by the number of events in Cox models. A rule of thumb is to have at least 10 events per variable at an initial stage (i.e., Total number of candidate variables. More accurately, it refers to the number of parameters to be estimated) for deriving a model and a minimum of 100 outcome events for validation cohorts[11]. Candidate prognostic factors should be chosen a priori on the basis of clinical knowledge, literature review, data quality and availability, and cost constraints. Often, a univariate analysis, using either the log-rank test or Cox regression, is applied to all predictors and then those potential variables with a *P*-value less than a pre-specified significant level (say, $P < 0.2$) are entered to multivariable Cox PH model with (backward) stepwise approach to further reduce the model complexity. Although pre-filtering by univariate selection seems attractive, it should be avoided where possible[12]. Moreover, a stepwise selection method is unstable especially with a low effective sample size. In such cases, model selection procedure by backward stepwise or elimination with a significance level of 0.157 [i.e., Akaike's information criterion selection as a default stopping criterion] is recommended[13].

To assess model fit, martingale residuals can be examined for checking the assumption of linear effect of covariates on log hazard rate for continuous predictors. If linearity assumption is violated, nonlinear relationships can be investigated using fractional polynomials or restricted cubic splines. In contrast, Schoenfeld residuals are used to test the assumption of proportional hazards, either by graphical or analytical methods. A risk score (linear combination of model predictors with regression coefficients offering weights) is calculated for each subject, followed by determining an optimal cut-off value to stratifying individuals into risk categories based on a pre-defined decision rule. The sensitivity and specificity at optimal cut point are subsequently estimated, together with Kaplan-Meier curves and the log-rank test can be used to evaluate the different risk profiles.

**Table 1. Statistical strategies for HCC risk scores**

| Scores | Formulae | Statistical strategies |
|---|---|---|
| Untreated patients | | |
| CU-HCC | Age > 50 (+3) + serum album ≤ 35 g/dL (+20) + serum total bilirubin > 18 umol/L (+1.5) + HBV DNA 4-6 $\log_{10}$ IU/mL (+1) OR > 6 $\log_{10}$ IU/mL (+4) + cirrhosis (+15) | Cox proportional hazard model Linear trend $\chi^2$ test |
| LSM-HCC | Age > 50 (+10) + serum album ≤ 35 g/dL (+1) HBV DNA 4 $\log_{10}$ IU/mL (+5) + liver stiffness measurement 8-12 kPa (+8) OR > 12 kPa (+12) | Cox proportional hazard model Youden's Index |
| Untreated patients | | |
| PAGE-B | Age U 30 (+2 to +10) + Male (+6) + Platelet < 200 (+6 to +9) | Cox proportional hazard model Points system |
| mPAGE-B | Age U 30 (+3 to +11) + Male (+2) + Platelet < 250 (+2 to +5) + Albumin < 40 g/dL (+1 to +3) | Cox proportional hazard model Points system |

HCC: hepatocellular carcinoma; LSM: liver stiffness measurement; HBV: hepatitis B virus

In addition, the time-dependent receiver operating characteristic curve ROC(t) and area under the ROC curve AUC(t) analyses for survival data can be employed at some specific times of interest to assess predictive power of the model[14]. Other performance metrics of model discrimination can also be computed including, among others, Harrell's concordance index (C-index) and Uno's concordance statistic. It may be preferable to report Uno's concordance statistic as C-index is affected by censoring[15]. For calibration, which seems to be often neglected, a measure proposed by Grønnesby and Borgan can be readily carried out by comparing the observed and predicted number of events based on dividing predicted risk scores into G different groups {where G = integer of [max(2, min(10, number of failures/40))]} to assess the overall goodness-of-fit in particular to the Cox model[16,17]. A calibration slope should also be presented routinely for both internal and external validation, of which a value close to 1 indicates good calibration. Conducting internal validation is crucial, preferably by bootstrap resampling[18]. This technique can not only evaluate the stability of selected predictors in a multivariable model, but also correct prognostic index obtained from the original sample for optimism. For external validation, the 'final' model derived from derivation cohort is utilized to a new population to judge generalizability and transportability (some executable STATA codes can be found in the Supplementary Material).

## STATISTICAL STRATEGIES FOR HCC RISK SCORES IN UNTREATED PATIENTS

### Examples: CU-HCC and LSM-HCC scores

CU-HCC and liver stiffness measurement (LSM)-HCC scores [Table 1] are clinical scoring systems derived from the hospital cohorts for the prediction of HCC in CHB patients[19,20]. The LSM-HCC score is a refined version of the CU-HCC score, which assigns a heavy weight to cirrhosis[20]. As the diagnosis of cirrhosis in CU-HCC score based on ultrasonography may be incorrect in some patients, cirrhosis is replaced by LSM, a more objective and accurate assessment for advanced liver fibrosis and cirrhosis[21]. Both CU-HCC and LSM-HCC scores have applied similar statistical strategies, namely Cox proportional hazard model for determining the relationship between HCC and clinical variables with the development of HCC (*e.g.,* HBV DNA level, LSM), and various discrimination methods for HCC risk group classification (*i.e.,* Youden's Index in LSM-HCC and linear trend $\chi^2$ test in CU-HCC).

The development of both the CU-HCC and LSM-HCC scores started with identifying significant risk factors of HCC. One approach is to include all categorized risk factors such as age, gender, and albumin (*i.e.,* ≤ 35 g/L, or > 35 g/L) into a multivariate Cox model first, followed by stepwise regression which selects an independent variable automatically in order to form the highest precision and most informative model. The resultant regression coefficient and standard errors would give rise to the Wald statistic that evaluates whether a model parameter is significant. After that, a simple scoring system is developed as the weighted sum of those significant risk factors, of which the new weights were defined as the quotient (rounded to the nearest integer) of corresponding $\chi^2$ score from the stepwise selection process divided by the smallest

**Table 2. Procedure for HCC risk scoring development**

| Step | Description |
|---|---|
| 1 | Categorizing all continuous risk factors into clinically meaningful categorical variables |
| 2 | Implementing Fine-Gray subdistribution hazard model to model the cumulative incidence of the event of interest as the Cox proportional hazard model overestimates the risk rate |
| 3 | Assigning zero weights for reference levels of the categorical variable |
| 4 | Defining weights by estimated regression coefficients which are multiplied by 10 and rounded to the nearest integer |
| 5 | Deriving the optimal cut-off values of a HCC risk score by maximizing the Youden's index |

HCC: hepatocellular carcinoma

$\chi^2$ score among all those factors. The $\chi^2$ score for a given variable is the value of the likelihood score test for testing the significance of the variable. The weights can then be interpreted as a prioritization of all significant risk factors. In the CU-HCC score, albumin (+20 points) and cirrhosis (+15 points) are two heavily-weighted components; whereas in the LSM-HCC score, age (+10 points) and LSM (+8 points if 8-12 kPa, +14 points if > 12 kPa) contribute the most[19,20].

There are several summary measures for determining the optimal cut-off values of a risk score, including cost analysis, likelihood ratios, and receiver operating characteristic (ROC) analysis. The use of different cut-off methods depends greatly on the medical condition. The LSM-HCC score was categorized into low risk and high risk groups with a cut-off value of highest sum of sensitivity and specificity value, which is similar to the Youden's index (*i.e.,* Youden's J statistic = sensitivity + specificity - 1). There is a trade-off relationship between sensitivity and specificity- as one increases, the other decreases. In the two HCC risk scores we discussed, selecting a cut-point by maximizing true positive and negative rates is preferred over merely optimizing the sensitivity. The Youden's index is less sensitive than the one associated with only the sensitivity, which would not inflate the false positive rate too much and therefore avoid patients with low HCC risk suffering from unnecessary HCC treatment. Hence the health care resources would be more efficiently allocated and utilized in the medium- or high-risk group. One can define cut-points by $\chi^2$ test for monotonicity like the CU-HCC score, as the multivariate Cox proportion hazard model can be written as a linear model. The procedure for HCC risk scoring development is summarized in Table 2[22].

## STATISTICAL STRATEGIES FOR HCC RISK SCORES IN TREATED PATIENTS
### Examples: PAGE-B and mPAGE-B scores
Current first-line oral HBV antiviral treatment suppresses HBV DNA replication effectively and prevents disease progression in CHB patients, yet does not completely eliminate the risk of HCC development[23,24]. Motivated by the modest performance of untreated-derived risk scores on treated patients, especially among the Caucasian population[23], the PAGE-B score [Table 1] was developed to specifically predict the risk of HCC in NA-treated CHB patients[25,26]. Subsequently, Korean investigators modified the PAGE-B score by adding serum albumin for accurate prediction in the Asian treated CHB population[25]. These two scores have been externally validated in several independent cohorts and achieve good prediction performance[10,27,28]. Likewise, other HCC risk scores have been derived and validated for treated CHB patients[25,29-31].

The PAGE-B score is calculated by summing up integer points that correspond to particular categories of the included risk factors. Based on multivariable Cox proportional hazards model, the authors demonstrated that advanced age, male gender, and low platelet counts are the three key risk factors to predict HCC development in the coming five years[26]. Instead of relying on a complex Cox model-based equation, they adopted the method described by Sullivan *et al.*[32] on simplifying the equation to a so-called "points system", which aims at easy calculation without aid of a calculator. This method is done by organizing every significant covariate into meaningful categories by cut-offs, followed by determining

reference value of each category. These cut-offs are predefined based on previous literature and clinical knowledge, or driven by data. For a category of continuous covariates like age or platelet counts, reference value is chosen as the mid-point of that category, *e.g.,* 34.5 is the reference value of the age group of 30-39 years. For an open-ended category of continuous covariates, usually the first and last categories such as platelet $< 100,000/mm^3$ and $\geq 200,000/mm^3$, respectively, the 1st and the 99th percentile of platelet counts of all patients are used as the lower bound and upper bound for calculating mid-point for the first and last category, respectively, to minimize the influence of outliers. The reference value is set to be 0 for the reference group of categorical covariate, which is female gender in PAGE-B score; any other categories of the categorical covariate, i.e., male gender, are assigned with reference value of 1.

After assigning all reference values, a base category is selected as the reference category for each risk factor. Usually the category with the lowest risk is chosen as the base category. The base category has 0 points in the points system. Following that, it is to determine how far each category is from the base category in terms of regression coefficient estimated by the original multivariable Cox regression. For each category of a continuous covariate, the distance is calculated as the product of the regression coefficient, i.e., natural logarithm of the adjusted hazard ratio, and the numerical difference of the reference value of that category from the reference value of the base category. The distance of each category of categorical covariate from the base category is exactly the estimated regression coefficient of that category. After that, a constant that represents the number of regression units that will correspond to one point in the points system is chosen. Then the point of each category of each risk factor is equal to its calculated distance divided by the constant, rounded to the nearest integer. Finally, the HCC risk score is calculated as the sum of integer point of each category that a patient falls into.

## COMMON PITFALL IN THE DEVELOPMENT AND VALIDATION OF HCC RISK SCORES

Existing HCC risk scores were mostly developed using traditional regression methods, or to be specific, the Cox proportional hazards regression. A point system is usually adopted by giving integer points to categories of each risk factor. In the old days, it was reasonable to reduce the complex regression equation into discrete scoring system so that clinicians can use the score with ease. Yet, as a trade-off, continuous covariates have to be divided into categories. Statistically speaking, part of the information carried by the covariates can be lost through categorization. Also, the overall performance of the risk score will rely on the choice of cut-offs. Sometimes, the value of the covariates themselves, for example platelet counts, is more objective than the cut-off, especially if the cut-off may be estimated using your own data. Data-driven cut-offs for covariates may not be generalizable to other patient populations if there is some unmeasured difference between populations. With the advancement of technology, nowadays even complex equations can be easily calculated with the help of a computer next to the clinicians when they see their patients. All they need to do would be to input the value of every covariate to the computer, if not the computer does that for them automatically. It is expected that in the future, instead of a point system, complex equations that can achieve even higher accuracy derived by big data approaches including machine learning or deep learning algorithms would play a more important role in prediction of HCC.

After calculating the HCC risk score, researchers have to explain to clinicians and patients the meaning of the value. To deal with that, traditionally cut-offs for HCC risk score are determined based on diagnostic accuracy to classify patients into low, intermediate, and high risk of HCC development. The cumulative incidence of HCC in each risk stratum would then be estimated by survival analysis. A drawback of the current way of determining cut-off is that the criteria used do not suit the target, hence the limited use of HCC risk score in clinical practice. Indeed, most of the determined low cut-offs of existing HCC risk scores achieve a high NPV to exclude a meaningful proportion of patients with low HCC risk[32]. HCC risk scores have the potential to guide HCC surveillance in the clinical setting, especially among non-cirrhotic patients, by identifying patients who have a low HCC risk in the near future[10]. HCC risk scores can be

more useful if a low cut-off is selected based on the low annual incidence of HCC in the low risk group, for instance, less than the suggested threshold by the American Association for the Study of Liver Diseases for cost-effective HCC surveillance for CHB patients, i.e., 0.2%[33].

Missing data is perhaps another important issue in developing a risk score. Many HCC risk scores involve laboratory measurements that may be missed in some of the patients. If ignored, a risk score developed based on solely complete cases can introduce selection bias and affect the precision of the effect estimates. Missing data should be probably handled by statistical methods such as multiple imputation to avoid bias. It is worth noting that apart from the PAGE-B score, existing HCC risk scores usually did not state explicitly on how missing data are handled, which can potentially affect their generalizability.

## CONCLUSIONS AND FUTURE PERSPECTIVE

With the knowledge of common statistical tests and strategies which have been adopted in the various HCC prediction models, the future is directed towards a more personalised approach. Continuous optimisation of the predictive accuracy of the models will be achieved by involving more serial parameters, as well as on-treatment data in NA-treated patients. HCC risk levels may change over time, as patients are getting older, at the same time the natural history has been modified by NA treatment, which leads to viral suppression, improvement in liver biochemistry, as well as regression of cirrhosis. Hence, accurate models should be able to identify such bidirectional changes of HCC risk over time. Whilst accuracy remains the most important aspect of an ideal prediction model, applicability and usability is just and important in order to translate HCC risk into clinical practice. Prediction models may be built into the computer systems for patient management with automated retrieval of relevant clinical parameters. The most-updated HCC risk level would be able to guide the optimal HCC surveillance intervals or modalities, by providing timely alerts in the computer system.

## DECLRATIONS

### Authors' contributions
Responsible for the interpretation of data and critical revision of the manuscript: Yip TCF, Hui VWK, Tse YK, Wong GLH

### Availability of data and materials
Not applicable.

### Conflicts of interest
Yip TCF has served as a speaker for Gilead Sciences; Wong GLH has served as an advisory committee member for Gilead Sciences and Janssen; and as a speaker for Abbott, Abbvie, Bristol-Myers Squibb, Echosens, Gilead Sciences, Janssen and Roche; Hui VWK and Tse YK declared that there are no conflicts of interest.

### Ethical approval and consent to participate
Not applicable.

### Consent for publication
Not applicable.

## Copyright

## REFERENCES

1.  World Health Organization. Combating hepatitis B and C to reach elimination by 2030 - advocacy brief. Available from: http://www.who.int/hepatitis/publications/hep-elimination-by-2030-brief/en/. [Last accessed on 28 Sep 2020]

2.  Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018 68:394-424.

3.  Sarin SK, Kumar M, Eslam M, et al. Liver diseases in the Asia-Pacific region: a lancet gastroenterology & hepatology commission. *Lancet Gastroenterol Hepatol* 2020;5:167-228.

4.  Yip TC, Liang LY, Wong GL. Assessment of HCC risk in patients with chronic HBV (REACH, PAGE-B, and beyond). *Curr Hep Rep* 2020; doi: 10.1007/s11901-020-00526-w.

5.  Chan SL, Wong AM, Lee K, Wong N, Chan AK. Personalized therapy for hepatocellular carcinoma: where are we now? *Cancer Treat Rev* 2016;45:77-86.

6.  Polaris Observatory C. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. *Lancet Gastroenterol Hepatol* 2018;3:383-403.

7.  Wong GL, Wong VW, Tan GM, et al. Surveillance programme for hepatocellular carcinoma improves the survival of patients with chronic viral hepatitis. *Liver Int* 2008;28:79-87.

8.  Omata M, Cheng AL, Kokudo N, et al. Asia-Pacific clinical practice guidelines on the management of hepatocellular carcinoma: a 2017 update. *Hepatol Int* 2017;11:317-70.

9.  Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR. A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* 2019;16:589-604.

10. Yip TC, Wong GL, Wong VW, et al. Reassessing the accuracy of PAGE-B-related scores to predict hepatocellular carcinoma development in patients with chronic hepatitis B. *J Hepatol* 2020;72:847-54.

11. Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;35:214-26.

12. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol* 1996;49:907-16.

13. Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;21:45-56.

14. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56:337-44.

15. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;30:1105-17.

16. Gronnesby JK, Borgan O. A method for checking regression models in survival analysis based on the risk score. *Lifetime Data Anal* 1996;2:315-28.

17. May S, Hosmer DW. A cautionary note on the use of the Gronnesby and Borgan goodness-of-fit test for the Cox proportional hazards model. *Lifetime Data Anal* 2004;10:283-91.

18. Meltzer PS, Kallioniemi A, Trent JM. Chromosome alterations in human solid tumors. In: Vogelstein B, Kinzler KW, editors. The genetic basis of human cancer. New York: McGraw-Hill; 2002. pp. 93-113.

19. Wong GL, Chan HL, Wong CK, et al. Liver stiffness-based optimization of hepatocellular carcinoma risk score in patients with chronic hepatitis B. *J Hepatol* 2014;60:339-45.

20. Wong VW, Chan SL, Mo F, et al. Clinical scoring system to predict hepatocellular carcinoma in chronic hepatitis B carriers. *J Clin Oncol* 2010;28:1660-5.

21. Wong GL. Non-invasive assessments for liver fibrosis: the crystal ball we long for. *J Gastroenterol Hepatol* 2018;33:1009-15.

22. Austin PC, Lee DS, D'Agostino RB, Fine JP. Developing points-based risk-scoring systems in the presence of competing risks. *Stat Med* 2016;35: 4056-72.

23. Arends P, Sonneveld MJ, Zoutendijk R, et al. Entecavir treatment does not eliminate the risk of hepatocellular carcinoma in chronic hepatitis B: limited role for risk scores in Caucasians. *Gut* 2015;64:1289-95.

24. Yip TC, Wong GL, Chan HL, et al. HBsAg seroclearance further reduces hepatocellular carcinoma risk after complete viral suppression with nucleos(t)ide analogues. *J Hepatol* 2019;70:361-70.

25. Kim JH, Kim YD, Lee M, et al. Modified PAGE-B score predicts the risk of hepatocellular carcinoma in Asians with chronic hepatitis B on antiviral therapy. *J Hepatol* 2018;69:1066-73.

26. Papatheodoridis G, Dalekos G, Sypsa V, et al. PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy. *J Hepatol* 2016;64:800-6.

27. Lee HW, Kim SU, Park JY, et al. External validation of the modified PAGE-B score in Asian chronic hepatitis B patients receiving antiviral therapy. *Liver Int* 2019;39:1624-30.

28. Kirino S, Tamaki N, Kaneko S, et al. Validation of hepatocellular carcinoma risk scores in Japanese chronic hepatitis B cohort receiving nucleot(s)ide analog. *J Gastroenterol Hepatol* 2020;35:1595-601.

29. Hsu YC, Yip TC, Ho HJ, et al. Development of a scoring system to predict hepatocellular carcinoma in Asians on antivirals for chronic hepatitis B. *J Hepatol* 2018;69:278-85.
30. Lee HW, Park SY, Lee M, et al. An optimized hepatocellular carcinoma prediction model for chronic hepatitis B with well-controlled viremia. *Liver Int* 2020;40:1736-43.
31. Lee HW, Yoo EJ, Kim BK, et al. Prediction of development of liver-related events by transient elastography in hepatitis B patients with complete virological response on antiviral therapy. *Am J Gastroenterol* 2014;109:1241-9.
32. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: the framingham study risk score functions. *Stat Med* 2004;23:1631-60.
33. Terrault NA, Lok ASF, McMahon BJ, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. *Hepatology* 2018;67:1560-99.