Research Article

JOURNAL OF HEPATOLOGY

PAGE-B predicts the risk of developing hepatocellular carcinoma in Caucasians with chronic hepatitis B on 5-year antiviral therapy

George Papatheodoridis[1,2,*], George Dalekos[3], Vana Sypsa[4], Cihan Yurdaydin[5], Maria Buti[6], John Goulis[7], Jose Luis Calleja[8], Heng Chi[9], Spilios Manolakopoulos[2], Giampaolo Mangia[10], Nikolaos Gatselis[3], Onur Keskin[5], Savvoula Savvidou[7], Juan de la Revilla[8], Bettina E. Hansen[9], Ioannis Vlachogiannakos[1], Kostantinos Galanis[3], Ramazan Idilman[5], Massimo Colombo[10], Rafael Esteban[6], Harry L.A. Janssen[9,11], Pietro Lampertico[10]

[1]*Department of Gastroenterology, School of Medicine, National and Kapodistrian University of Athens, Laiko General Hospital, Athens, Greece;* [2]*2nd Department of Internal Medicine, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece;* [3]*Department of Internal Medicine, Thessalia University Medical School, Larissa, Greece;* [4]*Department of Hygiene, Epidemiology & Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece;* [5]*Department of Gastroenterology, University of Ankara Medical School, Ankara, Turkey;* [6]*Hospital General Universitario Valle Hebron and Ciberehd, Barcelona, Spain;* [7]*4th Department of Internal Medicine, Aristotle University of Thessaloniki Medical School, Thessaloniki, Greece;* [8]*Hospital U Puerta de Hierro, IDIPHIM CIBERehd, Madrid, Spain;* [9]*Department of Gastroenterology & Hepatology, ErasmusMC, Rotterdam, Netherlands;* [10]*Division of Gastroenterology and Hepatology, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milano, Italy;* [11]*Liver Clinic, Toronto Western & General Hospital, University Health Network, Toronto, ON, Canada*

Background & Aims: Risk scores for hepatocellular carcinoma (HCC) developed in Asians offer poor-moderate predictability in Caucasian patients with chronic hepatitis B (CHB). This nine center cohort study aimed to develop and validate an accurate HCC risk score in Caucasian CHB patients treated with the current oral antivirals, entecavir/tenofovir.

Methods: We included 1815 adult Caucasians with CHB and no HCC at baseline who received entecavir/tenofovir for ⩾12 months. Using data from eight centers (derivation dataset, n = 1325), a HCC risk score was developed based on multivariable Cox models and points system for simplification. Harrell's c-index was used as discrimination, bootstrap for internal validation and the data from the 9th and largest center (validation dataset, n = 490) for external validation.

Results: The 5-year cumulative HCC incidence rates were 5.7% and 8.4% in the derivation and validation dataset, respectively. In the derivation dataset, age, gender, platelets and cirrhosis were independently associated with HCC. The PAGE-B score was developed based on age, gender and platelets (c-index = 0.82, 0.81 after bootstrap validation). The addition of cirrhosis did not substantially improve the discrimination (c-index = 0.84). The predictability of PAGE-B score was similar (c-index = 0.82) in the validation dataset. Patients with PAGE-B ⩽9, 10–17, ⩾18 had 5-year cumulative HCC incidence rates of 0%, 3%, 17% in the derivation and 0%, 4%, 16% in the validation dataset.

Conclusion: PAGE-B, which is based only on baseline patients' age, gender and platelets, represents a simple and reliable score for prediction of the 5-year HCC risk in Caucasian CHB patients under entecavir/tenofovir.

Introduction

Monotherapy with one of the current first-line oral nucleos(t)ide analogues (NAs), entecavir (ETV) and tenofovir disoproxil fumarate (TDF), results in long-term inhibition of hepatitis B virus (HBV) replication in almost all compliant patients with chronic hepatitis B (CHB), improves liver histological lesions, often achieves regression of cirrhosis, prevents or reverses hepatic decompensation, diminishes the need for liver transplantation and improves the overall survival [1]. However, hepatocellular carcinoma (HCC) still develops in CHB patients treated with NA(s) regardless of virological response [2–4] representing the major complication and a key challenge in the management of CHB patients.

Given that the early diagnosis of HCC increases the applicability of curative therapies and eventually the patients' prognosis [5], the identification and close surveillance of CHB patients at high risk for HCC is of great importance. Most of the HCC data in CHB come from cohort studies including untreated patients or patients treated with lamivudine and/or adefovir and more recently ETV [2,3,6–9]. Recently, risk scores (GAG-HCC, CU-HCC

Keywords: Hepatocellular carcinoma; Hepatitis B; Entecavir; Tenofovir.
Received 30 June 2015; received in revised form 25 November 2015; accepted 26 November 2015; available online 8 December 2015

* Corresponding author. Address: School of Medicine, National and Kapodistrian University of Athens, Laiko General Hospital of Athens, 17 Agiou Thoma Street, 11527 Athens, Greece. Tel.: +30 2107456513; fax: +30 2107462601.
E-mail address: gepapath@med.uoa.gr (G. Papatheodoridis).

Abbreviations: NA(s), nucleos(t)ide analogue(s); ETV, entecavir; TDF, tenofovir disoproxil fumarate; HBV, hepatitis B virus; CHB, chronic hepatitis B; HCC, hepatocellular carcinoma; ALT, alanine aminotransferase; IQR, interquartile range; HR, hazard ratio; CI, confidence interval(s).

and REACH-B) for prediction of HCC were developed and validated in cohorts of untreated Asian CHB patients [7–9], while their predictability was subsequently confirmed in Asian patients treated with entecavir [10]. We and others, however, have shown that the predictability of these HCC risk scores is poor to moderate in Caucasian CHB patients, for whom different risk scores seem to be required [11,12].

The aim of this large, multicenter, cohort study was to develop and validate an accurate HCC risk score in Caucasian CHB patients treated with the currently recommended oral antivirals, ETV or TDF.

Patients and methods

Patient population

This study was based on two datasets of Caucasian CHB patients selected by the same criteria from nine participating centers, as it has been previously described [11,12]. The derivation dataset including patients from eight centers was used as the training dataset to derive a score in predicting HCC, whereas the validation dataset including patients from the largest center (Milano, Italy) was used for external validation of the scoring system. All patients with CHB followed in the liver clinics of the nine participating centers were included if they were adults ($\geqslant$16 years old), Caucasians and had received treatment with ETV or TDF for $\geqslant$12 months. The participating centers were in Greece (Athens [2 centers], Larissa, Thessaloniki), Italy (Milano), Spain (Barcelona, Madrid), Netherlands (Rotterdam) and Turkey (Ankara). Patients naive to or previously treated with other NAs were included. Patients with decompensated cirrhosis, HCC diagnosed before the onset of ETV/TDF, patients with co-infection(s) with hepatitis D, hepatitis C or human immunodeficiency virus and liver transplant patients were excluded.

Follow-up – Definitions

CHB was diagnosed in patients with positive HBsAg for $\geqslant$6 months, elevated alanine aminotransferase (ALT) and serum HBV DNA >2000 IU/ml. Patients were classified according to their liver disease severity into: a) patients with CHB only (without cirrhosis) if they had a pretreatment liver biopsy without lesions of cirrhosis; and b) patients with compensated cirrhosis if they had histological findings and/or ultrasonographic findings (nodules in the hepatic parenchyma, spleen >12 cm, portal vein >16 mm) and/or endoscopic findings of cirrhosis (varices, portal gastropathy). Patients without a pretreatment liver biopsy and without any other sign of cirrhosis were considered as cases with unclassified disease severity.

All patients were treated with ETV and/or TDF and followed at each participating center according to international and/or national clinical practice guidelines. Clinical examination and routine laboratory tests were performed at least every 6 months. HBV DNA levels were determined every 6–12 months at the laboratory of each center by various polymerase chain reaction assays (sensitivities: 10–80 IU/ml). Virological remission was considered to be present in patients who achieved HBV DNA <80 IU/ml that was maintained throughout ETV/TDF therapy. Ultrasonography and/or alpha-fetoprotein levels were performed every 6 months in cirrhotic and every 12 months in non-cirrhotic patients. The diagnosis of HCC was based on standard histological and/or compatible radiological findings [5].

Entry into this study (baseline) was defined as the date of the onset of ETV/TDF. Follow-up was considered as the time interval between the study entry and the last available clinical information until May 2014, while treatment duration was considered the time interval between the study entry until the end of therapy or the last on-therapy follow-up. Analysis time was the time interval between the study entry and the diagnosis of HCC or the end of follow-up in the absence of HCC development.

Statistical analysis

All data were entered into and analyzed using the statistical package Stata 11.2 (StataCorp LP, USA) and R (version 3.2.1). Continuous variables are presented by their median values and interquartile range (IQR), unless otherwise stated. Their comparison was performed by the non-parametric Mann-Whitney *U*-test. The chi-squared or Fisher's exact test was used for comparisons of categorical variables. The cumulative probabilities of HCC occurrence were estimated by the Kaplan-Meier method and compared with the log-rank test. Univariable and multivariable Cox proportional hazards regression models were used to estimate the effect of various variables on the hazard of HCC occurrence. Hazard ratios (HR) and their 95% confidence intervals (CI) along with corresponding *p* values are presented. A *p* value of <0.05 was considered to be statistically significant. The proportional hazards assumption was tested on the basis of Schoenfeld residuals.

The prediction model was developed to predict the occurrence of HCC within 5 years after ETV/TDF initiation. The development of our HCC risk score was based on a multivariable Cox proportional hazards model using data from eight centers (derivation dataset). We accounted for the observed follow-up time for the patients from these centers up to the 75th percentile (5 years) to avoid the likely influence of a small number of participants with longer follow-up duration on model estimates. To develop the prediction model, we have used multiple imputation to deal with missing data in candidate predictor variables [13]. We imputed 10 values of the missing predictor for each patient. We applied backward elimination to each of the 10 completed data sets separately, resulting in 10 sets of selected predictors. The final set comprised those predictors that were selected in more than 50% of the 10 data sets. Given the finally selected predictors, a model was fitted in each of the 10 completed data sets. We used Rubin's rules to combine the estimated regression coefficients and variances from the 10 different completed data sets. To evaluate the predictive performance of the model, we examined discrimination and calibration measures. Discrimination was assessed using Harrell's c-index. A calibration plot was used to assess graphically the agreement between the 5-year probability of remaining HCC free as predicted by the model *vs.* the Kaplan-Meier estimate (observed probability). Per quintile of predicted probabilities, the Kaplan-Meier estimate and standard error were determined [14].

We performed validity assessment of the model using internal and external validation. Internal validation was performed using bootstrap. Bootstrap samples are random samples drawn with replacement from the original sample. We repeatedly fitted the model in 1000 bootstrap samples and evaluated its performance on the original sample. In external validation, the model developed in the derivation dataset was applied on the patients of the 9th (largest) center (validation dataset). The predictive performance of the model was assessed in the validation dataset as in the derivation dataset.

The next step was to develop a risk score based on a points system to simplify the computation of HCC risk estimate [15] (Supplementary material). We evaluated the agreement between risk estimates based on the points system and on the multivariable model (risk categories: <2%, 2–8.9%, $\geqslant$9%) using weighted kappa.

We assessed the discrimination and the calibration of the risk score in the derivation and the validation datasets by inspection of the Kaplan-Meier curves for risk groups stratified by the 25th and 75th percentiles of the risk score distribution [16].

We estimated the sensitivity, specificity, positive predictive value and negative predictive value (NPV) for various cut-offs of the risk score using appropriate methodology for censored data [17].

Results

There were 1325 patients in the derivation and 490 patients in the validation dataset. The patients in the two datasets differ in most of their characteristics at the onset of ETV/TDF (Table 1). The diagnosis of cirrhosis was based on histological findings before antiviral therapy in 172/269 (64%) and 164/234 (70%) cirrhotic patients in the derivation and validation datasets, respectively. Virological remission was achieved in 89% and 96% of patients in the derivation and validation datasets at year-1 of ETV/TDF therapy (p <0.001) and in 92% and 97% of patients beyond the first year of therapy, respectively (p <0.001). The median serum HBV DNA levels in patients without virological remission at year-1 in the derivation and validation datasets were 1000 (IQR: 6062) and 292 (913) IU/ml, respectively.

During a median follow-up of 50 (31–62) months, HCC was diagnosed in 51 (3.8%) patients in the derivation and 34 (6.9%) patients in the validation dataset. The cumulative 1-, 3- and 5-year rates of HCC were 0.9%, 3.1% and 5.7% in the derivation and 1.2%, 3.9% and 8.4% in the validation dataset, respectively (p = 0.108) (Fig. 1).