Predictive Factors for Hepatocellular Carcinoma in Chronic Hepatitis B Using Structural Equation Modeling: A Prospective Cohort Study

Laurent Lam, Hélène Fontaine, Marc Bourliere, Clovis Lusivika-Nzinga, Céline Dorival, Dominique Thabut, Fabien Zoulim, François Habersetzer, Tarik Asselah, Jean-Charles Duclos-Vallee, et al.

Predictive Factors for Hepatocellular Carcinoma in Chronic Hepatitis B Using Structural Equation Modeling: A Prospective Cohort Study

Short title: Structural Models for Hepatitis B Virus

Laurent LAM[a,b], Hélène FONTAINE[c], Marc BOURLIERE[d], Clovis LUSIVIKA- NZINGA[b], Céline DORIVAL[b], Dominique THABUT[e], Fabien ZOULIM[f], François HABERSETZER[g], Tarik ASSELAH[h], Jean-Charles DUCLOS-VALLEE[i], Jean-Pierre BRONOWICKI[j], Philippe MATHURIN[k], Thomas DECAENS[l], Nathalie GANNE[m], Dominique GUYADER[n], Vincent LEROY[o], Isabelle ROSA[p], Victor DE LEDINGHEN[q], Paul CALES[r], Xavier CAUSSE[s], Dominique LARREY[t], Olivier CHAZOUILLERES[u], Moana GELU-SIMEON[v], Véronique LOUSTAUD-RATTI[w], Sophie METIVIER[x], Laurent ALRIC[y], Ghassan RIACHI[z], Jérôme GOURNAY[aa], Anne MINELLO[ab], Albert TRAN[ac], Claire GEIST[ad], Armand ABERGEL[ae], François RAFFI[af], Louis D'ALTEROCHE[ag], Isabelle PORTAL[ah], Nathanaël LAPIDUS[a,b], Stanislas POL[c,ai], Fabrice CARRAT[a,b], for the ANRS/AFEF Hepather study group

[a] Assistance Publique-Hôpitaux de Paris. Sorbonne Université, Department of Public Health, Hôpital Saint-Antoine, Paris, France.

[b] Sorbonne Université, INSERM, Institut Pierre Louis d'Épidémiologie et de Santé Publique, Paris, France.

[c] Department of Hepatology, Assistance Publique-Hôpitaux de Paris, Hôpital Cochin, Paris, France.

[d] Department of Hepatology and Gastroenterology, hôpital Saint-Joseph, Marseille, France.

[e] Sorbonne Université, Department of Hepatology and Gastroenterology, Groupe Hospitalier Pitié-Salpétrière, AP-HP, INSERM UMR-S938, Paris, France.

[f] Department of Hepatology, Hospices Civils de Lyon, INSERM U1052, Université de Lyon, Lyon, France.

[g] CIC, Inserm 1110 et Pôle Hépato-digestif des Hôpitaux Universitaires de Strasbourg, Université de Strasbourg, Strasbourg, France.

[h] INSERM UMR 1149, Hepatology, Hospital Beaujon, Centre de Recherche sur l'Inflammation, (CRI), University Paris Diderot, Clichy, France.

[i] AP-HP Hôpital Paul-Brousse, Centre Hépato-Biliaire, Villejuif; UMR-S 1193, Université Paris-Saclay; DHU HEPATINOV, Villejuif, France.

[j] Inserm U1254 and Department of Hepato-Gastroenterology, University Hospital of Nancy Brabois, Université de Lorraine, Vandoeuvre-les-Nancy, France.

[k] Service des maladies de l'appareil digestif, Université Lille 2 and Inserm U795, France.

[l] Department of Hepatology and Gastroenterology, Centre Hospitalo-Universitaire, INSERM U1209, Université Grenoble Alpes, Grenoble, France.

[m] Department of Hepatology, Hôpitaux Universitaires Paris Seine-Saint-Denis, site Jean Verdier, AP-HP, Bondy, France; Université Paris 13, Sorbonne Paris Cité et INSERM UMR 1162, Paris France.

[n] CHU de Rennes, service d'hépatologie; Univ Rennes 1, Inra, Inserm, Institut NUMECAN (Nutrition, Métabolismes et Cancer), UMR A 1341, UMR S 1241, F-35033 Rennes, France.

[o] Department of Hepatology and Gastroenterology, Hôpital Henri Mondor, AP-HP, Université Paris-Est, INSERM U955, Créteil, France.

[p] Department of Hepatology and Gastroenterology, Centre Hospitalier Intercommunal, Créteil, France.

[q] Hepatology Unit, University Hospital Bordeaux and INSERM U1053, Bordeaux University, Pessac, France.

[r] Hepatology Department, University Hospital, Angers, France; HIFIH Laboratory, Angers University, Angers, France.

[s] Department of Hepatology and Gastroenterology, CHR Orléans, France.

[t] Liver Unit-IRB-INSERM 1183, Hôpital Saint Eloi, Montpellier, France.

[u] Department of Hepatology, Hôpital Saint-Antoine, AP-HP, Sorbonne université, Paris, France.

[v] Service d'Hépato-Gastroentérologie, CHU de la Guadeloupe - Faculté de médecine, Université des Antilles, Pointe-à-Pitre Cedex, F-97110, France - INSERM,UMR-S1085/IRSET, F-35043 Rennes, France.

[w] Department of Hepatology and Gastroenterology, CHU Limoges, U1248 INSERM, Univ. Limoges, F-87000 Limoges, France.

[x] Hepatology unit, CHU Rangueil, 31059 Toulouse, France.

[y] Department of Internal Medicine and Digestive Diseases, CHU Purpan, UMR 152 Pharma Dev, IRD Toulouse 3 University, France.

[z] Department of Hepatology and Gastroenterology, CHU Charles Nicolle, Rouen, France.

[aa] Gastroenterology and Hepatology Department, Institut des Maladies de l'Appareil Digestif, University Hospital of Nantes, Nantes, France.

[ab] Department of Hepatology and Gastroenterology, University hospital Dijon, INSERM UMR 1231, France.

[ac] Digestive Center, Centre Hospitalier Universitaire de Nice, INSERM U1065-8, Nice, France.

[ad] Department of Hepatology and Gastroenterology, Centre Hospitalier Régional, Metz, France.

[ae] Department of Digestive and Hepatobiliary Diseases, Estaing University Hospital, Clermont-Ferrand, France. UMR 6602 CNRS-Sigma-Université Clermont Auvergne, Clermont- Ferrand, France.

[af] Department of Infectious Diseases, Hotel-Dieu Hospital - INSERM CIC 1413, Nantes University Hospital, Nantes, France.

[ag] Unit of Hepatology, Hépatogastroentérologie, CHU Trousseau, 37044 Tours, France.

[ah] Service d'hépato-gastroentérologie, hôpital de la Timone, Aix-Marseille université, AP-HM, Marseille, France.

[ai] Université de Paris; Inserm U-1223 and ICD, Institut Pasteur, Paris, France.

Corresponding author:

Pr. Fabrice Carrat, MD, PhD

Institut Pierre Louis d'Épidémiologie et de Santé Publique

27, rue Chaligny

75571, Paris Cedex 12, France

Tel: +33(0)144738643

Email: fabrice.carrat@iplesp.upmc.fr

Electronic word count:

Abstract: 250 words

Manuscript (including the figure and table legends): 3,999 words

References: 50 references

Tables and Figures:

Number of Main Tables: 3

Number of Main Figures: 3

Conflict of interest statement

HF reports personal fees and invitations for medical meeting from Gilead, Abbvie, BMS, MSD, Janssen, MSD outside this work. MB reports grants and personal fees from AbbVie, grants and personal fees from Gilead, personal fees from MSD, personal fees from Janssen, personal fees from Boehringher Ingelheim, personal fees from intercept, personal fees from BMS, outside the submitted work. DT reports personal fees from MSD, Abbvie, Gilead Sciences, outside the submitted work. FZ reports personal fees from Abbvie, Gilead, during the conduct of the study. FH reports personal and travel fees from Abbvie and Gilead, during the conduct of the study. TA reports grants from ANRS-INSERM, personal fees from ABBVIE, GILEAD, JANSSEN, MSD, outside the submitted work. J-PB reports personal fees from Abbvie, Gilead, MSD, BMS, JANSSEN outside the submitted work. PM reports personal fees from MSD, Abbvie, Gilead Sciences, outside the submitted work. TD reports personal fees from Bayer, MSD, Abbvie, Gilead Sciences, outside the submitted work. NG reports grants from Echosens, personal fees and non-financial support from Gilead Sciences, personal fees and non-financial support from Abbvie, non-financial support from MSD, personal fees and non-financial support from Bayer, outside the submitted work. DG reports personal fees and non-financial support from Gilead sciences, MSD, and Abbbvie, grants, personal fees and non-financial support from Janssen, during the conduct of the study. VL reports grants, personal fees and non-financial support from Abbvie, grants, personal fees and non-financial support from BMS, grants, personal fees and non-financial support from Gilead, personal fees and non-financial support from MSD, personal fees from echosens, outside the submitted work. VdL has received consulting and/or lecturing fees from Gilead, AbbVie, Echosens, Intercept Pharma, SuperSonic Imagine, Indivior, Spimaco, Pfizer, Bristol Myers Squibb, Myr-Pharma. XC reports personal fees from AbbVie, Gilead, Intercept, Janssen, MSD, during the conduct of the study, outside the submitted work. OC reports personal fees

from Abbvie, Gilead Sciences, Intercept, Pliant Therapeutics , outside the submitted work. MG-S reports invitations for medical meeting from Gilead, Abbvie. VL-R is a Clinical Investigator/Speaker for AbbVie, Gilead Sciences, Merck Sharp & Dohme. LA reports personal fees from MSD, Abbvie, Gilead Sciences, outside the submitted work. GR reports personal fees from MSD, Abbvie, Gilead Sciences, outside the submitted work. JG reports personal fees from Gilead, MSD, Abbvie, during the conduct of the study; personal fees from Intercept, outside the submitted work. AM reports personal fees from for Gilead, Abbvie, Intercept, during the conduct of the study, outside the submitted work. AA reports personal fees from AbbVie, Gilead Sciences, Merck Sharp & Dohme, Intercept, during the conduct of the study, outside the submitted work. FR reports personal fees from for Gilead, ViiV, Abbvie, MSD, Janssen during the conduct of the study, outside the submitted work. LDA reports invitations for medical meeting from Gilead, Abbvie, BMS, MSD and Janssen outside this work. IP reports personal fees from MSD, Abbvie, Gilead Sciences, outside the submitted work. SP received consulting and lecturing fees from Bristol-Myers Squibb, Janssen, Gilead, Roche, Biotest, Shinogi, ViiV, MSD and Abbvie, and grants from Bristol-Myers Squibb, Gilead and MSD. FC reports grants from INSERM-ANRS, during the conduct of the study; personal fees from Imaxio, outside the submitted work. LL, CL-N, CD, J-CD-V, CG, PC, DL, SM, AT, IR and NL have nothing to disclose.

Financial support statement: The ANRS CO22 HEPATHER cohort is sponsored and funded by INSERM- ANRS and conducted in collaboration with Association Française pour l'étude du Foie (AFEF). The cohort received supports from ANR (Agence Nationale de la Recherche), DGS (Direction Générale de la Santé), MSD, Janssen, Gilead, Abbvie, BMS, and Roche. The public/private partnership is built in total transparency through a specific contract. The pharmaceutical companies are not involved in scientific decisions. The sponsor contributed to the study design and drafting of the study. The sponsor played no role in data

collection, data analysis or data interpretation. The other funding sources played no role in study design, data collection, data analysis, data interpretation, or drafting the study. FC had full access to all data in the study and made the final decision to submit the study for publication.

Author contributions

Study idea and design: LL and FC. Data acquisition: DT, FZ, FH, TA, J-CD-V, J-PB, PM, TD, NG, DG, VL, IR, MB, VdL, PC, XC, DL, OC, MG-S, VL-R, SM, LA, GR, JG, AM, AT, CG, AA, FR, LDA, IP, MB, CL- N and CD. Statistical analysis: LL and FC. Data analysis and data interpretation: LL, FC, HF, SP and NL. Manuscript writing : LL, FC, HF and SP wrote the report, and all authors reviewed the manuscript for important intellectual content. All authors approve the final version of the submission.

Lay summary:

- We explored the pathways leading to hepatocellular carcinoma in hepatitis B disease.
- The relationships between clinical characteristics and the cancer were analyzed.
- Liver fibrosis has a key role in the occurrence of hepatocellular carcinoma.
- The risk of occurrence is increased among elderly men with a metabolic syndrome.
- Excessive alcohol consumption is also associated with a higher risk.

Abstract

Background & Aims: The factors predicting hepatocellular carcinoma (HCC) occurrence in chronic hepatitis B need to be precisely known to improve its detection. We identified

pathways and individual predictive factors associated with HCC in the ANRS CO22 HEPATHER cohort.

Methods: The study analyzed HBV-infected patients recruited at 32 French expert hepatology centers from August 6, 2012, to December 31, 2015. We excluded patients with chronic HCV, HDV and a history of HCC, decompensated cirrhosis or liver transplantation. Structural equation models were developed to characterize the causal pathways leading to HCC occurrence. The association between clinical characteristics (age, gender, body-mass index, liver fibrosis, alcohol consumption, smoking status, diabetes, hypertension, dyslipidemia, alpha-fetoprotein, HBV DNA levels, antiviral therapy) and incident HCC was quantified.

Results: Among the 4,489 patients included, 33 patients reported incident HCC. The median follow-up was 45.2 months. Age (β = 0.18 by decade, 95% CI 0.14-0.23), male gender (β = 0.23, 95% CI 0.18-0.29), metabolic syndrome (β = 0.28, 95% CI 0.22-0.33), alcohol consumption (β = 0.09, 95% CI 0.05-0.14) and HBV DNA (β = 0.25, 95% CI 0.17-0.34) had a significant and direct effect on the occurrence of advanced liver fibrosis. Liver fibrosis (β = 0.71, 95% CI 0.55-0.87) predicted, in turn, the occurrence of HCC.

Conclusions: Liver fibrosis mediates the effects of age, gender, alcohol, metabolic syndrome and HBV DNA on the occurrence of HCC. Elderly men with chronic hepatitis B, risky alcohol use, advanced liver fibrosis, metabolic syndrome and high HBV DNA levels should be monitored closely to detect the development of HCC.

Keywords: Carcinoma, Hepatocellular - Hepatitis B virus – Hepatocarcinogenesis - Liver Cancer – Epidemiology

Abbreviations: AFEF, Association Française pour l'étude du Foie; AFP, alpha-fetoprotein; ALP, alkaline phosphatase; ANR, Agence Nationale de la Recherche; ANRS, Agence

nationale de recherche sur le sida et les hépatites virales; AST = aspartate aminotransferase ; BMI, body-mass index; CFI, comparative fit index; CPP, Comités de protection des personnes; DGS, Direction Générale de la Santé; DNA, deoxyribonucleic acid; EASL, European Association for the Study of the Liver; ENT, entecavir; FIB-4 index, Fibrosis-4; GGT, gamma-glutamyl transferase; HBeAg, Hepatitis B e antigen; HBV, hepatitis B virus; HCC, Hepatocellular carcinoma; HCV, hepatitis C virus; HDL, high-density lipoprotein; HIV, human immunodeficiency virus; HTN, hypertension; INR, International Normalised Ratio; INSERM, Institut national de la santé et de la recherche médicale; MELD, model for end-stage liver disease; NAFLD, non-alcoholic fatty liver disease; NNFI, Non-Normed Fit Index; PAL, alkaline phosphatase; PNPLA3, patatin-like phospholipase domain-containing 3; RMSEA, Root Mean Square Error of Approximation absolute adequacy index; SEM, structural equation model; SRMR, Standardized Root Mean square Residual; TDF, tenofovir disoproxil fumarate; ULN, upper limit of normal

Introduction

Nearly 250 million people are infected with the hepatitis B virus (HBV) worldwide with significant geographical variations in prevalence [1,2]. Hepatocellular carcinoma (HCC) is the most common primary liver cancer and is one of the leading causes of incident cancer and deaths worldwide [3]. HBV and HCV infections are the main risk factors for HCC [4–6], the leading cause of death from HCC and up to 10% of liver transplant causes. Chronic HBV infection is found in about 50-55% of HCC cases [7]. Seventy to 80% of HCC in chronic HBV-infected patients occur following a cirrhosis although HCC can also occur in chronic HBV patients without cirrhosis [8]. As such, the HBV is one of the oncogenic viruses whose oncogenic mechanisms are well-established. The integration of the HBV viral genome into the DNA of the liver host cell is one of the main mechanisms involved in the occurrence of HCC [9]. Other factors typically associated with the development of HCC include cirrhosis [10], autoimmune hepatitis, alcoholic steatosis, HIV [11], HBV/HCV co-infection [12], aflatoxin B1 [13], hemochromatosis [6], Budd-Chiari syndrome [10] and diabetes [14]. Age, gender [15], certain ethnic groups [16], HBV DNA levels [17], family history of HCC as well as lifestyle factors such as excessive alcohol consumption [10] and smoking [18] are also identified risk factors for HCC. Treatment with nucleos(t)ide analogue agents has been shown to decrease the risk of HCC in chronic HBV-infected patients [19], however a lower risk of HCC with tenofovir disoproxil fumarate (TDF) treatment compared to entecavir (ENT) is still a controversial issue [20,21].

In HCC, surgical resection, radiological thermo-ablation and liver transplantation are the only curative therapies [22], however, patients with advanced, bilobal or multifocal HCC often cannot benefit from these therapies due to the diagnosis frequently occurring at a late stage [23].

To improve patient survival and the chances of curative therapies, early detection of HCC is required [22]. The risk factors for HCC and their interrelationships need to be more precisely described to enhance the performance of this early detection. However, in chronic HBV-infected patients, a model describing how clinical and biological determinants can contribute to HCC is still lacking. Structural equation modeling would allow the study of the complex relationships between factors related to the development of HCC and therefore provide a better understanding of the underlying pathways.

The objectives of this study are: 1) to propose a Structural Equation Model (SEM) that schematizes the causal relationships between the characteristics of chronic HBV-infected patients and the occurrence HCC, including direct and indirect (i.e. mediated) effects, 2) to identify and quantify the risks of incident HCC associated with the clinical, biological and therapeutic characteristics of HBV-infected patients in a French cohort study.

Materials and Methods

Data source and analysis population

The data source was based on the analysis of the ANRS CO22 HEPATHER cohort, promoted by the ANRS and initiated in 2012, which aims to improve knowledge about viral hepatitis [24]. This cohort allows for a multicenter observational study with prospective data collection and biological collections of patients who have had hepatitis B and/or C. It includes 21,277 patients with hepatitis B or C. Thirty-two hospital clinical centers, spread over the French territory, are involved in data collection and the constitution of biological collections. The criteria for non-inclusion in the HEPATHER cohort are as follows: 1) HIV coinfection; 2) so-called vulnerable population (minors, persons under guardianship or tutelage, or persons deprived of their liberty by a judicial or administrative decision, treatment against hepatitis C in progress or arrested for less than 3 months, end-of-life patients); 3) women whose pregnancy is known. Our study included all HBV-infected patients from the ANRS CO22 HEPATHER cohort included from August 6, 2012, to December 31, 2015. We selected all HBV-infected patients (n=6,248) and excluded patients with the following co-morbidities: 1) HCV or HDV co-infection; 2) past or present HCC diagnosed at inclusion; 3) past or present episode of decompensated cirrhosis 4) liver transplant recipients; 5) acute HBV infection or cured HBV infection; 6) HBV chronic infection according to EASL definition due to the different prognostic and management of this group. Written informed consent was obtained from each patient before inclusion. The protocol was carried out in accordance with the Declaration of Helsinki and the French law on biomedical research and was approved by the Ethics Committee of the CPP of Ile-de-France 3 (Paris, France).

Procedures

Clinical predictors assessed at entry in the cohort included demographic data on age, gender, body-mass index (BMI), and the geographical origin. The following clinical characteristics were also gathered during the inclusion visit using an electronic case-report form: family history of HCC, past and current alcohol or tobacco consumption, existence of medical comorbidities at inclusion (diabetes, hypertension, dyslipidemia), liver fibrosis staging and history of past treatment, results at inclusion of FibroTest and FibroScan tests, model for end-stage liver disease (MELD) score in patients with cirrhosis, Child-Pugh scores and results of biological samples (transaminase levels, hemoglobin, creatinine, neutrophils, glucose, total cholesterol, HDL cholesterol, triglycerides, prothrombin time, International Normalised Ratio (INR), albumin, total and conjugated bilirubin, gamma-glutamyl transferase (GGT), alkaline phosphatase (PAL), ferritin and alpha-fetoprotein (AFP)). Biological, virological and liver fibrosis stage analyses were performed in each participating center according to the routine procedures available in the biochemistry, virology and histopathology units. Information on anti-HBV treatments received during follow-up by patients was extracted. The different modalities of treatment with nucleoside inhibitors, ETV or TDF, were collected from the cohort data. Blood and urine samples were obtained and stored in a centralized biobank (Cell & Co Biorepository, Pont du Château, France). Follow-up included systematic visits once a year and spontaneous reports for particular events, which were recorded on specific data forms (eg, death, HCC, decompensated cirrhosis, the onset of treatment).

Liver fibrosis assessment was done closest to the date of inclusion, but less than 1 year before and up to 3 months after inclusion. Fibrosis was assessed either by liver biopsy or another noninvasive method (liver stiffness measurement; FibroScan, Echosens, Paris, France), FibroTest (Biopredictive, Paris, France), Fibrometer (Echosens), FIB-4 index or the Hepascore [25]. If a recent measurement of liver fibrosis was unavailable, or in case of discrepancies between noninvasive fibrosis markers, clinicians were asked to assess the level

of liver fibrosis based on past fibrosis scores and the patient's history of liver-related comorbidities. Mild liver fibrosis (F0–F2), severe fibrosis (F3), and cirrhosis (F4) were defined by the Metavir score [26]. Cutoffs for severe fibrosis by noninvasive methods were 8.5 kPa with FibroScan, 0.59 with FibroTest and 1.75 with the FIB-4 index.

Incident HCC outcome

For all patients who had an HCC after inclusion, data included number and localizations of nodules, the sum of the diameters of the nodules, diagnostic imaging procedures, histological data, and treatment.

Statistical methods

The categorization of continuous covariates, including biological variables, was based on clinically relevant, a priori determined thresholds. The clinical and demographic characteristics of patients were compared between groups with and without an incident HCC during follow-up by univariate Cox models. To develop an explanatory model, SEMs were conducted to further investigate the hypothetical causal relationships between the clinical and demographic variables at the time of inclusion and the occurrence of HCC. Consequently, we modelled time-to-event event outcomes and censored survival time results in our SEMs. The models were selected based on model fit criteria and clinical plausibility. The causal diagrams proposed by the SEM provide a visual summary of the most relevant causal relationships between patient inclusion characteristics and disease. A latent variable is a construct that cannot be directly observed or measured but is inferred from other indicator measures [27]. The following latent variables (and their component measures) were selected to best operationalize the underlying theoretical constructs: liver fibrosis, liver biology, metabolic syndrome. The metabolic risk factors that usually lead to a metabolic syndrome have been identified to define the corresponding latent variable in the SEM. Covariates without

significant paths were trimmed from the model to improve the overall model fit. The Root Mean Square Error of Approximation absolute adequacy index (RMSEA) was determined to assess the quality of the SEM. An RMSEA smaller than 0.08 is generally considered a good fit, while an RMSEA higher than 0.1 corresponds to a poor fit [28]. The following scores were also determined: not significant model chi-square (χ^2), Standardized Root Mean square Residual (SRMR), comparative fit index (CFI), and Non-Normed Fit Index (NNFI) [29]. For easier interpretation of the coefficients, we reported standardized estimates with p-values ≤ 0.05. Standardized estimates (ranging from -1 for a completely negative association to 1 for a completely positive association) can be interpreted relative to other estimates and the relative strength of associations can be compared. Two SEM have been developed. The main model concerned all patients. In a sensitivity analysis to explore the pathways linking patient characteristics to the occurrence of HCC in a subset selected according to the liver fibrosis stage, we estimated a SEM in patients with F3-F4 fibrosis evaluated at entry in the cohort – in this SEM, liver fibrosis was no longer considered as a latent variable.

Multiple Imputations by Chained Equations were used to address missing data. The imputation model included all clinical and adjustments variables. Complete data on more than 75% of characteristics and confounders were available for 2852 participants eligible for the study. The covariates incorporated in the imputed model were considered missing at random. Fifty randomly generated datasets were imputed, and analyses were performed on these data. Data were analyzed using R version 3.6.1. The SEMs was developed with the Lavaan 0.6-4 package. All indicated p-values are obtained in bilateral tests and p-values below 0.05 were considered statistically significant.

Results

Flow diagram of the study

Figure 1 shows the construction of the samples analyzed for the study. 4,489 chronic HBV-infected patients were extracted and analyzed. Among these patients, 3,781 patients had F0-F2 fibrosis and 708 patients had F3-F4 fibrosis. 33 (0.7%) patients with incident HCC were identified, including 11 (33.3%) patients with F0-F2 fibrosis and 22 (66.6%) patients with F3-F4 fibrosis. The HCC incidence rate was 0.21 (95% CI 0.14-0.29) cases per 100 person-years in the total population and 0.87 (95% CI 0.55-1.33) in subjects with F3-F4 fibrosis. We identified 822 (18.3%) and 968 (21.6%) patients treated with ETV and TDF alone, respectively. Of these, 90% remained on this treatment for the duration of the study. Additionally, we observed that 8% of untreated patients at entry in the cohort, initiated treatment with ETV or TDF during the follow-up period of the study.

Liver fibrosis at inclusion was determined by liver biopsy and noninvasive methods with reasonable reliability in 521 (11.6%) and 1666 (37.1%) patients respectively. The level of liver fibrosis had to be confirmed by a physician based on clinical examination, previous fibrosis scores and the patient history of liver related comorbidities in 2310 (51.5%) patients.

The median follow-up time for patients was 45.5 months (IQR 30.6-57.9).

Comparison of patient characteristics

Table 1 describes the clinical-demographic characteristics of patients at inclusion and the distribution of patients by whether or not an incident HCC occurred during follow-up. The treatments received by patients are also detailed. In univariate analyses, compared to patients who did not develop HCC, patients who had an incident HCC during follow-up are on average older, are predominantly male, smoke more, have higher alcohol consumption,

arterial hypertension, advanced fibrosis (≥ F3), elevated AFP and more pejorative results on FibroTest and FibroScan ($p < 0.05$ for all these analyses).

SEMs mapping the relationships between variables

Figure 2 presents the SEM proposed for this study by imagining the numerous links that can link the variables. The absolute adequacy index of RMSEA was 0.051 (95% CI: 0.049 - 0.053) after convergence, which allows considering the model has good adequacy to the data. Table 2 presents estimates of the parameters of the SEM. The latent variable liver fibrosis appears to be both a consequence of a set of patient characteristics at inclusion and a cause of the occurrence of HCC and thus an intermediate factor in the causal relationship. Severe liver fibrosis (B = 0.71) was directly and significantly associated with incident HCC. Besides, age, male gender, excessive alcohol consumption, metabolic syndrome (latent variable) and elevated level of HBV DNA were indirectly associated with the occurrence of HCC (all via the liver fibrosis). The standardized total effects on HCC of age (B = 0.13), male gender (B = 0.16), elevated HBV DNA level ≥ 10^5 copies/ml (B = 0.18) were similar, smaller than that of an advanced liver fibrosis (B = 0.71), and greater than that of excessive alcohol consumption (B = 0.06).

Figure 3 represents the SEM developed specifically for patients with F3-F4 liver fibrosis stage at baseline. This model is specified in Table 3.

Discussion

This study identifies the socio-demographic and clinical characteristics predicting the occurrence of HCC among chronic HBV-infected patients in France. The liver fibrosis mediated the effects of age, male gender, alcohol consumption, metabolic syndrome and HBV DNA level on the occurrence of HCC. The incidence of HCC in chronic HBV-infected patients increases significantly in elderly men with a history of excessive alcohol consumption, fibrosis ≥ F3 and a metabolic syndrome.

Cirrhosis is a recognized risk factor for the occurrence of HCC [30–33] with an increased risk of more than 10 times in cohorts with HBV infection [32,33]. Our study considered this comorbidity through a SEM with satisfactory adequacy parameters [28,34]. This model considered many causal patterns that explain the occurrence of HCC based on previous data in the literature. Importantly, the liver fibrosis not only had a direct relation with HCC but also was a key mediator for other factors as a common pathway. One of the strengths of our study was the proposal of a SEM model to approximate as closely as possible the causal relationship between the characteristics of the subjects and the occurrence of HCC. Dependencies between both observed variables and the latent construct by using a path model were highlighted using SEM.

The increased risk of HCC occurrence with the age of individuals in this study was consistent with data already widely observed in previous studies [35,36].

Serum alpha-fetoprotein is a serological marker widely used in practice in combination with liver ultrasound to detect HCC in chronic HBV-infected patients. We observed a significant association between incident HCC in chronic HBV-infected patients with AFP levels above 10 ng/ml at inclusion (B = 0.18). Patients with HCC have elevated rates of AFP at 12, 9, 6 and 3 months before the onset of HCC [37]. Moreover, an AFP level above 9 ng/ml at

inclusion is an independent risk factor for the occurrence of HCC in advanced stages (stage B or C - BCLC) despite proper patient monitoring [37]. According to our results, particular attention should, therefore, be paid to the patients with an AFP level above 10 ng/ml who are at high risk of developing HCC.

The results of this study did not indicate a significant association between tobacco consumption and the risk of HCC in chronic HBV-infected patients. Previous studies have found contradictory results regarding the link between smoking and liver cancer [18,38], however, strong pharmacological hypotheses based on the carcinological nature of some tobacco metabolites suggest this association. This is the case, for example, for DNA adducts of 4-aminobiphenyl [39,40], or vinyl chloride [41]. This association with tobacco was also detected in a Chinese non-prospective study of a chronic HBV population [42]. However, none of these studies explored this risk factor employing a prospective methodology in a European population with chronic HBV infection. Because of divergent results on these epidemiological studies, prospective studies involving more geographical territories with greater representativeness of ethnic groups are required.

The occurrence of incident HCC was positively associated with male gender. The literature data support an increased risk of HCC of two to four times higher in men compared to women [15,36]. The reasons why liver cancer rates are higher in men than in women are not fully known, androgenic hormones and increased genetic susceptibility may partly explain this risk in men.

Several publications support an increased risk of HCC due to diabetes [14,43,44]. This increased risk was also not significantly detected in a study involving an Australian cohort of approximately 128,000 patients [36]. Although diabetes appears to be slightly associated with an increased risk of HCC, this risk was limited in our patients, suggesting that this

co-morbidity represents a low risk compared to other characteristics included our models. Furthermore, our study had probably limited power to demonstrate this association in chronic HBV-infected patients.

The secondary SEM analysis of patients with F3-F4 fibrosis suggested that the metabolic syndrome may be an independent predictor of the development of HCC in chronic HBV patients with advanced fibrosis. The association between non-alcoholic fatty liver disease (NAFLD), the hepatic manifestation of the metabolic syndrome, and HCC is supported by several epidemiological studies [45,46]. Interestingly, this association is mainly limited to those with advanced fibrosis and cirrhosis [46,47], which is consistent with our SEM performed specifically in patients with F3-F4 fibrosis. Although there is some evidence to support, in non-alcoholic fatty liver disease, an association of host genetic variants, in particular the gene coding for the patatin-like phospholipase domain-containing 3 (PNPLA3), with the development of HCC, independent of the activity and the extent of liver damage [48,49], this positive direct effect not mediated by fibrosis has not been demonstrated in the primary SEM including all patients.

In the same way, high levels of HBV DNA independently led to the occurrence of HCC in patients with F3-F4 fibrosis. Indeed, regardless of the presence of cirrhosis, an elevated HBV DNA at baseline is a strong predictor of HCC [50]. Therefore, the direct effect of the pathway involving HBV DNA level on the occurrence of HCC was illustrated in patients with F3-F4 fibrosis.

Limitations

Our analyses were limited principally by the small number of HCC cases identified during the prospective follow-up, which led us to make a careful choice of variables to include in our SEM. In addition, some patients might have undergone less regular screening for HCC

carcinoma than recommended, resulting in potentially missed diagnoses. The assessment of fibrosis was based on patient records upon arrival in the cohort, determined by different methods and not updated during follow-up. However, fibrosis probably progressed in some patients, which explains the development of liver-related complications in patients classified without advanced fibrosis at entry in the cohort. In addition, possible misclassification of fibrosis cannot be ruled out, although we aimed to limit this by asking physicians to clearly provide their interpretation of true levels of fibrosis at baseline using the available measures. Consequently, the staging of fibrosis remains a limitation of this study as fibrosis may be evolutive over time and there was a lack of standardization of this assessment involving either invasive (liver biopsy) and noninvasive means, or physician evaluation based on clinical examination, previous fibrosis scores and the patient's medical history to confirm fibrosis. Finally, since our study is observational, it is possible that some confounding factors have not been taken into account, which could lead to residual confusion bias in our results. Similarly, we cannot formally conclude the associations between the clinical characteristics of the subjects, and the incidence of HCC reflect cause-and-effect relationships. However, we can assume plausible mechanisms. Age, male gender, metabolic syndrome, alcohol consumption and an elevated viral replication of HBV DNA cause inflammation and regeneration of the liver, which can lead to fibrosis and ultimately to the development of HCC.

Conclusions

Liver fibrosis mediates the effects of age, male gender, alcohol consumption, metabolic syndrome and HBV DNA level on the occurrence of HCC in chronic HBV-infected patients. Considering the extremely negative prognosis associated with HCC, a more targeted and close follow-up of elderly men with a history of excessive alcohol consumption, liver fibrosis stage

≥ F3, a metabolic syndrome and elevated level of HBV DNA is necessary to detect HCC as early as possible in chronic HBV-infected patients.

Acknowledgments

We thank participants and participating clinicians at each study site.

References

[1] Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. The Lancet 2015;386:1546–55. https://doi.org/10.1016/S0140-6736(15)61412-X.

[2] Razavi-Shearer D, Gamkrelidze I, Nguyen MH, Chen D-S, Van Damme P, Abbas Z, et al. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. Lancet Gastroenterol Hepatol 2018;3:383–403. https://doi.org/10.1016/S2468-1253(18)30056-6.

[3] 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. https://doi.org/10.3322/caac.21492.

[4] Pisani P, Parkin DM, Muñoz N, Ferlay J. Cancer and infection: estimates of the attributable fraction in 1990. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol 1997;6:387–400.

[5] Yu MW, Chen CJ. Hepatitis B and C viruses in the development of hepatocellular carcinoma. Crit Rev Oncol Hematol 1994;17:71–91.

[6] But D-Y-K, Lai C-L, Yuen M-F. Natural history of hepatitis-related hepatocellular carcinoma. World J Gastroenterol 2008;14:1652–6. https://doi.org/10.3748/wjg.14.1652.

[7] Fattovich G, Bortolotti F, Donato F. Natural history of chronic hepatitis B: Special emphasis on disease progression and prognostic factors. J Hepatol 2008;48:335–52. https://doi.org/10.1016/j.jhep.2007.11.011.

[8] El-Serag HB. Hepatocellular Carcinoma. N Engl J Med 2011;365:1118–27. https://doi.org/10.1056/NEJMra1001683.

[9] Levrero M, Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. J Hepatol 2016;64:S84–101. https://doi.org/10.1016/j.jhep.2016.02.021.

[10] Nagaoki Y, Hyogo H, Aikata H, Tanaka M, Naeshiro N, Nakahara T, et al. Recent trend of clinical features in patients with hepatocellular carcinoma. Hepatol Res Off J Jpn Soc Hepatol 2012;42:368–75. https://doi.org/10.1111/j.1872-034X.2011.00929.x.

[11] Clifford GM, Rickenbach M, Polesel J, Dal Maso L, Steffen I, Ledergerber B, et al. Influence of HIV-related immunodeficiency on the risk of hepatocellular carcinoma. AIDS Lond Engl 2008;22:2135–41. https://doi.org/10.1097/QAD.0b013e32831103ad.

[12] Shi J, Zhu L, Liu S, Xie W-F. A meta-analysis of case-control studies on the combined effect of hepatitis B and C virus infections in causing hepatocellular carcinoma in China. Br J Cancer 2005;92:607–12. https://doi.org/10.1038/sj.bjc.6602333.

[13] Zhang JY, Wang X, Han SG, Zhuang H. A case-control study of risk factors for hepatocellular carcinoma in Henan, China. Am J Trop Med Hyg 1998;59:947–51. https://doi.org/10.4269/ajtmh.1998.59.947.

[14] Davila JA, Morgan RO, Shaib Y, McGlynn KA, El-Serag HB. Diabetes increases the risk of hepatocellular carcinoma in the United States: a population based case control study. Gut 2005;54:533–9. https://doi.org/10.1136/gut.2004.052167.

[15] McGlynn KA, London WT. Epidemiology and natural history of hepatocellular carcinoma. Best Pract Res Clin Gastroenterol 2005;19:3–23. https://doi.org/10.1016/j.bpg.2004.10.004.

[16] Nguyen MH, Whittemore AS, Garcia RT, Tawfeek SA, Ning J, Lam S, et al. Role of ethnicity in risk for hepatocellular carcinoma in patients with chronic hepatitis C and

cirrhosis. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2004;2:820–4.
[17] Chen C-J, Yang H-I, Su J, Jen C-L, You S-L, Lu S-N, et al. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. JAMA 2006;295:65–73. https://doi.org/10.1001/jama.295.1.65.
[18] Abdel-Rahman O, Helbling D, Schöb O, Eltobgy M, Mohamed H, Schmidt J, et al. Cigarette smoking as a risk factor for the development of and mortality from hepatocellular carcinoma: An updated systematic review of 81 epidemiological studies. J Evid-Based Med 2017;10:245–54. https://doi.org/10.1111/jebm.12270.
[19] Liaw Y-F, Sung JJY, Chow WC, Farrell G, Lee C-Z, Yuen H, et al. Lamivudine for Patients with Chronic Hepatitis B and Advanced Liver Disease. N Engl J Med 2004;351:1521–31. https://doi.org/10.1056/NEJMoa033364.
[20] Choi J, Kim HJ, Lee J, Cho S, Ko MJ, Lim Y-S. Risk of Hepatocellular Carcinoma in Patients Treated With Entecavir vs Tenofovir for Chronic Hepatitis B: A Korean Nationwide Cohort Study. JAMA Oncol 2019;5:30–6. https://doi.org/10.1001/jamaoncol.2018.4070.
[21] Flemming JA, Terrault NA. Tenofovir vs Entecavir for Hepatocellular Carcinoma Prevention in Patients With Chronic Hepatitis B: One of These Things Is Not Like the Other. JAMA Oncol 2019;5:17. https://doi.org/10.1001/jamaoncol.2018.4039.
[22] Yuen M-F, Cheng C, J Lauder I, Lam S, Ooi G, Lai C. Early detection of hepatocellular carcinoma increases the chance of treatment: Hong Kong experience. Hepatol Baltim Md 2000;31:330–5. https://doi.org/10.1002/hep.510310211.
[23] Singal AG, El-Serag HB. Hepatocellular Carcinoma From Epidemiology to Prevention: Translating Knowledge into Practice. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2015;13:2140–51. https://doi.org/10.1016/j.cgh.2015.08.014.
[24] Pol S, Bourliere M, Lucier S, Hezode C, Dorival C, Larrey D, et al. Safety and efficacy of daclatasvir-sofosbuvir in HCV genotype 1-mono-infected patients. J Hepatol 2017;66:39–47. https://doi.org/10.1016/j.jhep.2016.08.021.
[25] Adams LA, Bulsara M, Rossi E, DeBoer B, Speers D, George J, et al. Hepascore: an accurate validated predictor of liver fibrosis in chronic hepatitis C infection. Clin Chem 2005;51:1867–73. https://doi.org/10.1373/clinchem.2005.048389.
[26] Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. Hepatol Baltim Md 1996;24:289–93. https://doi.org/10.1002/hep.510240201.
[27] Ullman JB. Structural equation modeling: reviewing the basics and moving forward. J Pers Assess 2006;87:35–50. https://doi.org/10.1207/s15327752jpa8701_03.
[28] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55. https://doi.org/10.1080/10705519909540118.
[29] Hooper D, Coughlan J, R. Mullen M. Structural Equation Modelling: Guidelines for Determining Model Fit. The Electronic Journal of Business Research Methods 2008;6:53–60.
[30] Chu CM. Natural history of chronic hepatitis B virus infection in adults with emphasis on the occurrence of cirrhosis and hepatocellular carcinoma. J Gastroenterol Hepatol 2000;15 Suppl:E25-30.
[31] Sherman M. Hepatocellular carcinoma: New and emerging risks. Dig Liver Dis Off J Ital Soc Gastroenterol Ital Assoc Study Liver 2010;42 Suppl 3:S215-222. https://doi.org/10.1016/S1590-8658(10)60508-7.

[32] Beasley RP, Hwang LY, Lin CC, Chien CS. Hepatocellular carcinoma and hepatitis B virus. A prospective study of 22 707 men in Taiwan. Lancet Lond Engl 1981;2:1129–33. https://doi.org/10.1016/s0140-6736(81)90585-7.
[33] Yu MW, Hsu FC, Sheen IS, Chu CM, Lin DY, Chen CJ, et al. Prospective study of hepatocellular carcinoma and liver cirrhosis in asymptomatic chronic hepatitis B virus carriers. Am J Epidemiol 1997;145:1039–47. https://doi.org/10.1093/oxfordjournals.aje.a009060.
[34] MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1:130–49. https://doi.org/10.1037/1082-989X.1.2.130.
[35] Ikeda K, Saitoh S, Koida I, Arase Y, Tsubota A, Chayama K, et al. A multivariate analysis of risk factors for hepatocellular carcinogenesis: a prospective observation of 795 patients with viral and alcoholic cirrhosis. Hepatol Baltim Md 1993;18:47–53.
[36] Walter SR, Thein H-H, Gidding HF, Amin J, Law MG, George J, et al. Risk factors for hepatocellular carcinoma in a cohort infected with hepatitis B or C. J Gastroenterol Hepatol 2011;26:1757–64. https://doi.org/10.1111/j.1440-1746.2011.06785.x.
[37] Chon YE, Jung KS, Kim M-J, Choi J-Y, An C, Park JY, et al. Predictors of failure to detect early hepatocellular carcinoma in patients with chronic hepatitis B who received regular surveillance. Aliment Pharmacol Ther 2018;47:1201–12. https://doi.org/10.1111/apt.14578.
[38] Chuang S-C, Lee Y-CA, Hashibe M, Dai M, Zheng T, Boffetta P. Interaction between cigarette smoking and hepatitis B and C virus infection on the risk of liver cancer: a meta-analysis. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol 2010;19:1261–8. https://doi.org/10.1158/1055-9965.EPI-09-1297.
[39] Wang LY, Chen CJ, Zhang YJ, Tsai WY, Lee PH, Feitelson MA, et al. 4-Aminobiphenyl DNA damage in liver tissue of hepatocellular carcinoma patients and controls. Am J Epidemiol 1998;147:315–23. https://doi.org/10.1093/oxfordjournals.aje.a009452.
[40] Chen S-Y, Wang L-Y, Lunn RM, Tsai W-Y, Lee P-H, Lee C-S, et al. Polycyclic aromatic hydrocarbon-DNA adducts in liver tissues of hepatocellular carcinoma patients and controls. Int J Cancer 2002;99:14–21. https://doi.org/10.1002/ijc.10291.
[41] Grosse Y, Baan R, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. Carcinogenicity of 1,3-butadiene, ethylene oxide, vinyl chloride, vinyl fluoride, and vinyl bromide. Lancet Oncol 2007;8:679–80.
[42] Li X, Xu H, Gao P. Diabetes Mellitus is a Risk Factor for Hepatocellular Carcinoma in Patients with Chronic Hepatitis B Virus Infection in China. Med Sci Monit Int Med J Exp Clin Res 2018;24:6729–34. https://doi.org/10.12659/MSM.911702.
[43] El–Serag HB, Hampel H, Javadi F. The Association Between Diabetes and Hepatocellular Carcinoma: A Systematic Review of Epidemiologic Evidence. Clin Gastroenterol Hepatol 2006;4:369–80. https://doi.org/10.1016/j.cgh.2005.12.007.
[44] El-Serag HB. Epidemiology of hepatocellular carcinoma in USA. Hepatol Res Off J Jpn Soc Hepatol 2007;37 Suppl 2:S88-94. https://doi.org/10.1111/j.1872-034X.2007.00168.x.
[45] Kasmari AJ, Welch A, Liu G, Leslie D, McGarrity T, Riley T. Independent of Cirrhosis, Hepatocellular Carcinoma Risk Is Increased with Diabetes and Metabolic Syndrome. Am J Med 2017;130:746.e1-746.e7. https://doi.org/10.1016/j.amjmed.2016.12.029.
[46] Kim H, El-Serag HB. The Epidemiology of Hepatocellular Carcinoma in the USA. Curr Gastroenterol Rep 2019;21:17. https://doi.org/10.1007/s11894-019-0681-x.

[47] Yasui K, Hashimoto E, Komorizono Y, Koike K, Arii S, Imai Y, et al. Characteristics of Patients With Nonalcoholic Steatohepatitis Who Develop Hepatocellular Carcinoma. Clin Gastroenterol Hepatol 2011;9:428–33. https://doi.org/10.1016/j.cgh.2011.01.023.
[48] Negro F. Natural history of NASH and HCC. Liver Int 2020;40:72–6. https://doi.org/10.1111/liv.14362.
[49] Liu Y-L, Patman GL, Leathart JBS, Piguet A-C, Burt AD, Dufour J-F, et al. Carriage of the PNPLA3 rs738409 C >G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma. J Hepatol 2014;61:75–81. https://doi.org/10.1016/j.jhep.2014.02.030.
[50] Chen C-J. Risk of Hepatocellular Carcinoma Across a Biological Gradient of Serum Hepatitis B Virus DNA Level. JAMA 2006;295:65. https://doi.org/10.1001/jama.295.1.65.

Figure legends:

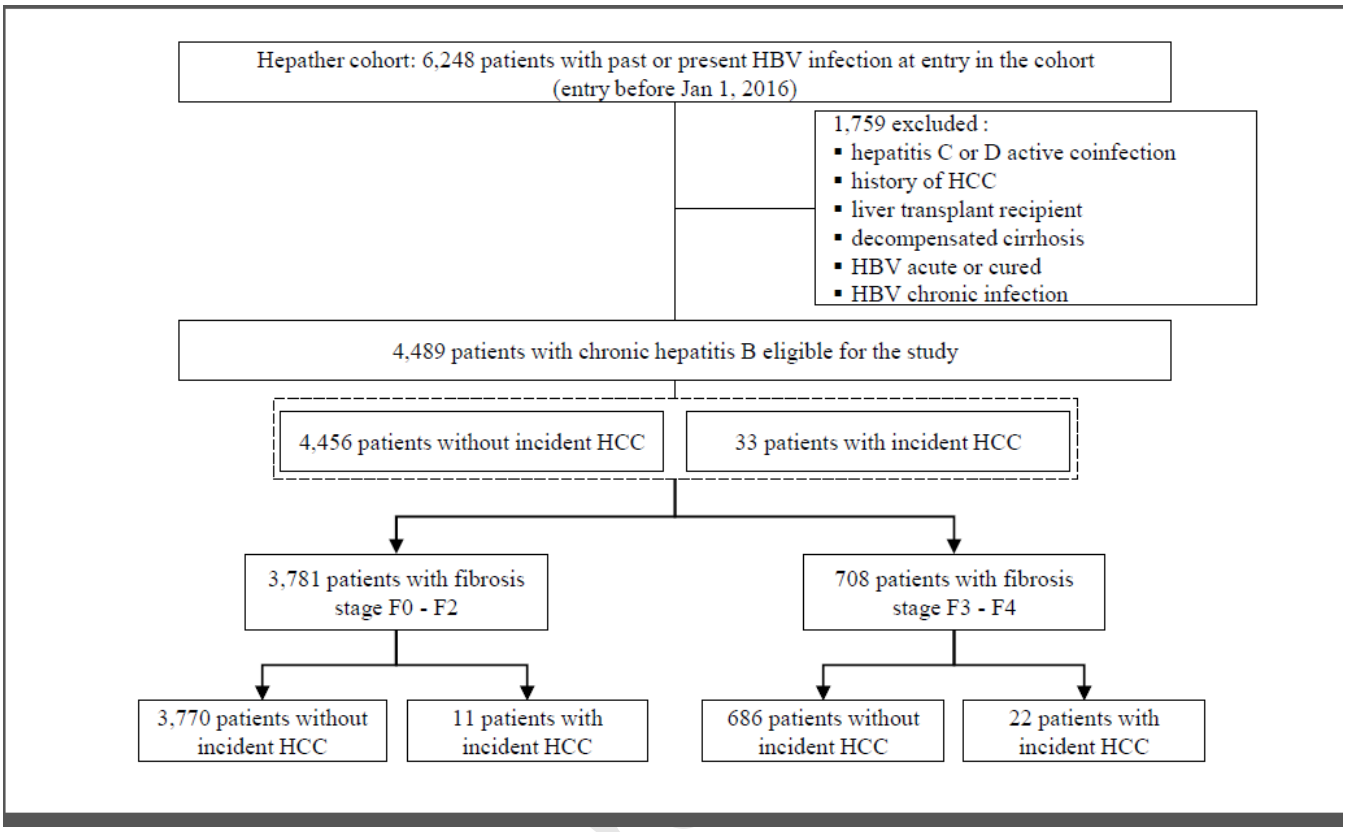

Figure 1. Flowchart showing the disposition of the patients included in the study (n = 4,489)

ETV = entecavir. HBV = hepatitis B virus. HCV = hepatitis C virus. HCC = hepatocellular carcinoma. TDF = Tenofovir Disoproxil Fumarate.

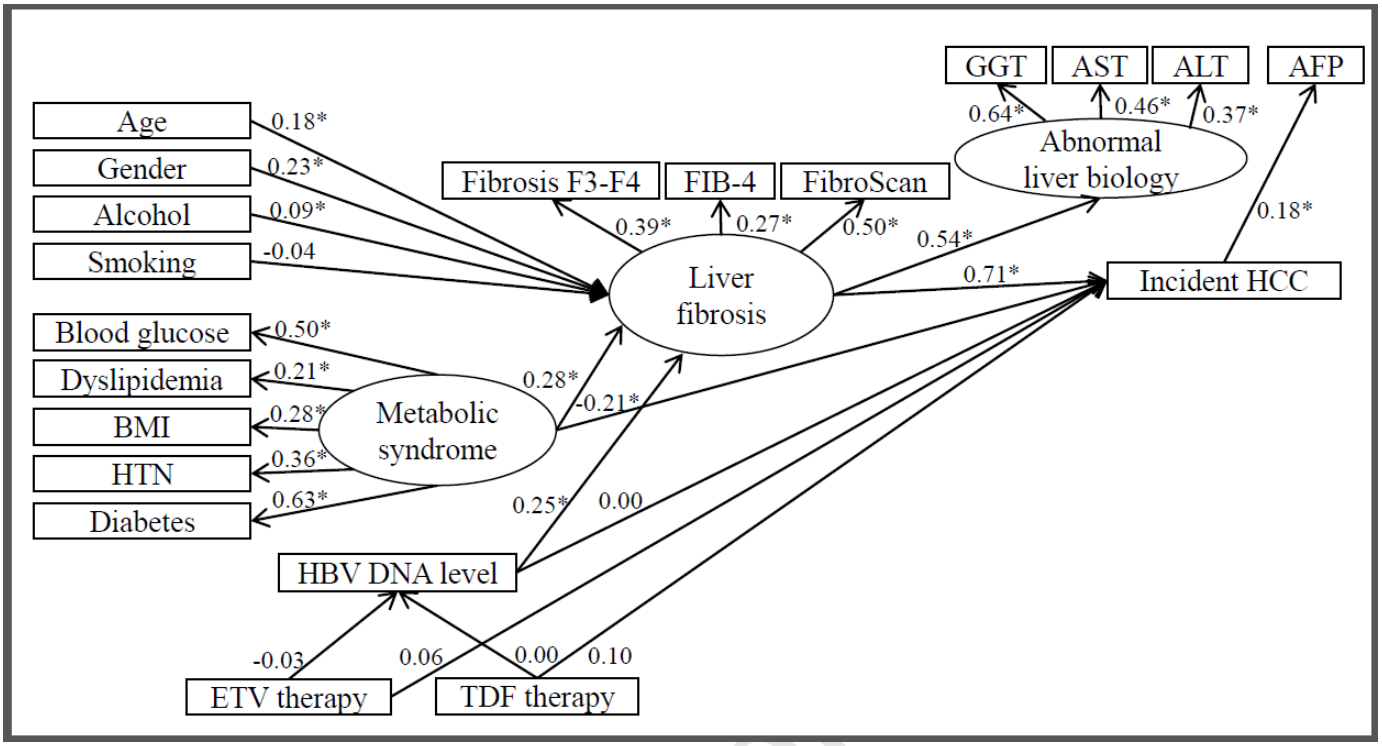

Figure 2. Hypothetical sequential process with the effects of different clinical characteristics on the occurrence of HCC. All patients analyzed (n = 4,489)

Numbers corresponding to the arrows represent the standardized regression coefficients of direct paths. Detailed standardized regression coefficients (beta weights) obtained through respective regression equations in the SEM are presented in Table 2. Arrows indicate structural component and direct effect. Analyses were performed using Weighted Least Squares (WLS) as the estimator. Ovals = factors (ie, latent variables). Rectangles = measured variables. Model fit: root- mean- square error of approximation (RMSEA) = 0.051, 95% CI for RMSEA = [0.049, 0.053]. ALT = alanine aminotransferase. AFP = alpha-fetoprotein. AST = aspartate aminotransferase. BMI = body mass index, GGT = gamma-glutamyl transferase. HBV = hepatitis B virus. HTN = hypertension. * statistically significant pathways at $p < 0.05$.

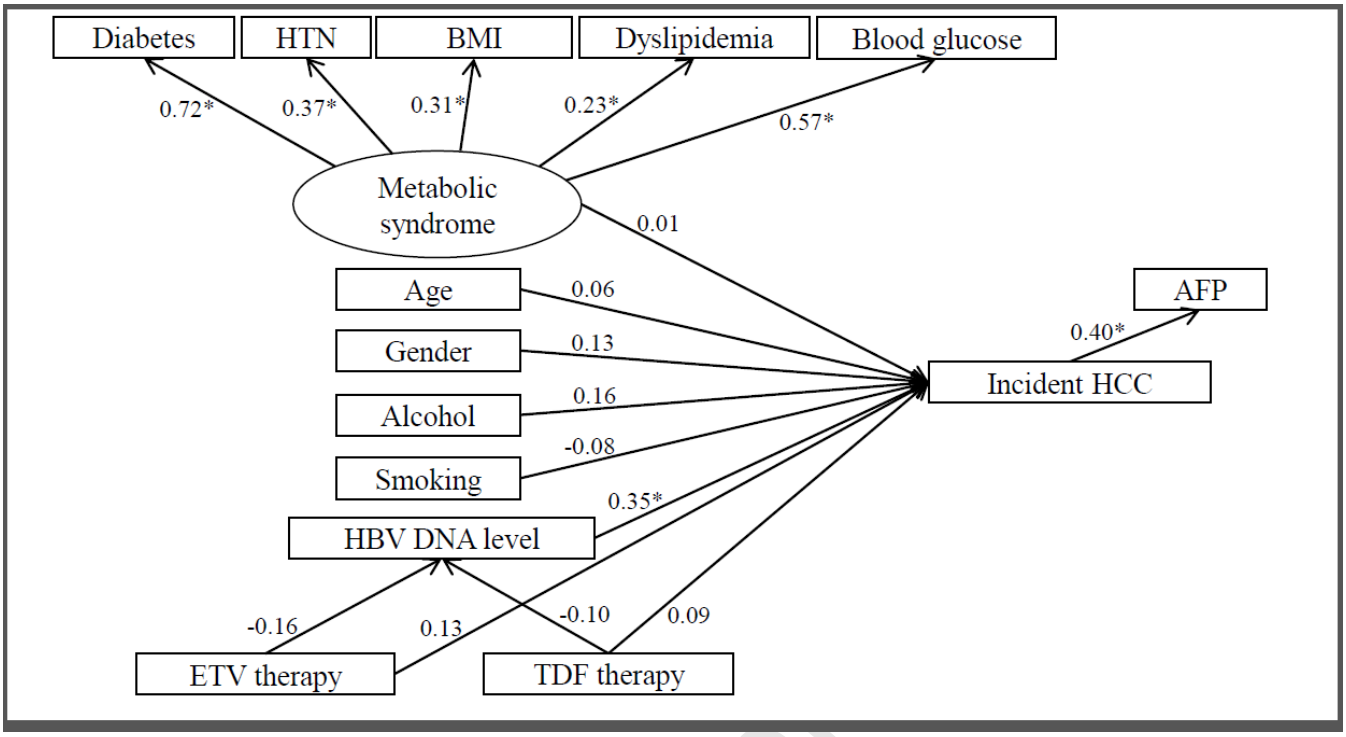

Figure 3. Hypothetical sequential process with the effects of different clinical characteristics on the occurrence of HCC in patients with advanced fibrosis (n = 708)

Numbers corresponding to the arrows represent the standardized regression coefficients of direct paths. Detailed standardized regression coefficients (beta weights) obtained through respective regression equations in the SEM are presented in Table 3. Arrows indicate structural component and direct effect. Analyses were performed using Weighted Least Squares (WLS) as the estimator. Ovals = factors (ie, latent variables). Rectangles = measured variables. Model fit: root- mean- square error of approximation (RMSEA) = 0.052, 95% CI for RMSEA = [0.044, 0.060]. AFP = alpha-fetoprotein. BMI = body mass index. ETV = entecavir. HBV = hepatitis B virus. HCC = hepatocellular carcinoma. HTN = hypertension. TDF = Tenofovir Disoproxil Fumarate. * statistically significant pathways at $p < 0.05$.

Tables with captions

Table 1. Characteristics of patients at entry in the cohort

	All patients	HCC (-)	HCC (+)	Uni. HR (95% CI)	*P*-value
n	4489	4456	33		
Age (years), mean (SD)	45.52 (13.89)	45.42 (13.86)	59.89 (10.00)	1.07 (1.05-1.10)	< .001
Gender, n (%)					
Male	2997 (66.8)	1487 (33.4)	5 (15.2)	2.75 (1.06-7.11)	0.037
Female	1492 (33.2)	2969 (66.6)	28 (84.8)	*ref.*	
Geographical origin, n (%)					
France and Europe	1516 (33.8)	1496 (33.6)	20 (60.6)	*ref.*	
North Africa	507 (11.3)	503 (11.3)	4 (12.1)	0.61 (0.21-1.77)	0.359
Sub-Saharan Africa	1564 (34.8)	1559 (35.0)	5 (15.2)	0.27 (0.10-0.72)	0.009
Asia	902 (20.1)	898 (20.2)	4 (12.1)	0.35 (0.12-1.03)	0.056
Family history of HCC, n (%)					
No	4350 (96.9)	4320 (96.9)	30 (90.9)	*ref.*	
Yes	139 (3.1)	136 (3.1)	3 (9.1)	3.19 (0.97-10.45)	0.055
Smoking, n (%)					
No	2845 (63.4)	2831 (63.5)	14 (42.4)	*ref.*	
Yes	1644 (36.6)	1625 (36.5)	19 (57.6)	2.28 (1.14-4.53)	0.020
Excessive alcohol consumption†, n (%)					
No	4055 (90.3)	4032 (90.5)	23 (69.7)	*ref.*	
Yes	434 (9.7)	424 (9.5)	10 (30.3)	4.06 (1.93-8.52)	< .001
BMI (kg/m^2), n (%)					
≤ 29.9	3852 (85.8)	3823 (85.8)	29 (87.9)	*ref.*	
≥ 30	637 (14.2)	633 (14.2)	4 (12.1)	0.84 (0.29-2.38)	0.741
Diabetes, n (%)					
No	4141 (92.2)	4111 (92.3)	30 (90.9)	*ref.*	
Yes	348 (7.8)	345 (7.7)	3 (9.1)	1.19 (0.36-3.91)	0.771
Hypertension, n (%)					
No	3706 (82.6)	3687 (82.7)	19 (57.6)	*ref.*	
Yes	783 (17.4)	769 (17.3)	14 (42.4)	3.29 (1.65-6.56)	< .001
Dyslipidemia, n (%)					
No	4074 (90.8)	4043 (90.7)	31 (93.9)	*ref.*	
Yes	415 (9.2)	413 (9.3)	2 (6.1)	0.59 (0.14-2.46)	0.468
HBV classification, n (%)					
Chronic hepatitis HBeAg -	3850 (85.8)	3825 (85.8)	25 (75.8)	*ref*	
Chronic hepatitis HBeAg +	639 (14.2)	631 (14.2)	8 (24.2)	1.74 (0.79-3.86)	0.172
Liver fibrosis, n (%)					
< F3	3781 (84.2)	3770 (84.6)	11 (33.3)	*ref.*	
≥ F3	708 (15.8)			10.58 (5.13-	< .001
		686 (15.4)	22 (66.7)	21.82)	
Cirrhosis, n (%)	312 (7.0)	298 (6.7)	14 (42.4)		
FibroTest score, n (%)					
< 0.59	3630 (80.9)	3623 (81.3)	7 (21.2)	*ref.*	
≥ 0.59	859 (19.1)			14.89 (6.46-	< .001
		833 (18.7)	26 (78.8)	34.30)	
FibroScan score, n (%)					
< 8.5 kPa	3773 (84.0)	3761 (84.4)	12 (36.4)	*ref.*	
≥ 8.5 kPa	716 (16.0)	695 (15.6)	21 (63.6)	8.98 (4.42-18.25)	< .001
FIB-4 index					
≤ 1.75	3240 (72.2)	3237 (72.6)	3 (9.1)	*ref.*	
> 1.75	1249 (27.8)			24.11 (7.36-	< .001
		1219 (27.4)	30 (90.9)	79.00)	
HBV DNA level, n (%)					
< 10^5 copies/ml	4209 (93.8)	4178 (93.8)	31 (93.9)	*ref.*	
≥ 10^5 copies/ml	280 (6.2)	278 (6.2)	2 (6.1)	1.03 (0.25-4.32)	0.964
AST, n (%)					

	All patients	HCC (-)	HCC (+)	Uni. HR (95% CI)	*P*-value
< 1 ULN	3530 (78.6)	3510 (78.8)	20 (60.6)	*ref.*	
≥ 1 ULN	959 (21.4)	946 (21.2)	13 (39.4)	2.48 (1.23-4.99)	0.011
ALT, n (%)					
< 1 ULN	2451 (54.6)	2435 (54.6)	16 (48.5)	*ref.*	
≥ 1 ULN	2038 (45.4)	2021 (45.4)	17 (51.5)	1.29 (0.65-2.55)	0.469
ALP, n (%)					
< 1 ULN	408 (9.1)	408 (9.2)	0 (0.0)		
≥ 1 ULN	4081 (90.9)	4048 (90.8)	33 (100.0)		
GGT, n (%)					
< 1 ULN	3842 (85.6)	3826 (85.9)	16 (48.5)	*ref.*	
≥ 1 ULN	647 (14.4)	630 (14.1)	17 (51.5)	6.15 (3.11-12.17)	< .001
Total bilirubin, n (%)					
< 20 mmol/l	4124 (91.9)	4094 (91.9)	30 (90.9)	*ref.*	
≥ 20 mmol/l	365 (8.1)	362 (8.1)	3 (9.1)	1.19 (0.36-3.91)	0.772
Conjugated bilirubin, n (%)					
< 5 mmol/l	3719 (82.8)	3697 (83.0)	22 (66.7)	*ref.*	
≥ 5 mmol/l	770 (17.2)	759 (17.0)	11 (33.3)	2.68 (1.30-5.53)	0.008
Prothrombin time, n (%)					
> 70%	4327 (96.4)	159 (3.6)	3 (9.1)	0.37 (0.11-1.23)	0.104
≤ 70%	162 (3.6)	4297 (96.4)	30 (90.9)	*ref.*	
INR, n (%)					
< 1.15	4021 (89.6)	3997 (89.7)	24 (72.7)	*ref.*	
≥ 1.15	468 (10.4)	459 (10.3)	9 (27.3)	3.32 (1.54-7.13)	0.002
Hemoglobin level, n (%)					
≥ 12 g/dL	4124 (91.9)	4094 (91.9)	30 (90.9)	0.78 (0.24-2.57)	0.686
< 12 g/dL	365 (8.1)	362 (8.1)	3 (9.1)	*ref.*	
Neutrophil level, n (%)					
≥ 1500 SI/mm3	4028 (89.7)	4001 (89.8)	27 (81.8)	0.48 (0.20-1.16)	0.104
< 1500 SI/mm3	461 (10.3)	455 (10.2)	6 (18.2)	*ref.*	
Platelet count, n (%)					
< 10^5 per µL	89 (2.0)	86 (1.9)	3 (9.1)	*ref.*	
≥ 10^5 per µL	4400 (98.0)	4370 (98.1)	30 (90.9)	0.21 (0.06-0.69)	0.010
Creatinine, n (%)					
< 13 mg/l	4250 (94.7)	4219 (94.7)	31 (93.9)	ref.	
≥ 13 mg/l	239 (5.3)	237 (5.3)	2 (6.1)	1.18 (0.28-4.92)	0.823
Albumin, n (%)					
< 35 IU	179 (4.0)	177 (4.0)	2 (6.1)	ref.	
≥ 35 IU	4310 (96.0)	4279 (96.0)	31 (93.9)	0.62 (0.15-2.61)	0.518
Gamma globulins, n (%)					
< 7 g/L	116 (2.6)	115 (2.6)	1 (3.0)	ref.	
≥ 7 g/L	4373 (97.4)	4341 (97.4)	32 (97.0)	0.87 (0.12-6.40)	0.894
Ferritin, n (%)					
< 300 ng/ml	3848 (85.7)	3815 (85.6)	33 (100.0)		
≥ 300 ng/ml	641 (14.3)	641 (14.4)	0 (0.0)		
AFP, n (%)					
< 10 ng/mL	4232 (94.3)	4204 (94.3)	28 (84.8)	ref.	
≥ 10 ng/mL	257 (5.7)	252 (5.7)	5 (15.2)	3.02 (1.17-7.83)	0.023
Blood glucose, n (%)					
< 1.1 g/L	3901 (86.9)	3877 (87.0)	24 (72.7)	ref.	
≥ 1.1 g/L	588 (13.1)	579 (13.0)	9 (27.3)	2.48 (1.15-5.34)	0.020
Total cholesterol, n (%)					
< 2.5 g/L	4158 (92.6)	4127 (92.6)	31 (93.9)	ref.	
≥ 2.5 g/L	331 (7.4)	329 (7.4)	2 (6.1)	0.84 (0.20-3.52)	0.815
HDL cholesterol, n (%)					
< 0.5 g/L	2209 (49.2)	2186 (49.1)	23 (69.7)	ref.	
≥ 0.5 g/L	2280 (50.8)	2270 (50.9)	10 (30.3)	0.43 (0.21-0.91)	0.026
Triglycerides, n (%)					
< 1.5 g/L	3842 (85.6)	3815 (85.6)	27 (81.8)	ref.	
≥ 1.5 g/L	647 (14.4)	641 (14.4)	6 (18.2)	1.22 (0.51-2.96)	0.655

	All patients	HCC (-)	HCC (+)	Uni. HR (95% CI)	*P*-value
Treatments, n (%)					
ETV monotherapy	822 (18.3)	813 (18.2)	9 (27.3)	ref.	
TDF monotherapy	968 (21.6)	955 (21.4)	13 (39.4)	1.24 (0.53-2.90)	0.623
ETV + TDF combination	113 (2.5)	113 (2.5)	0 (0.0)		
Combination with ETV	19 (0.4)	19 (0.4)	0 (0.0)		
Combination with TDF	159 (3.5)	157 (3.5)	2 (6.1)	1.02 (0.22-4.71)	
Not treated	2264 (50.4)	2257 (50.7)	7 (21.2)	0.32 (0.12-0.85)	0.022
Others	144 (3.2)	142 (3.2)	2 (6.1)	1.27 (0.28-5.89)	0.757

Data are mean (SD), n (%) unless otherwise stated. AFP = alpha-fetoprotein. ALP = alkaline phosphatase. ALT = alanine aminotransferase. AST = aspartate aminotransferase.

BMI = body mass index. DNA = Deoxyribonucleic acid. ETV = entecavir.

GGT = gamma-glutamyl transferase. HBeAg = Hepatitis B e antigen. HDL = high-density lipoproteins. LDL = low-density lipoproteins. HBV = hepatitis B virus. INR = International Normalized Ratio. MELD = model for end-stage liver disease. TDF = Tenofovir Disoproxil Fumarate. ULN = upper limit of normal. Uni. HR = univariate Cox regression with Hazard Ratio for the relationship between clinical characteristics and survival time of HCC patients.

† Defined as at least 15 alcoholic drinks (150 g) per week for a woman or 22 alcoholic drinks (220 g) per week for a man, or at least six consecutive alcoholic drinks (60 g) on at least one occasion per week.

Table 2. Estimates of the SEM. All patients analyzed (n = 4,489)

Model pathways	Std. Estimate	S.E.	95% CI	*P*-value
Advanced fibrosis†, yes → incident HCC	0.71	0.08	[0.55, 0.87]	< .001
HBV DNA level‡, high → incident HCC	0.00	0.12	[-0.24, 0.24]	0.984
Metabolic syndrome, yes → incident HCC	-0.21	0.07	[-0.35, -0.06]	0.004
ETV therapy, yes → incident HCC	0.06	0.07	[-0.07, 0.19]	0.367
TDF therapy, yes → incident HCC	0.10	0.07	[-0.04, 0.24]	0.148
Age → advanced fibrosis	0.18	0.02	[0.14, 0.23]	< .001
Male gender → advanced fibrosis	0.23	0.03	[0.18, 0.29]	< .001
Smoking status, yes → advanced fibrosis	-0.04	0.02	[-0.08, 0.01]	0.108
Alcohol consumption, yes → advanced fibrosis	0.09	0.02	[0.05, 0.14]	< .001
Metabolic syndrome, yes → advanced fibrosis	0.28	0.03	[0.22, 0.33]	< .001
HBV DNA level, elevated → advanced fibrosis	0.25	0.04	[0.17, 0.34]	< .001
Advanced fibrosis, yes → abnormal liver biology	0.67	0.03	[0.61, 0.74]	< .001
Incident HCC, yes → AFP§ level, high	0.18	0.06	[0.06, 0.29]	0.002
ETV therapy, yes → HBV DNA level, high	-0.03	0.04	[-0.10, 0.04]	0.393
TDF therapy, yes → HBV DNA level, high	0.00	0.04	[-0.06, 0.08]	0.861
Latent variables				
Metabolic syndrome ← diabetes	0.63	0.03	[0.58, 0.68]	< .001
Metabolic syndrome ← dyslipidemia	0.21	0.02	[0.17, 0.24]	< .001
Metabolic syndrome ← BMI, high	0.28	0.02	[0.23, 0.33]	< .001
Metabolic syndrome ← blood sugar, high	0.50	0.02	[0.45, 0.55]	< .001
Metabolic syndrome ← arterial hypertension	0.36	0.02	[0.32, 0.40]	< .001
Advanced fibrosis ← liver fibrosis ≥ F3¶	0.39	0.03	[0.33, 0.44]	< .001
Advanced fibrosis ← FibroScan score ≥ 8.5 kPa	0.50	0.03	[0.44, 0.56]	< .001
Advanced fibrosis ← FIB-4 index ≥ 1.75	0.27	0.02	[0.22, 0.31]	< .001
Abnormal liver biology ← elevated ALT# level	0.37	0.04	[0.29, 0.45]	< .001
Abnormal liver biology ← elevated AST# level	0.46	0.03	[0.40, 0.53]	< .001
Abnormal liver biology ← elevated GGT# level	0.64	0.03	[0.58, 0.71]	< .001

The SEM predicts HCC, with severity of fibrosis as the mediating variable**.** Standardized regression coefficients can be interpreted as the change in y (dependent variable) in y standard deviation units for a standard deviation change in x (independent variable).

AFP = alpha-fetoprotein. ALT= alanine aminotransferase. AST = aspartate aminotransferase. BMI = body mass index. DNA = Deoxyribonucleic acid. ETV = entecavir. HBV= hepatitis B virus. HCC = hepatocellular carcinoma. TDF= tenofovir disoproxil fumarate. † Fibrosis ≥ F3. ‡ HBV DNA level ≥ 10^5 copies/ml, § AFP ≥ 10 ng/ml. ¶ The fibrosis stage refers to fibrosis stage at cohort entry. # ALT/AST/ GGT ≥ 1 ULN.

Table 3. SEM: prediction of the occurrence of HCC in patients with F3-F4 liver fibrosis (n = 708)

Model pathways	Std. Estimate	S.E.	95% CI	*P*-value
Age → incident HCC	0.06	0.11	[-0.16, 0.27]	0.612
Male gender → incident HCC	0.13	0.10	[-0.08, 0.33]	0.218
Metabolic syndrome, yes → incident HCC	0.01	0.08	[-0.14, 0.17]	0.873
HBV DNA level†, high → incident HCC	0.35	0.11	[0.13, 0.57]	0.002
Smoking status, yes → incident HCC	-0.08	0.09	[-0.26, 0.10]	0.379
Alcohol consumption, yes -> incident HCC	0.16	0.08	[0.00, 0.32]	0.055
ETV therapy, yes → incident HCC	0.13	0.10	[-0.06, 0.32]	0.191
TDF therapy, yes → incident HCC	0.09	0.11	[-0.12, 0.30]	0.389
Incident HCC → AFP level‡, high	0.40	0.08	[0.25, 0.55]	< .001
ETV therapy, yes → HBV DNA level, high	-0.16	0.10	[-0.35, 0.03]	0.107
TDF therapy, yes → HBV DNA level, high	-0.10	0.08	[-0.26, 0.06]	0.221
Latent variables				
Metabolic syndrome ← diabetes	0.72	0.06	[0.60, 0.83]	< .001
Metabolic syndrome ← dyslipidemia	0.23	0.04	[0.15, 0.31]	< .001
Metabolic syndrome ← BMI, high	0.31	0.06	[0.19, 0.42]	< .001
Metabolic syndrome ← blood sugar, high	0.57	0.06	[0.46, 0.69]	< .001
Metabolic syndrome ← arterial hypertension	0.37	0.05	[0.27, 0.47]	< .001

Standardized regression coefficients can be interpreted as the change in y (dependent variable) in y standard deviation units for a standard deviation change in x (independent variable). AFP= alpha-fetoprotein. ALT= alanine aminotransferase. AST = aspartate aminotransferase. BMI = body mass index. DNA = Deoxyribonucleic acid. ETV = entecavir. HBV = hepatitis B virus. HCC = hepatocellular carcinoma. TDF = Tenofovir disoproxil fumarate. † HBV DNA level $\geq 10^5$ copies/ml. ‡ AFP $\geq$ 10 ng/ml.