Alimentary Pharmacology & Therapeutics

WILEY

AP&T Alimentary Pharmacology & Therapeutics

ORIGINAL ARTICLE OPEN ACCESS

Development and Validation of a PIVKA-II-Based Model for HCC Risk Stratification in Patients With HCV-Related Cirrhosis Successfully Treated With DAA

Gian Paolo Caviglia[1] | Piero Fariselli[2] | Roberta D'Ambrosio[3] | Piero Colombatto[4] | Elisabetta Degasperi[3] | Gabriele Ricco[4] | Maria Lorena Abate[1] | Giovanni Birolo[2] | Giulia Troshina[1] | Francesco Damone[4] | Barbara Coco[4] | Daniela Cavallone[4] | Riccardo Perbellini[3] | Sara Monico[3] | Giorgio Maria Saracco[1,5] | Maurizia Rossana Brunetto[4,6] | Pietro Lampertico[3,7] | Alessia Ciancio[1,5]

[1]Department of Medical Sciences, Liver Unit, University of Turin, Turin, Italy | [2]Department of Medical Sciences, Computational Biomedicine, University of Turin, Turin, Italy | [3]Division of Gastroenterology and Hepatology, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy | [4]Hepatology Unit and Laboratory of Molecular Genetics and Pathology of Hepatitis Viruses, Reference Center of the Tuscany Region for Chronic Liver Disease and Cancer, University Hospital of Pisa, Pisa, Italy | [5]Gastroenterology Unit, Città della Salute e della Scienza di Torino—Molinette Hospital, Turin, Italy | [6]Institute of Biostructure and Bioimaging, National Research Council, Naples, Italy | [7]CRC "A. M. and A. Migliavacca" Center for Liver Disease, Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

Correspondence: Gian Paolo Caviglia (gianpaolo.caviglia@unito.it)

Received: 17 June 2024 | **Revised:** 22 July 2024 | **Accepted:** 11 November 2024

Handling Editor: Grace L-H Wong

Funding: The authors received no specific funding for this work.

Keywords: biomarkers | DAA | HCC | hepatitis C | non-invasive tests

ABSTRACT

Background and Aims: Patients with hepatitis C virus (HCV)-related cirrhosis with sustained virological response (SVR) to direct-acting antivirals (DAA) remain at risk of developing hepatocellular carcinoma (HCC). Recently, serum protein induced by vitamin K absence or antagonist-II (PIVKA-II) has shown promising results as an HCC-predictive biomarker. We aimed to develop and validate a PIVKA-II-based model for HCC risk stratification in cirrhotic patients with SVR to DAA.

Methods: A total of 1220 consecutive patients (Turin, $n = 531$; Pisa, $n = 335$; Milan, $n = 354$) with HCV-related cirrhosis treated with DAA were included in the study. Patients were retrospectively allocated to the training cohort (Turin+Pisa; median follow-up [FU] 39, 22–55 months; incident HCC: 93 [10.7%]) and validation cohort (Milan; median FU 49.0, 35.0–52.0 months; incident HCC: 19 [5.4%]). Serum PIVKA-II levels were measured using the LumipulseG system (Fujirebio, Japan) at SVR12 (Turin and Pisa cohorts) or the end of treatment (Milan cohort).

Results: Using Cox regression analysis, a model including PIVKA-II combined with age, sex, ALT, AST, γGT, platelet count, albumin and total bilirubin was derived from the training cohort (C-index = 0.72). In the validation cohort, the model showed a C-index of 0.71 with an area under the curve of 0.84 for identifying patients who developed HCC during the first 12 months of FU. When patients were grouped into three risk categories, the cumulative incidence of HCC was 2.7%, 4.0% and 14.3% in the low-, medium- and high-risk groups, respectively ($p < 0.001$). Notably, no HCC occurred within 3 years of FU in the low-risk group.

Conclusions: Our PIVKA-II-based model showed satisfactory accuracy for HCC risk stratification and may represent a valuable tool for implementing risk-based surveillance protocols in patients with HCV-related cirrhosis with SVR to DAA.

Maurizia Rossana Brunetto, Pietro Lampertico and Alessia Ciancio contributed equally to this work as co-last authors.

1 | Introduction

The introduction of direct acting antivirals (DAAs) is a major advance in the treatment of patients with chronic hepatitis C virus (HCV) infection [1]. More than 98% of patients treated with DAA therapy achieve a sustained virological response (SVR), improving hepatic and extrahepatic diseases, quality of life and survival [2]. However, patients with advanced liver disease remain at risk of developing hepatocellular carcinoma (HCC) despite being cured of HCV infection. The incidence of *de novo* HCC is estimated as 2.1 per 100 person-years after HCV treatment in patients with cirrhosis [3], with an increased risk of HCC that may persist for up to 10 years after HCV eradication [4]. Therefore, all patients with cirrhosis cured of HCV infection should undergo long-term surveillance; timely detection of early HCC is critical for the possibility of receiving curative treatments and improving patient survival.

According to international recommendations, HCC surveillance should be performed using biannual ultrasound (US) examination with or without α-fetoprotein (AFP) [5, 6]. However, US has several limitations, including suboptimal accuracy for early-stage HCC, inter-operator variability and poor adherence, whereas AFP lacks adequate sensitivity and specificity [7]. More importantly, the current surveillance policy does not consider the patients' individual risk of HCC development but follows a one-size-fits-all approach.

Protein induced by vitamin K absence or antagonist-II (PIVKA-II), also known as des-γ-carboxy prothrombin, is an immature form of prothrombin produced by malignant hepatocytes. PIVKA-II can be easily measured in serum samples by using highly sensitive and standardised assays that are widely available in biomedical laboratories. To date, several case-control studies have investigated the performance of PIVKA-II in HCC detection, reporting diagnostic accuracies higher than those of AFP [8]. Remarkably, two recent studies demonstrated that PIVKA-II can predict HCC development among at-risk patients with encouraging accuracy [9, 10]. However, the use of serological non-invasive tests (NITs) and composite HCC risk scores combining different demographic, clinical and biochemical variables could represent a valuable approach for the stratification of HCC risk in patients under surveillance [11, 12]. Their major advantage is cost-effectiveness, which makes NITs and risk scores particularly suitable for the surveillance of patients at risk of HCC.

Based on this premise, we aimed to develop and validate a model based on PIVKA-II in combination with variables normally collected during standard clinical practice to stratify the risk of HCC development in cirrhotic patients successfully treated with DAAs during long-term surveillance.

2 | Materials and Methods

2.1 | Patients

This multi-centre retrospective study enrolled consecutive patients with HCV-related cirrhosis who underwent DAA treatment in tertiary referral clinics between June 2014 and December 2020. Patients were included in the study if they fulfilled the following criteria: age ≥ 18 years at DAA initiation, SVR to DAA treatment, diagnosis of cirrhosis, availability of a stored frozen serum sample collected after DAA therapy and at least 6 months of follow-up (FU) from the date of serum sample collection.

The exclusion criteria were previous or active HCC at DAA initiation, HCC diagnosis within 6 months from serum sample collection, FU < 6 months from the date of sample collection, no SVR to antiviral treatment, concomitant treatment with vitamin K antagonists, human immunodeficiency virus coinfection, hepatitis B surface antigen positivity and weekly alcohol consumption of > 140 g in women and > 210 g in men. Patients with missing clinical and biochemical data and those who did not provide written informed consent were also excluded.

Cirrhosis was diagnosed histologically or using vibration-controlled transient elastography (VCTE; FibroScan, Echosens, Paris, France) showing liver stiffness measurement (LSM) > 11.9 kPa [13], and/or by indirect signs of portal hypertension (i.e. abdominal collateral circles, platelet count $< 150 \times 10^9$/L, oesophageal varices). The Albumin-Bilirubin (ALBI) grade was calculated according to Johnson et al. [14] SVR to DAA was defined as the clearance of HCV RNA 12 weeks after the end of treatment (EOT) [15].

The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Institutional Ethics Committee of each centre. Written informed consent was obtained from all the patients included in the study.

2.2 | Post Therapy Follow-Up and Study Endpoints

All patients underwent regular clinical FU with scheduled visits at month 1, 3 and 6 months after EOT, and every 6 months thereafter. At each time point, routine laboratory tests were performed and the results were collected in a dedicated database. US examination was performed at SVR12 (training cohort) and EOT (validation cohort) to exclude HCC at baseline and every 3–6 months thereafter until HCC detection (primary outcome), death or last FU. HCC was diagnosed based on non-invasive criteria and/or pathology following the European Association for the Study of the Liver (EASL) recommendations; the Barcelona Clinic Liver Cancer (BCLC) staging system was adopted for HCC staging and treatment allocation [5].

2.3 | Non-Invasive Tests and Prognostic Scores

We investigated the predictive value of the following non-invasive tests, calculated according to their original formula:

- Aspartate aminotransferase (AST) to platelets ratio index (APRI): [16] (AST/AST upper limit normal)/platelets $[10^9/L] \times 100$
- Fibrosis 4 (FIB-4) index: [17] (Age [years] × AST [U/L])/(platelets $[10^9/L] \times \sqrt{}$alanine aminotransferase (ALT) [U/L]).

The prognostic score age–male–ALBI–platelets (aMAP) and the Toronto HCC Risk Index (THRI) were calculated as follows:

- aMAP score: [18] $((0.06 \times \text{age [years]} + 0.89 \times \text{sex [male: 1; female: 0]} + (0.48 \times (\log \text{total bilirubin } (\mu\text{mol/L}) \times 0.66) + (\text{albumin [g/L]} \times -0.085)) - 0.01 \times \text{platelets } (10^3/\text{mm}^3) + 7.4)/14.77) \times 100$
- THRI: [19] age (<45 years = 0 points; 45–60 years = 50 points; >60 years = 100 points) + aetiology (HCV SVR = 0 points) + gender (female = 0 points; male = 80 points) + platelet count (>200 × 10^9/L = 0 points; 140–200 × 10^9/L = 20 points; 80–139 × 10^9/L = 70 points; <80 × 109/L = 89 points).

2.4 | Serum PIVKA-II Measurement

PIVKA-II was retrospectively measured on patients' serum samples collected at SVR12 (Turin and Pisa cohorts) or EOT (Milan cohort), and stored at −20°C. The personnel performing the analyses were blinded to the patients' clinical characteristics. PIVKA-II serum values were measured using chemiluminescent enzyme immunoassay on the fully automated LUMIPULSE G600 II or G1200 platform (Fujirebio Inc. Tokyo, Japan) using Lumipulse G PIVKA-II immunoreaction cartridges (precision <4.4%; lower limit of detection = 1.37 mAU/mL), according to manufacturer recommendation. To further check the analytical stability of PIVKA-II in serum, we performed a study in samples stored at different designated temperatures (−20°C, +4°C and room temperature) for different periods (0, 7 and 14 days). The coefficients of variation were calculated as the ratio of the standard deviation to the mean and reported as percentages (CV%).

2.5 | Statistical Analysis

Continuous variables were reported as the median and interquartile range (IQR), while categorical variables were reported as number (n) and percentage (%). Data normality was checked using the D'Agostino–Pearson test (Table S1). Accordingly, continuous variables between unpaired groups were compared using the non-parametric Mann–Whitney *U* test. Fisher's exact test was used to compare the distribution of dichotomous variables between unpaired groups. Non-parametric correlation analysis was performed using Spearman's correlation test. Cox regression analysis was performed to evaluate the association between baseline variables and the occurrence of HCC, and the likelihood of developing the event was reported as a hazard ratio (HR) and 95% confidence interval (CI). Survival analysis was performed using Kaplan–Meier curves; *p* values were calculated by log-rank test. Patients that did not develop HCC were censored at last FU or death.

We performed a model evaluation of the predictive capability of different machine learning methods using cross-validation procedures for the training cohort (Pisa + Turin), keeping the test cohort (Milan) as an external and blind set. We evaluated two linear and additive methods, namely Cox proportional hazards (PH) and penalised Cox regression [20], and three non-linear methods, namely Random Survival Forests [21], Gradient Boosting [22], and DeepCoxNN [23]. The Scikit-Survival Python implementation is the library used to evaluate the methods mentioned above [24], with the exception of DeepCoxNN implemented in-house. Feature relevance was assessed using an ablation test, which consisted of removing each covariate from the initial model and testing the predictive performance of various feature-reduced models. The validity of the CoxPH model assumptions was checked using the residuals method described by Grambsch et al. [25] and chi-squared statistics based on the lifelines software [26]. X-tile plots (X-tile Software v. 3.6.1; https://medicine.yale.edu/lab/rimm/research/software/) were used to identify the two optimal cut-off values with the highest chi-square (χ^2) value to separate patients into low-, medium- and high-risk (https://shap.readthedocs.io). Conventional statistical analysis was performed using MedCalc software, v. 18.9.1 (MedCalc Software Ltd. Ostend, Belgium).

3 | Results

3.1 | Characteristics of Study Cohorts

A total of 1220 patients (Turin cohort, *n* = 531; Pisa cohort, *n* = 335; Milan cohort, *n* = 354) with HCV-related cirrhosis treated with DAAs were included in the study. Based on the sample size, timing of data collection and timing of serum PIVKA-II measurement, patients enrolled in the Liver Unit of the Department of Medical Sciences, University of Torino (Turin) and at the Hepatology Unit of the University Hospital of Pisa (Pisa) were assigned to the training cohort, whereas patients enrolled at the Division of Gastroenterology and Hepatology, Ospedale Maggiore Policlinico di Milano (Milan) were assigned to the validation cohort (Figure 1).

The principal demographic, biochemical and clinical characteristics of the study cohort are presented in Table 1. In the entire population, the median age was 66 (57–76) years; most patients were men (*n* = 707; 58.0%). In comparison with the training cohort, patients in the validation cohort showed a higher rate of markers consistent with previous hepatitis B virus (HBV) exposure (39.0% vs. 18.1%; $p < 0.001$), lower serum albumin values (4.2, 4.0–4.4 g/dL vs. 4.4, 4.1–4.6 g/dL; $p < 0.001$), higher total bilirubin levels (0.8, 0.6–1.3 mg/dL vs. 0.7, 0.5–0.9 mg/dL; $p < 0.001$), higher platelet count (146, 100–200 × 10^9/L vs. 131, 93–173 × 10^9/L; $p < 0.001$), lower total cholesterol values (156, 136–182 mg/dL vs. 168, 146–194 mg/dL; $p < 0.001$), lower LSM (16.9, 13.3–21.3 kPa vs. 19.6, 15.1–26.4 kPa; $p < 0.001$), higher proportion of Child-Turcotte-Pugh (CTP) score A (97.7% vs. 94.8%; $p = 0.020$) and lower PIVKA-II serum concentration (34, 26–47 mAU/mL vs. 46, 37–58 mAU/mL; $p < 0.001$). Of note, serum PIVKA-II values were significantly higher in patients with CTP-B as compared to patients with CTP-A (49, 39–74 mAU/mL vs. 42, 33–56 mAU/mL; $p = 0.004$) and were weakly correlated to LSM ($r_s = 0.14$, 95% CI 0.09–0.20; $p < 0.001$).

Overall, median APRI and FIB-4 values were 0.5 (0.3–0.7) and 2.73 (1.80–4.34), respectively. While APRI values were comparable between the training and validation cohort (0.5, 0.3–0.7 vs. 0.4, 0.3–0.8; $p = 0.107$), FIB-4 values were significantly higher in the former as compared to the latter (2.81, 1.84–4.41 vs. 2.43, 1.63–4.18; $p = 0.013$). Concerning the aMAP score, the median value in the entire cohort of 1220 patients was 61 (57–66); no

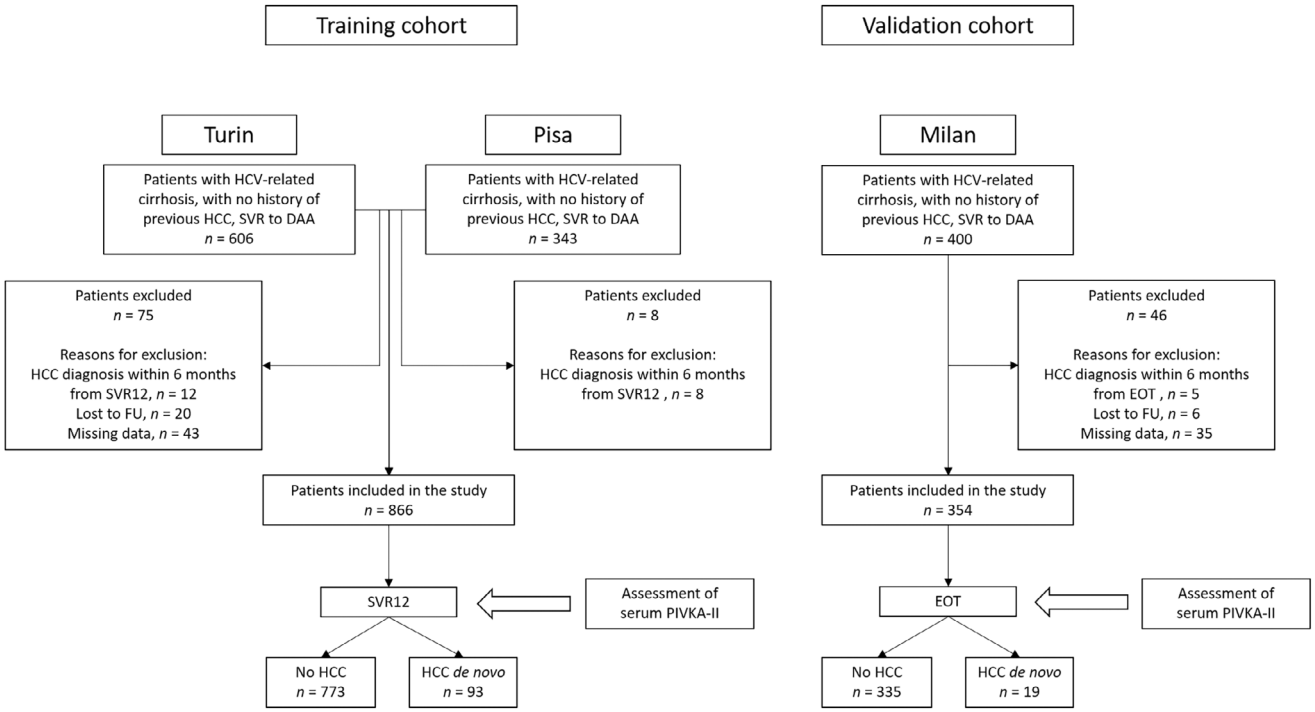

FIGURE 1 | Flow chart of patients' enrolment in the training (Turin + Pisa) and validation cohorts (Milan). Patients of the Milan validation cohort were retrieved from Degasperi et al. [9] Patients HCC-free at last visit/US examination but with FU < 6 months from SVR12 (training cohort) or EOT (validation cohort) were defined as lost to FU. DAA, direct-acting antivirals; EOT, end-of-treatment; FU, follow-up; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; PIVKA-II, protein induced by vitamin K absence or antagonist II; SVR, sustained virologic response.

differences were observed between the training and validation cohorts (62, 57–66 vs. 61, 56–66; $p = 0.073$). Finally, the median THRI value in the whole study cohort was 170 (130–200); higher THRI values were observed in the training cohort compared to the validation cohort (180, 130–200 vs. 170, 120–200; $p = 0.003$).

During a median FU of 44 (25–54) months (FU < 12 months, $n = 119$ [9.8%]; FU = 12–24 months, $n = 170$ [13.9%]; FU > 24 months, $n = 931$ [76.3%]), 112 of 1220 patients developed HCC (incidence rate: 2.71 per 100 person-years). In the training cohort, 93 patients developed HCC during a median FU of 39 (22–55) months (cumulative incidence: 10.7%; incidence rate: 3.24 per 100 person-year), while in the validation cohort, 19 patients developed HCC during a median FU of 49 (35–52) months (cumulative incidence: 5.4%; incidence rate: 1.51 per 100 person-year). At HCC diagnosis, the proportion of early tumour (BCLC 0/A) was 80.6% in the training cohort and 84.2% in the validation cohort ($p = 0.718$). Serum PIVKA-II values were significantly associated with HCC occurrence both in the training (HR = 2.51, 95% CI 1.61–3.90; $p < 0.001$) and in the validation cohorts (HR = 13.92, 95% CI 5.33–36.37; $p < 0.001$).

3.2 | Assessment of PIVKA-II Stability in Serum Samples

Serum samples from three different patients collected for PIVKA-II measurement and then stored at −20°C, were retrospectively analysed. The PIVKA-II levels at serum sample collection (baseline) were 2642 mAU/mL (sample ID: #303), 237 mAU/mL (sample ID: #335) and 50 mAU/mL (sample ID: #299). After a mean of 20 ± 6 days of storage at −20°C, the 3 samples were thawed and tested for PIVKA-II. Then, 4 aliquots for each serum sample were prepared; 2 aliquots were stored at +4°C and 2 at room temperature (average temperature: +21°C). After 7 and 14 days, the aliquoted serum samples were tested for PIVKA-II (Table S2); overtime percentage changes in samples stored at +4°C and at room temperature are reported in Figure 2. The overall CV% of PIVKA-II measurements was 3.5%, ranging from 2.0% (sample #303, stored at room temperature) to 6.2% (sample #335, stored at room temperature); no differences in CV% were observed between the whole set of aliquots stored at +4°C and aliquots stored at room temperature (3.4% vs. 3.7%; $p = 0.726$).

3.3 | Development and Validation of the PIVKA-II-Based Model for HCC Prediction

First, we selected demographic and biochemical features reported to be associated with HCC occurrence in previous studies as candidate features for model construction. Apart from PIVKA-II which was selected a priori, we considered the following covariates: age, sex, ALT, AST, γ-glutamyl transferase (γGT), platelet count, albumin and total bilirubin. A log transformation was made to PIVKA-II due to extreme skewness (Figure S1). Then, four machine learning-based methods, namely, Gradient Boosting, Random Survival Forests, DeepCoxNet and CoxPH regression, were trained and compared to determine the best prediction model. Non-linear methods showed no improvement in terms of the C-index over the CoxPH regression (Table S3); thus, we chose the CoxPH model and tested the relevance of the features using an ablation test. Because we observed that, in all simulations, the removal of a single covariate led to a reduction

TABLE 1 | Baseline demographic, biochemical and clinical characteristics of the patients included in the study according to training (Turin + Pisa) and validation cohorts (Milan).

Variables	Overall ($n=1220$)	Training[a] ($n=866$)	Validation[b] ($n=354$)	p
Age (years), median (range)	66 (57–76)	66 (57–76)	66 (57–75)	0.293
Sex (male), *n* (%)	707 (58.0%)	507 (58.5%)	200 (56.5%)	0.511
Caucasian, *n* (%)	1184 (97.0%)	855 (98.7%)	329 (92.9%)	<0.001
BMI (kg/m^2), median (IQR)	24.9 (22.7–27.7)	25.0 (22.8–27.7)	24.8 (22.4–27.5)	0.858
T2DM, *n* (%)	243 (19.9%)	175 (20.2%)	68 (19.2%)	0.692
Previous HBV exposure, *n* (%)	295 (24.2%)	157 (18.1%)	138 (39.0%)	<0.001
ALT (U/L), median (IQR)	19 (15–26)	19 (15–26)	19 (14–27)	0.999
AST (U/L), median (IQR)	24 (20–30)	24 (20–29)	25 (20–30)	0.382
γGT (U/L), median (IQR)	26 (18–40)	27 (18–41)	25 (18–35)	0.094
Albumin (g/dL), median (IQR)	4.3 (4.1–4.6)	4.4 (4.1–4.6)	4.2 (4.0–4.4)	<0.001
Total bilirubin (mg/dL), median (IQR)	0.7 (0.5–1.0)	0.7 (0.5–0.9)	0.8 (0.6–1.3)	<0.001
Platelet count ($\times 10^9$/L), median (IQR)	134 (94–181)	131 (93–173)	146 (100–200)	0.001
Total cholesterol (mg/dL), median (IQR)	166 (143–191)	168 (146–194)	156 (136–182)	<0.001
Triglycerides (mg/dL), median (IQR)	92 (74–116)	91 (72–120)	95 (79–108)	0.420
AFP (ng/mL), median (IQR)[c]	/	/	5.0 (3.0–8.0)	/
PIVKA-II (mAU/mL), median (IQR)	43 (33–56)	46 (37–58)	34 (26–47)	<0.001
LSM (kPa), median (IQR)	18.2 (14.5–25.7)	19.6 (15.1–26.4)	16.9 (13.3–21.3)	<0.001
LSM < 15.0 kPa, *n* (%)	338 (27.7%)	208 (24.0%)	130 (36.7%)	
LSM 15.0–19.9 kPa, *n* (%)	406 (33.3%)	289 (33.4%)	117 (33.1%)	<0.001
LSM ≥ 20.0 kPa, *n* (%)	476 (39.0%)	369 (42.6%)	107 (30.2%)	
CTP score A, *n* (%)	1167 (95.7%)	821 (94.8%)	346 (97.7%)	0.020
ALBI score, median (IQR)	−2.98 (−3.21– −2.70)	−3.05 (−3.25– −2.81)	−2.79 (−3.02– −2.56)	<0.001
De novo HCC, *n* (%)	112 (9.2%)	93 (10.7%)	19 (5.4%)	0.003
BCLC staging at diagnosis				
0, *n* (%)	33 (29.5%)	31 (33.3%)	2 (10.5%)	0.299
A, *n* (%)	58 (51.8%)	44 (47.3%)	14 (73.7%)	
B, *n* (%)	12 (10.7%)	11 (11.8%)	1 (5.3%)	
C, *n* (%)	9 (8.0%)	7 (7.5%)	2 (10.5%)	
Monofocal HCC, *n* (%)	69 (61.6%)	61 (65.6%)	9 (47.4%)	0.193
Major HCC nodule size (mm), median (IQR)	21 (15–34)	21 (15–35)	20 (15–23)	0.411
FU (months), median (IQR)	44 (25–54)	39 (22–55)	49 (35–52)	0.001
HCC incidence rate (per 100 person/years)	2.71	3.24	1.51	0.017

Note: p value for BCLC comparison between the training and validation cohort was calculated by χ^2 test for trend.
Abbreviations: AFP, α-fetoprotein; ALBI, albumin-bilirubin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancers; BMI, body mass index; CTP, Child-Turcotte-Pugh; FU, follow-up; γGT, gamma-glutamyl transpeptidase; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; IQR, interquartile range; *n*, number; LSM, liver stiffness measurement; PIVKA-II, protein induced by vitamin K absence or antagonist II; T2DM, type 2 diabetes mellitus.
[a]Data were collected at SVR12.
[b]Data were collected at EOT.
[c]AFP values were systematically available only for the validation cohort.

in model performance (ablation test; Figure S2), all the initially selected variables were retained in the final CoxPH model. Furthermore, all the nine variables included in the model met the assumptions of CoxPH (χ^2 $p>0.05$) (Table S4). The final CoxPH model included age, sex, ALT, AST, γGT, platelet count, albumin, total bilirubin and PIVKA-II (event-per-predictor

parameter ratio: 12.4). Finally, we implemented a web server calculator that is freely available to the scientific community at the website https://compbiomed.hpc4ai.unito.it/pivka-hcc/.

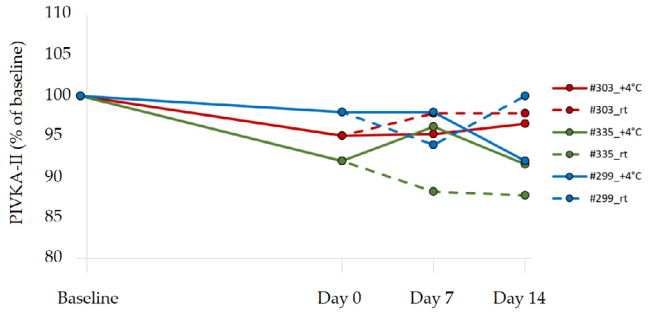

FIGURE 2 | Overtime stability of PIVKA-II in serum according to storage temperature. All samples were stored at −20°C from baseline to day 0 for a mean of 20 ± 6 days. Red lines indicate the sample at high serum PIVKA-II concentration (baseline: 2642 mAU/mL), green lines indicate the sample at intermediate serum PIVKA-II concentration (baseline: 237 mAU/mL), and blue lines indicate the sample at low serum PIVKA-II concentration (baseline: 50 mAU/mL). Solid lines indicate aliquots stored at +4°C from day 0 to day 14, while dashed lines indicate aliquots stored at room temperature from day 0 to day 14. PIVKA-II, protein induced by vitamin K absence or antagonist II; rt., room temperature.

The model showed a C-index of 0.72 in the training cohort, which was calculated using a 1000-fold bootstrapping approach. The performance of the model was significantly better than that of APRI (C-index = 0.64), FIB-4 (C-index = 0.65), aMAP (C-index = 0.65) and THRI (C-index = 0.63). Similarly, the model confirmed adequate accuracy for HCC prediction in the validation cohort (C-index = 0.71) (feature relevance is shown as SHAP beeswarm plot in Figure S3), with significantly superior performance compared to APRI (C-index = 0.60), FIB-4 (C-index = 0.57), aMAP (C-index = 0.61), THRI (C-index = 0.57) and AFP (C-index = 0.67) (Table 2). Dynamic AUC analysis revealed that our PIVKA-II-based model had a persistently good accuracy over time (Figure 3), with a 12-month AUC of 0.73 and 0.84 in training and validation cohort, respectively (Table S5). Results from direct comparison between the performance of the PIVKA-II-based model and the PIVKA-II-ablated model (model with 8 covariates without PIVKA-II) are reported in Figure S4.

3.4 | HCC Risk Stratification According to PIVKA-II-Based Model

Based on X-tile plots and the recommendation of the EASL policy statement on risk-based surveillance [27, 28], we identified

TABLE 2 | Performance of non-invasive scores and PIVKA-II-based model for HCC prediction according to training (Turin + Pisa) and validation cohorts (Milan).

Variables	Overall	Training	Validation
	C-index (95% CI)	C-index (95% CI)	C-index (95% CI)
APRI	0.64 (0.58–0.70)	0.64 (0.57–0.70)	0.60 (0.45–0.75)
FIB-4	0.64 (0.59–0.70)	0.65 (0.59–0.71)	0.57 (0.42–0.72)
aMAP	0.65 (0.60–0.70)	0.65 (0.59–0.71)	0.61 (0.46–0.75)
THRI	0.63 (0.58–0.68)	0.63 (0.57–0.69)	0.57 (0.44–0.71)
AFP	/	/	0.67 (0.53–0.81)
PIVKA-II-based model	0.72 (0.67–0.77)	0.72 (0.69–0.75)	0.71 (0.66–0.75)

Note: PIVKA-II-based model: age + sex + ALT + AST + γGT + platelet count + albumin + total bilirubin + PIVKA-II.
Abbreviations: AFP, α-fetoprotein; aMAP, age–male–ALBI–platelets score; APRI, AST to platelets ratio index; CI, confidence interval; FIB-4, fibrosis 4 index; PIVKA-II, protein induced by vitamin K absence or antagonist II; THRI, Toronto HCC risk index.

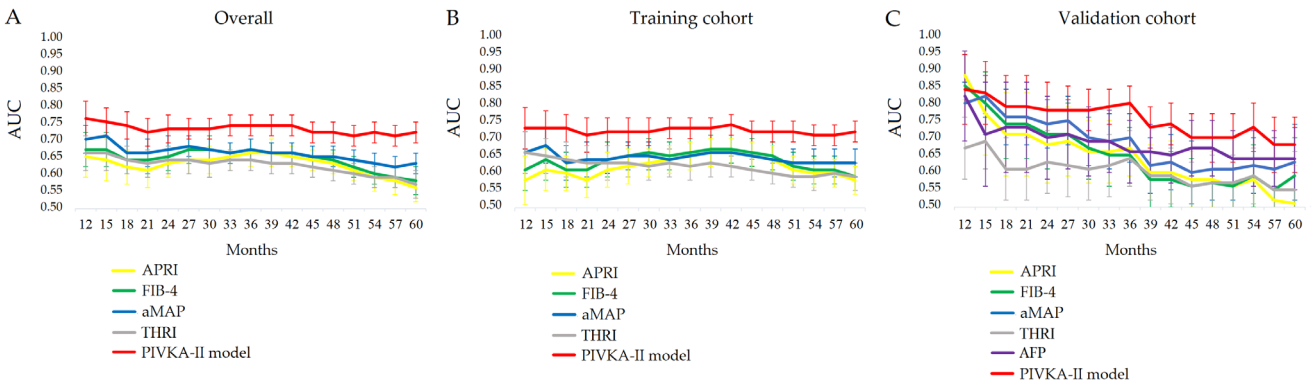

FIGURE 3 | Dynamic AUC of non-invasive scores and PIVKA-II-based model for HCC detection in the overall (A), training (Turin + Pisa) (B) and validation cohorts (Milan) (C). Depicted data refer to the AUC ± SE at each time point. AFP, α-fetoprotein; aMAP, age–male–ALBI–platelets score; APRI, AST to platelets ratio index; AUC, area under the curve; FIB-4, fibrosis 4 index; PIVKA-II, protein induced by vitamin K absence or antagonist II; SE, standard error; THRI, Toronto HCC Risk Index.

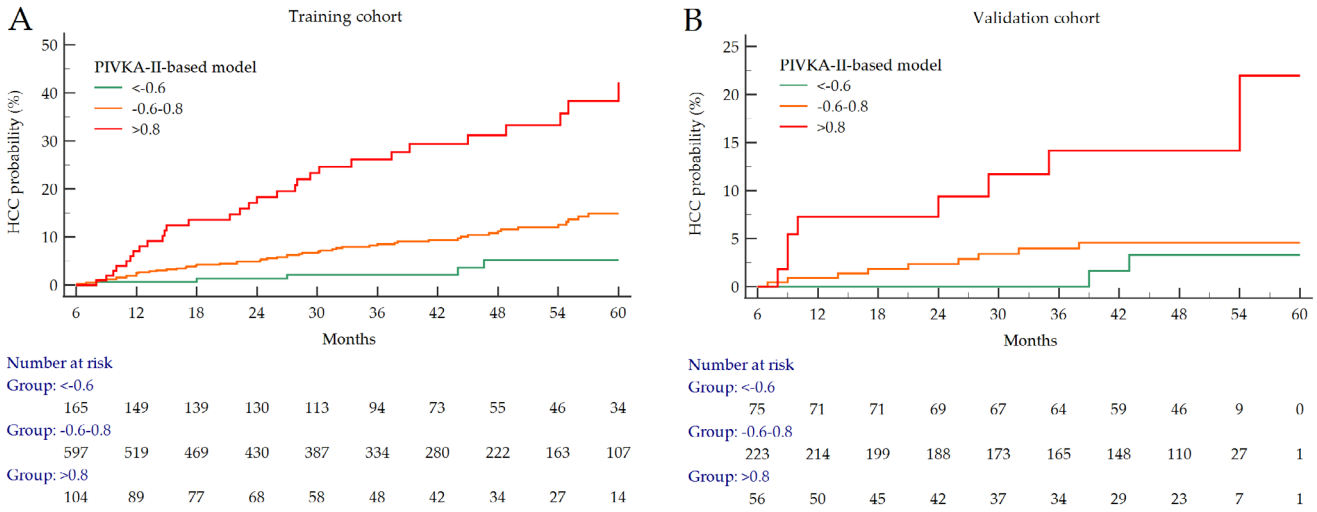

FIGURE 4 | Patients' stratification according to PIVKA-II-based model risk categories in the training (A) and validation cohorts (B). HCC, hepatocellular carcinoma; PIVKA-II, protein induced by vitamin K absence or antagonist II.

two optimal cut-off values to separate the training cohort into low- (< −0.6), medium- (−0.6–0.8) and high-risk groups (>0.8) (Figure S5); accordingly, 165 (19.1%) patients were assigned to the low-risk group, 597 (68.9%) to the medium-risk group and 104 (12.0%) to the high-risk group. Thus, in the training cohort, the cumulative incidence of HCC increased stepwise from low (3.6%) to medium (9.4%) and further increased in the high-risk group (29.8%) (log-rank test, $p < 0.001$) (Figure 4A). As compared with the low-risk group, both the medium- and high-risk groups showed a significantly higher likelihood of developing HCC during the FU (HR = 2.67, 95% CI 1.59–4.48 and HR = 9.28, 95% CI 4.32–19.95, respectively).

Next, we applied the same cut-off values to the validation cohort and observed good discrimination among the three risk groups; the cumulative incidence of HCC was 2.7% (2/75), 4.0% (9/223) and 14.3% (8/56) in the low-, medium- and high-risk groups, respectively (log-rank test, $p < 0.001$) (Figure 4B). The likelihood of developing HCC of the medium- and high-risk group as compared to the low-risk group was HR = 1.64 (95% CI 0.55–4.92) and HR = 6.31 (95% CI 1.39–28.60), respectively. Of note, only two patients were incorrectly predicted as low-risk; however, in both cases, HCC developed more than 3 years after the PIVKA-II-based model assessment. The major contributors to prediction failure are reported in SHAP (Figure S6).

Finally, we performed a subanalysis stratifying patients according to LSM ≥ 15.0 kPa (highly suggestive of compensated advanced chronic liver disease) and ≥ 20.0 kPa (clinically significant portal hypertension) [29], and to ALBI grade 1 and ALBI grade 2. The effectiveness of the prognostic stratification was maintained in all patient subgroups (Tables S6 and S7).

Overall, among patients classified as low risk, no HCC occurred within 18 and 39 months of FU in the training and validation cohorts, respectively. The PIVKA-II-based model provided more effective patient stratification than the APRI, FIB-4, IMAP and THRI in the training and validation cohorts (Table 3, Table S8, Figures S7 and S8). Sensitivity, specificity, positive-predictive values (PPVs) and negative-predictive values (NPVs) of APRI, FIB-4, aMAP, THRI and PIVKA-II-based model at different cut-offs were summarised in Table S9.

In the entire cohort of 1220 patients, the HCC incidence was 0.93 per 100 person-years in the low-risk group (PIVKA-II-based model < −0.6), 2.34 per 100 person-years in the medium-risk group (PIVKA-II-based model −0.6–0.8) and 7.96 per 100 person-years in the high-risk group (PIVKA-II-based model > 0.8).

4 | Discussion

In the present study, we developed and validated a PIVKA-II-based model for predicting HCC development in patients with HCV-related cirrhosis who were cured of HCV infection using DAAs. The model consists of nine variables: age, sex, ALT, AST, γGT, platelet count, albumin, total bilirubin and PIVKA-II collected after treatment completion. The model showed satisfactory accuracy for predicting HCC in both the training (C-index = 0.72) and validation cohorts (C-index = 0.71). Furthermore, our model proved to perform well (AUC > 0.8) for the identification of patients who developed HCC in short-term FU. The model allowed for the proper stratification of patients into three different risk categories according to the individual probability of HCC development, which may help define personalised surveillance strategies.

Since the availability of DAA, the number of patients cured of chronic HCV infection has steadily grown [30]. Whether or not we will achieve the ambitious World Health Organisation's target for hepatitis C elimination by 2030, the actual cohorts of aging patients at risk of HCC development will remain a medical issue for the years to come. Indeed, a recent projection estimated a 55% increase in the burden of primary liver cancer by 2040, with 905,700 new cases diagnosed in 2020 and 1.4 million people diagnosed in 2040 [31]. In our series, we reported an overall HCC incidence of 2.71 per 100 person-years despite HCV treatment. This percentage is far beyond the cost-effectiveness threshold of 1.5%, warranting post-treatment surveillance for the risk of HCC development [32]. The use of a blood-based

TABLE 3 | Incidence and hazard ratios of HCC development according to non-invasive scores and PIVKA-II-based model risk categories in the overall cohort, and in the training and validation cohorts.

	Overall			Training			Validation		
	HCC incidence			HCC incidence			HCC incidence		
	n (%)	*p*	HR (95% CI)	*n* (%)	*p*	HR (95% CI)	*n* (%)	*p*	HR (95% CI)
APRI									
< 0.5	38/666 (5.7%)	< 0.001	Ref.	30/470 (6.4%)	< 0.001	Ref.	8/196 (4.1%)	< 0.001	Ref.
0.5–1.5	54/488 (11.1%)		1.80 (1.23–2.64)	47/346 (13.4%)		1.91 (2.16–2.91)	7/142 (4.9%)		1.22 (0.48–3.09)
> 1.5	20/66 (30.3%)		6.60 (2.60–16.74)	16/50 (32.0%)		5.99 (2.24–15.99)	4/16 (25.0%)		8.15 (0.67–98.60)
FIB-4									
< 1.3	6/124 (4.8%)	< 0.001	Ref.	3/82 (3.7%)	< 0.001	Ref.	3/42 (7.1%)	0.349	Ref.
1.3–3.25	36/615 (5.9%)		1.10 (0.57–2.13)	29/429 (6.8%)		1.62 (0.76–3.48)	7/186 (3.8%)		0.51 (0.11–2.18)
> 3.25	70/481 (14.6%)		2.72 (1.38–2.13)	61/355 (17.2%)		4.03 (1.87–3.80)	9/126 (7.1%)		1.00 (0.22–4.59)
aMAP									
< 50	4/66 (6.1%)	< 0.001	Ref.	1/45 (2.2%)	0.006	Ref.	3/21 (14.3%)	0.013	Ref.
50–60	22/449 (4.9%)		0.79 (0.33–1.89)	20/308 (6.5%)		2.80 (1.03–7.63)	2/141 (1.4%)		0.10 (0.01–0.70)
> 60	86/705 (12.2%)		1.88 (0.80–4.40)	72/513 (14.0%)		5.46 (2.06–14.46)	14/192 (7.3%)		0.56 (0.08–3.73)
THRI									
< 120	4/138 (2.9%)	< 0.001	Ref.	1/83 (1.2%)	< 0.001	Ref.	3/55 (5.5%)	0.518	Ref.
120–240	74/890 (8.3%)		2.61 (1.41–4.83)	62/639 (9.7%)		6.84 (3.23–14.48)	12/251 (4.8%)		0.86 (0.24–3.04)
> 240	34/192 (17.7%)		5.42 (2.60–11.33)	30/144 (20.8%)		13.63 (5.78–32.16)	4/48 (8.3%)		1.66 (0.30–9.14)
PIVKA-II-model									
< −0.6	8/240 (3.3%)	< 0.001	Ref.	6/165 (3.6%)	< 0.001	Ref.	2/75 (2.7%)	< 0.001	Ref.
−0.6–0.8	65/820 (7.9%)		2.51 (1.57–4.00)	56/597 (9.0%)		2.67 (1.59–4.48)	9/223 (4.0%)		1.64 (0.55–4.92)
> 0.8	39/160 (24.4%)		8.44 (4.31–16.53)	31/104 (29.8%)		9.29 (4.32–19.95)	8/56 (14.3%)		6.31 (1.39–28.60)

Note: p values were calculated by log-rank test. HR and 95% CI were estimated by CoxPH regression analysis. PIVKA-II-based model cut-offs were calculated by X-tile software, while NIT values were categorised according to the most widely adopted cut-offs.

Abbreviations: aMAP, age–male–ALBI–platelets score; APRI, AST to platelets ratio index; CI, confidence interval; FIB-4, fibrosis 4 index; HR, hazard ratio; *n*, number PIVKA-II, protein induced by vitamin K absence or antagonist II; THRI, Toronto HCC Risk Index.

predictive model may be useful to personalise the modality and intensity of surveillance. Considering that the number of HCC surveillance candidates cured of HCV infection will increase more than 6-fold by 2030 [33], the identification of patients deserving less intensive surveillance due to the minimal risk of HCC development will save resources for closer monitoring of patients at higher risk, improving early diagnosis and thus patients' outcomes. In this regard, our PIVKA-II-based model allowed the identification of a subgroup of patients at low risk of HCC development (0.66 per 100 person-years) in which the cost-effectiveness of the current surveillance strategy might be questioned, and another subgroup of patients at a higher risk of de novo HCC (8.17 per 100 person-years) in which more intensive surveillance (such as screening by abbreviated magnetic resonance imaging instead of US) might be recommended due to the considerable probability of tumour development.

Among the different components of the model, PIVKA-II was the only parameter selected a priori to build the model. Compared to other potential HCC biomarkers, PIVKA-II has several remarkable advantages from both analytical and clinical perspectives. First, the protein is highly stable in the serum, mitigating the pre-analytical source of variation linked to logistic issues and sample processing. Second, the availability of fully automated assays ensures accurate interpretation and standardisation of results among different laboratories and analytical platforms [34]. Third, among the biomarkers that underwent phase II and early phase III validation [7], PIVKA-II showed satisfactory performance for the detection of HCC irrespective of liver disease activity and aetiology [35]. Furthermore, recent data showed that time-related changes of serum PIVKA-II in patients with cirrhosis of different aetiologies (HBV, active HCV, non-viral) were able to predict HCC occurrence prior to imaging discovery, suggesting that elevations in PIVKA-II circulating levels may reflect a cancer-permissive tissue milieu [36, 37]. Similar results have been observed in patients with cirrhosis of viral aetiology undergoing antivirals [9, 10]. Degasperi et al. showed that serum PIVKA-II ≥ 41 mAU/mL at EOT was significantly associated with a higher risk of HCC development in patients with HCV-related cirrhosis treated with DAA (HR = 9.49, 3.58–25.14); the 4-year cumulative probability of HCC was 24% in patients with EOT PIVKA-II ≥ 41 mAU/mL vs. 2% in patients with EOT PIVKA-II < 41 mAU/mL ($p < 0.001$) [9]. Notably, among patients who remained HCC-free during the FU, serum PIVKA-II values did not significantly change from DAA start to EOT (34, 12–520 mAU/mL vs. 34, 16–867 mAU/mL, p = ns) [9]. In patients with HBV-related cirrhosis undergoing nucleos(t)ide analogues treatment, Su et al. found that PIVKA-II > 50 mAU/mL at virologic remission was an independent predictor of HCC occurrence (HR = 2.46, 1.35–4.49); remarkably, in the subgroup of patients with low AFP, PIVKA-II > 50 mAU/mL significantly increased the likelihood of subsequent HCC (HR = 2.45, 1.18–5.10 in patients with AFP ≤ 10 ng/mL and HR = 3.16, 1.74–5.73 in patients with AFP ≤ 20 ng/mL) [10]. In addition, in a subgroup of patients with serum samples collected every 6–12 months until HCC diagnosis or last FU, the authors observed no significant change of PIVKA-II levels in patients without HCC after a mean FU of 58 months. Conversely, in patients that developed HCC serum PIVKA-II values increased from baseline (median: 39 mAU/mL) to 7 months before HCC detection (median: 82 mAU/mL), and further increased at the time of HCC diagnosis (median: 130 mAU/mL) [10]. Taken together, the results from these independent studies support the absence of significant over-time variation of serum PIVKA-II values within HCC-free patients.

In addition to PIVKA-II, our model included demographic features and biochemical parameters normally collected during standard clinical practice. Older age and male sex are well-known risk factors associated with HCC development, as well as biochemical parameters reflecting liver damage (ALT, AST and γGT) and impaired hepatocellular function (platelet count, albumin and total bilirubin) [38]; unsurprisingly, these parameters have already been included in previous non-invasive scores [17, 39, 40]. The inclusion of only demographic and biochemical features that can be objectively measured confers model reliability and reproducibility, thereby limiting potential sources of bias. In addition, the model can be easily implemented in clinical laboratories with the potential to increase adherence to surveillance. Finally, our model proved superior performance compared to other non-patented scores in terms of accuracy for HCC prediction and patients' stratification. However, the relatively low number of events (i.e. incident HCC) in the validation cohort and the cut-off adopted for the definition of the three different risk categories by APRI, FIB-4, aMAP and THRI may have partially affected the results (Supplementary Discussion).

Previous studies have shown that HCC prediction models constructed using artificial intelligence approaches, including complex supervised and non-linear machine learning algorithms outperform traditional regression models built using conventional statistics [41]. In the present study, tests with more complex models, such as Random Survival Forests, Gradient Boosting and DeepCoxNN, showed no improvement over CoxPH, indicating an intrinsically linear and additive effect in risk assignment without synergetic effects in the covariates.

This study has some limitations that need to be acknowledged. First, the retrospective nature of the study may represent a potential source of bias. However, all patients included in the study were consecutively enrolled at each centre, thereby limiting the risk of selection bias. In addition, the model achieved good predictive performance in both the training and validation cohorts, despite some differences in baseline characteristics, HCC incidence rate and discrepancies in the timing of data collection (SVR12 in the training cohort and EOT in the validation cohort). However, we believe that these temporal and clinical differences strengthen the reliability of our model [42]. Indeed, model validation in an external cohort which differed from that used for model development, demonstrated its generalisability and transportability and avoided multiple looks until satisfactory results were obtained, as in the case of random splitting of the whole dataset [43]. Second, all patients were enrolled in Italy and were mostly Caucasians; therefore, we could not evaluate whether the model might confer appropriate performance on patients of different ethnicities. Third, we did not evaluate the time-related changes in the risk of HCC within each patient through repeated assessments of the model; thus, we were unable to recalculate the probability of HCC development at different time points and its association with patient outcomes. In addition, we did not assess the impact of other potential risk modifiers such as a family history of HCC, incident type 2 diabetes, alcohol

consumption and other metabolic comorbidities on liver disease progression and HCC occurrence. However, we would like to emphasise that our model should be intended to complement the decision-making process for patient allocation to the most appropriate surveillance protocol according to the individual risk of HCC development. Last, although our PIVKA-II-based model allowed an effective stratification of the risk of HCC development, the efficiency resulted suboptimal (PPV = 25.6% in the entire population). This is consistent with results from previous studies [44, 45], suggesting that HCC predictive models could be more useful to rule out rather than ruling in patients at risk. Further studies are needed to improve the PIVKA-II-model efficiency possibly through stepwise surveillance approaches [46], which may allow the identification of super high-risk patients among patients at high risk of HCC development.

Finally, we acknowledge the strengths of the proposed model. Compared to other available tools, our model provides not only the calculation of a mere risk class, but also a tool (i.e. SHAP force plot), allowing clinicians to understand which parameters mostly determine the calculation of HCC risk, and thus evaluate the reliability of the prediction based on patients' characteristics. Paradigmatic is the case of the patient (ID: #304) incorrectly predicted to be at low risk of HCC development due to female sex and the conserved hepatic function despite elevated serum PIVKA-II values; the possibility of checking which parameters mostly contributed to patients' risk class allocation, could help clinicians to assess the reliability of prognostic prediction.

In conclusion, we developed and validated a model based on PIVKA-II combined with variables normally collected during standard clinical practice that showed adequate accuracy for the prediction of HCC development in patients with HCV-related cirrhosis successfully treated with DAA; in such patients, our PIVKA-II-based model may allow the transition from a 'one-size-fits-all' to a 'risk-based' surveillance strategy. In particular, our model allowed the identification of patients at low risk of HCC development that may not need surveillance, reducing costs and harms from surveillance. Further studies are warranted to provide prospective validation in a larger cohort of patients and to investigate the performance of the model in large cohorts of patients at risk of HCC with different liver disease aetiologies.

Author Contributions

Gian Paolo Caviglia: conceptualization, formal analysis, investigation, writing – original draft, data curation, methodology, validation, visualization, resources, project administration. **Piero Fariselli:** conceptualization, formal analysis, supervision, writing – original draft, software, methodology, investigation, resources, validation, visualization. **Roberta D'Ambrosio:** data curation, writing – original draft, investigation, resources, validation. **Piero Colombatto:** writing – original draft, data curation, validation, investigation. **Elisabetta Degasperi:** data curation, writing – review and editing, investigation, resources. **Gabriele Ricco:** writing – review and editing, data curation, resources, investigation. **Maria Lorena Abate:** investigation, resources, validation, writing – original draft. **Giovanni Birolo:** investigation, software, formal analysis, methodology, resources, visualization, writing – original draft. **Giulia Troshina:** data curation, writing – review and editing. **Francesco Damone:** data curation, writing – review and editing, investigation. **Barbara Coco:** data curation, writing – review and editing, investigation, resources. **Daniela Cavallone:** investigation, writing – review and editing, validation. **Riccardo Perbellini:** investigation, writing – review and editing, data curation. **Sara Monico:** data curation, writing – review and editing, investigation. **Giorgio Maria Saracco:** writing – review and editing, supervision, resources. **Maurizia Rossana Brunetto:** supervision, writing – review and editing, validation, resources. **Pietro Lampertico:** supervision, writing – review and editing, validation, resources. **Alessia Ciancio:** writing – review and editing, supervision, validation, resources.

Acknowledgements

The authors thank Fujirebio Europe NV for partly supplying PIVKA-II test reagents.

Ethics Statement

The study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Institutional Ethics Committee of each centre. Written informed consent was obtained from all the patients included in the study.

Conflicts of Interest

Gian Paolo Caviglia: research grant from Fujirebio Diagnostics AB; Roberta D'Ambrosio: speaking, teaching and advisory board for AbbVie and Gilead; consultant for Takeda; Piero Colombatto: speaker bureau for AbbVie and Gilead, advisory board for AbbVie; Elisabetta Degasperi: speaking, teaching and research grant for Gilead; Maurizia Rossana Brunetto: Advisory Board and Speakers' Bureau AbbVie, Gilead, Janssen, EISAI-MSD, Roche; Pietro Lampertico: Advisory board/speaker bureau for AbbVie, Aligos, Altona, Antios, Eiger, Gilead Sciences, GlaxoSmithKline, Grifols, Janssen, MYR, Roboscreen, Roche Pharma/Diagnostics, Vir; Alessia Ciancio, research grant for Gilead Sciences.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. C. W. Spearman, G. M. Dusheiko, M. Hellard, and M. Sonderup, "Hepatitis C," *Lancet* 394, no. 10207 (2019): 1451–1466, https://doi.org/10.1016/S0140-6736(19)32320-7.

2. R. Nevola, L. Rinaldi, L. Zeni, et al., "Changes in Clinical Scenarios, Management, and Perspectives of Patients With Chronic Hepatitis C After Viral Clearance by Direct-Acting Antivirals," *Expert Review of Gastroenterology & Hepatology* 15, no. 6 (2021): 643–666, https://doi.org/10.1080/17474124.2021.1877136.

3. I. Lockart, M. G. H. Yeo, B. Hajarizadeh, G. J. Dore, and M. Danta, "HCC Incidence After Hepatitis C Cure Among Patients With Advanced Fibrosis or Cirrhosis: A Meta-Analysis," *Hepatology* 76, no. 1 (2022): 139–154, https://doi.org/10.1002/hep.32341.

4. G. N. Ioannou, L. A. Beste, P. K. Green, et al., "Increased Risk for Hepatocellular Carcinoma Persists up to 10 Years After HCV Eradication in Patients With Baseline Cirrhosis or High FIB-4 Scores," *Gastroenterology* 157, no. 5 (2019): 1264–1278.e4, https://doi.org/10.1053/j.gastro.2019.07.033.

5. European Association for the Study of the Liver, "EASL Clinical Practice Guidelines: Management of Hepatocellular Carcinoma," *Journal of Hepatology* 69, no. 1 (2018): 182–236, https://doi.org/10.1016/j.jhep.2018.03.019.

6. J. A. Marrero, L. M. Kulik, C. B. Sirlin, et al., "Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases," *Hepatology* 68, no. 2 (2018): 723–750, https://doi.org/10.1002/hep.29913.

7. N. D. Parikh, N. Tayob, and A. G. Singal, "Blood-Based Biomarkers for Hepatocellular Carcinoma Screening: Approaching the End of the Ultrasound Era?," *Journal of Hepatology* 78, no. 1 (2023): 207–216, https://doi.org/10.1016/j.jhep.2022.08.036.

8. G. P. Caviglia, D. G. Ribaldone, M. L. Abate, et al., "Performance of Protein Induced by Vitamin K Absence or Antagonist-II Assessed by Chemiluminescence Enzyme Immunoassay for Hepatocellular Carcinoma Detection: A Meta-Analysis," *Scandinavian Journal of Gastroenterology* 53, no. 6 (2018): 734–740, https://doi.org/10.1080/00365521.2018.1459824.

9. E. Degasperi, R. Perbellini, R. D'Ambrosio, et al., "Prothrombin Induced by Vitamin K Absence or Antagonist-II and Alpha Foetoprotein to Predict Development of Hepatocellular Carcinoma in Caucasian Patients With Hepatitis C-Related Cirrhosis Treated With Direct-Acting Antiviral Agents," *Alimentary Pharmacology & Therapeutics* 55, no. 3 (2022): 350–359, https://doi.org/10.1111/apt.16685.

10. T. H. Su, C. Y. Peng, S. H. Chang, et al., "Serum PIVKA-II and Alpha-Fetoprotein at Virological Remission Predicts Hepatocellular Carcinoma in Chronic Hepatitis B Related Cirrhosis," *Journal of the Formosan Medical Association* 121 (2022): 703–711, https://doi.org/10.1016/j.jfma.2021.08.003.

11. R. D'Ambrosio, E. Degasperi, and P. Lampertico, "Predicting Hepatocellular Carcinoma Risk in Patients With Chronic HCV Infection and a Sustained Virological Response to Direct-Acting Antivirals," *Journal of Hepatocellular Carcinoma* 8 (2021): 713–719, https://doi.org/10.2147/JHC.S292139.

12. H. Y. Kim, P. Lampertico, J. Y. Nam, et al., "An Artificial Intelligence Model to Predict Hepatocellular Carcinoma Risk in Korean and Caucasian Patients With Chronic Hepatitis B," *Journal of Hepatology* 76, no. 2 (2022): 311–318, https://doi.org/10.1016/j.jhep.2021.09.025.

13. M. Fraquelli, C. Rigamonti, G. Casazza, et al., "Reproducibility of Transient Elastography in the Evaluation of Liver Fibrosis in Patients With Chronic Liver Disease," *Gut* 56, no. 7 (2007): 968–973, https://doi.org/10.1136/gut.2006.111302.

14. P. J. Johnson, S. Berhane, C. Kagebayashi, et al., "Assessment of Liver Function in Patients With Hepatocellular Carcinoma: A New Evidence-Based Approach-The ALBI Grade," *Journal of Clinical Oncology* 33 (2015): 550–558, https://doi.org/10.1200/JCO.2014.57.9151.

15. European Association for the Study of the Liver, "EASL Recommendations on Treatment of Hepatitis C 2016," *Journal of Hepatology* 66, no. 1 (2017): 153–194, https://doi.org/10.1016/j.jhep.2016.09.001.

16. C. T. Wai, J. K. Greenson, R. J. Fontana, et al., "A Smple Noninvasive Index Can Predict Both Significant Fibrosis and Cirrhosis in Patients With Chronic Hepatitis C," *Hepatology* 38, no. 2 (2003): 518–526, https://doi.org/10.1053/jhep.2003.50346.

17. R. K. Sterling, E. Lissen, N. Clumeck, et al., "Development of a Simple Noninvasive Index to Predict Significant Fibrosis in Patients With HIV/HCV Coinfection," *Hepatology* 43, no. 6 (2006): 1317–1325, https://doi.org/10.1002/hep.21178.

18. R. Fan, G. Papatheodoridis, J. Sun, et al., "aMAP Risk Score Predicts Hepatocellular Carcinoma Development in Patients With Chronic Hepatitis," *Journal of Hepatology* 73, no. 6 (2020): 1368–1378, https://doi.org/10.1016/j.jhep.2020.07.025.

19. S. A. Sharma, M. Kowgier, B. E. Hansen, et al., "Toronto HCC Risk Index: A Validated Scoring System to Predict 10-Year Risk of HCC in Patients With Cirrhosis," *Journal of Hepatology* 68, no. 1 (2018): 92–99, https://doi.org/10.1016/j.jhep.2017.07.033.

20. N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, "Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent," *Journal of Statistical Software* 39, no. 5 (2011): 1–13, https://doi.org/10.18637/jss.v039.i05.

21. H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, "Random Survival Forests," *Annals of Applied Statistics* 2, no. 3 (2008): 841–860, https://doi.org/10.1214/08-AOAS169.

22. J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics* 29, no. 5 (2001): 1189–1232, https://doi.org/10.1214/aos/1013203451.

23. D. Faraggi and R. Simon, "A Neural Network Model for Survival Data," *Statistics in Medicine* 14, no. 1 (1995): 73–82, https://doi.org/10.1002/sim.4780140108.

24. S. Pölsterl, "Scikit-Survival: A Library for Time-To-Event Analysis Built on Top of Scikit-Learn," *Journal of Machine Learning Research* 21, no. 212 (2020): 1–6.

25. P. M. Grambsch and T. M. Therneau, "Proportional Hazards Tests and Diagnostic Based on Weighted Residuals," *Biometrika* 81, no. 3 (1994): 515–526, https://doi.org/10.1093/biomet/81.3.515.

26. C. Davidson-Pilon, "Lifelines: Survival Analysis in Python," *Journal of Open Source Software* 4, no. 40 (2019): 1317, https://doi.org/10.21105/joss.01317.

27. R. L. Camp, M. Dolled-Filhart, and D. L. Rimm, "X-Tile: A New Bio-Informatics Tool for Biomarker Assessment and Outcome-Based Cut-Point Optimization," *Clinical Cancer Research* 10, no. 21 (2004): 7252–7259, https://doi.org/10.1158/1078-0432.CCR-04-0713.

28. European Association for the Study of the Liver. EASL Policy Statement: Risk-Based Surveillance for Hepatocellular Carcinoma Among Patients with Cirrhosis, accessed March 11, 2024, http://easl.eu/publication/easl-policy-statement-risk-based/.

29. R. de Franchis, J. Bosch, G. Garcia-Tsao, T. Reiberger, and C. Ripoll, "Baveno VII - Renewing Consensus in Portal Hypertension," *Journal of Hepatology* 76, no. 4 (2022): 959–974, https://doi.org/10.1016/j.jhep.2021.12.022.

30. Q. Chen, T. Ayer, E. Bethea, et al., "Changes in Hepatitis C Burden and Treatment Trends in Europe During the Era of Direct-Acting Antivirals: A Modelling Study," *British Medical Journal Open* 9, no. 6 (2019): e026726, https://doi.org/10.1136/bmjopen-2018-026726.

31. H. Rumgay, M. Arnold, J. Ferlay, et al., "Global Burden of Primary Liver Cancer in 2020 and Predictions to 2040," *Journal of Hepatology* 77, no. 6 (2022): 1598–1606, https://doi.org/10.1016/j.jhep.2022.08.021.

32. F. P. Sarasin, E. Giostra, and A. Hadengue, "Cost-Effectiveness of Screening for Detection of Small Hepatocellular Carcinoma in Western Patients With Child-Pugh Class A Cirrhosis," *American Journal of Medicine* 101, no. 4 (1996): 422–434, https://doi.org/10.1016/S0002-9343(96)00197-0.

33. Q. Chen, T. Ayer, M. G. Adee, X. Wang, F. Kanwal, and J. Chhatwal, "Assessment of Incidence of and Surveillance Burden for Hepatocellular Carcinoma Among Patients With Hepatitis C in the Era of Direct-Acting Antiviral Agents," *Journal of the American Medical Association Network Open* 3, no. 11 (2020): e2021173, https://doi.org/10.1001/jamanetworkopen.2020.21173.

34. M. R. Ryu, E. S. Kang, and H. D. Park, "Performance Evaluation of Serum PIVKA-II Measurement Using HISCL-5000 and a Method Comparison of HISCL-5000, LUMIPULSE G1200, and ARCHITECT i2000," *Journal of Clinical Laboratory Analysis* 33, no. 6 (2019): e22921, https://doi.org/10.1002/jcla.22921.

35. G. Ricco, D. Cavallone, C. Cosma, et al., "Impact of Etiology of Chronic Liver Disease on Hepatocellular Carcinoma Biomarkers," *Cancer Biomarkers* 21, no. 3 (2018): 603–612, https://doi.org/10.3233/CBM-170551.

36. G. Ricco, C. Cosma, G. Bedogni, et al., "Modeling the Time-Related Fluctuations of AFP and PIVKA-II Serum Levels in Patients With Cirrhosis Undergoing Surveillance for Hepatocellular Carcinoma," *Cancer Biomarkers* 29, no. 2 (2020): 189–196, https://doi.org/10.3233/CBM-190118.

37. L. Dong, X. Qiu, F. Gao, K. Wang, and X. Xu, "Protein Induced by Vitamin K Absence or Antagonist II: Experience to Date and Future Directions," *Biochimica Et Biophysica Acta. Reviews on Cancer* 1878, no. 6 (2023): 189016, https://doi.org/10.1016/j.bbcan.2023.189016.

38. G. N. Ioannou, "HCC Surveillance After SVR in Patients With F3/F4 Fibrosis," *Journal of Hepatology* 74, no. 2 (2021): 458–465, https://doi.org/10.1016/j.jhep.2020.10.016.

39. P. J. Johnson, S. J. Pirrie, T. F. Cox, et al., "The Detection of Hepatocellular Carcinoma Using a Prospectively Developed and Validated Model Based on Serological Biomarkers," *Cancer Epidemiology, Biomarkers & Prevention* 23, no. 1 (2014): 144–153, https://doi.org/10.1158/1055-9965.EPI-13-0870.

40. G. P. Caviglia, G. Troshina, U. Santaniello, et al., "Long-Term Hepatocellular Carcinoma Development and Predictive Ability of Non-Invasive Scoring Systems in Patients With HCV-Related Cirrhosis Treated With Direct-Acting Antivirals," *Cancers (Basel)* 14, no. 3 (2022): 828, https://doi.org/10.3390/cancers14030828.

41. J. Calderaro, T. P. Seraphin, T. Luedde, and T. G. Simon, "Artificial Intelligence for the Prevention and Clinical Management of Hepatocellular Carcinoma," *Journal of Hepatology* 76, no. 6 (2022): 1348–1361, https://doi.org/10.1016/j.jhep.2022.01.014.

42. C. L. Ramspek, K. J. Jager, F. W. Dekker, C. Zoccali, and M. van Diepen, "External Validation of Prognostic Models: What, Why, How, When and Where?," *Clinical Kidney Journal* 14, no. 1 (2020): 49–58, https://doi.org/10.1093/ckj/sfaa188.

43. J. Futoma, M. Simons, T. Panch, F. Doshi-Velez, and L. A. Celi, "The Myth of Generalisability in Clinical Research and Machine Learning in Health Care," *Lancet Digital Health* 2, no. 9 (2020): e489–e492, https://doi.org/10.1016/S2589-7500(20)30186-2.

44. S. Wu, N. Zeng, F. Sun, et al., "Hepatocellular Carcinoma Prediction Models in Chronic Hepatitis B: A Systematic Review of 14 Models and External Validation," *Clinical Gastroenterology and Hepatology* 19, no. 12 (2021): 2499–2513, https://doi.org/10.1016/j.cgh.2021.02.040.

45. G. L. Wong, V. W. Hui, Q. Tan, et al., "Novel Machine Learning Models Outperform Risk Scores in Predicting Hepatocellular Carcinoma in Patients With Chronic Viral Hepatitis," *JHEP Reports* 4, no. 3 (2022): 100441, https://doi.org/10.1016/j.jhepr.2022.100441.

46. R. Fan, L. Chen, S. Zhao, et al., "Novel, High Accuracy Models for Hepatocellular Carcinoma Prediction Based on Longitudinal Data and Cell-Free DNA Signatures," *Journal of Hepatology* 79, no. 4 (2023): 933–944, https://doi.org/10.1016/j.jhep.2023.05.039.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.