Virology Journal

RESEARCH

Open Access

Machine learning model for HBsAg seroclearance after 48-week pegylated interferon therapy in inactive HBsAg carriers: a retrospective study

Jianxia Dong[1†], Shan Ren[1†], Jing Zhao[1], Pengxuan Wu[1], Haitian Yu[1], Yao Xie[2], Junliang Fu[3], Xiaorong Mao[4], Zhiliang Gao[5], Bingliang Lin[5], Qingfa Ruan[6], Yongfang Jiang[7], Xiulan Xue[8], Yueyong Zhu[9], Haidong Zhao[10], Haifang Cao[11], Xinyue Chen[1*] and Sujun Zheng[1*]

Abstract

Purpose To identify early predictive factors for hepatitis B surface antigen (HBsAg) clearance at week 48 following pegylated interferon (Peg-IFN) therapy in inactive HBsAg carriers (IHC), and to develop an early machine learning-based model to assist clinical decision-making.

Methods This retrospective analysis was based on a multicenter, prospective cohort and included 777 IHC patients who received at least 48 weeks of Peg-IFN therapy. Least Absolute Shrinkage and Selection Operator (LASSO) regression and the Boruta algorithm were applied to select predictive variables. Nine machine learning models—including logistic regression (LR), decision tree (DT), and random forest (RF)—were constructed and evaluated using 10-fold cross-validation. An external validation cohort ($n = 167$) from three medical centers in Beijing was used for model validation. SHapley Additive exPlanations (SHAP) values were used to interpret variable contributions.

Results The overall HBsAg clearance rate at week 48 was 29.9% (232/777). Key predictors included baseline HBsAg level, HBsAg decline > 1 log IU/mL at week 12, and the ratio of alanine aminotransferase (ALT) to HBsAg at week 12. The RF model demonstrated the best performance with an area under the curve (AUC) of 0.829 (95% CI: 0.784–0.874) and specificity of 0.774 in the training set, and an AUC of 0.838 (95% CI: 0.759–0.917) with specificity of 0.968 in the external validation set. SHAP analysis showed that baseline HBsAg had the highest predictive importance.

Conclusions The RF-based model accurately predicts HBsAg clearance in IHC patients undergoing Peg-IFN therapy and offers a promising tool for early identification of candidates for individualized treatment strategies.

Keywords Inactive HBsAg carrier, HBsAg clearance, Machine learning, Predictive model, Random forest

†Jianxia Dong and Shan Ren contributed equally to this work.

*Correspondence:
Xinyue Chen
chenxydoc@ccmu.edu.cn
Sujun Zheng
zhengsujun@ccmu.edu.cn

Full list of author information is available at the end of the article

Background

Chronic hepatitis B virus (HBV) infection remains a major global public health threat. According to data from 2022, approximately 254 million individuals worldwide are living with chronic HBV infection. Viral hepatitis and its complications cause around 3,500 deaths every day, with HBV being the leading cause of these fatalities [1]. Clearance of hepatitis B surface antigen (HBsAg) not only slows down the progression of liver fibrosis but also significantly reduces the risk of hepatocellular carcinoma (HCC), making it a key therapeutic target [2].

Inactive HBsAg carriers (IHC) represent a special subgroup of chronic HBV-infected patients characterized by low HBsAg levels, suppressed viral replication, and minimal liver inflammation. However, long-term outcomes among IHC individuals vary by ethnicity. Asian IHC patients exhibit a higher risk of disease progression than their Western counterparts. Studies have shown that compared to healthy individuals (HBsAg-negative), IHC patients have a 4.6-fold increased risk of HCC and a 2.1-fold increase in liver-related mortality [3, 4]. Therefore, active antiviral treatment and achievement of HBsAg clearance are clinically meaningful in this population.

Pegylated interferon (Peg-IFN) therapy has demonstrated an HBsAg clearance rate of 25%–50% at 48 weeks in IHC patients [5, 6], which is markedly higher than the spontaneous clearance rate of 1%–2% per year. However, even among IHC individuals, treatment response to Peg-IFN is highly heterogeneous. Approximately 50–70% of patients fail to achieve HBsAg clearance, while being exposed to potential adverse effects and economic burden [7].

Previous studies have identified baseline HBsAg levels, early decline in HBsAg, and elevation of alanine aminotransferase (ALT) as important predictors of treatment response to Peg-IFN. However, most existing predictive models rely on single or limited variables, limiting their overall predictive power [8–10]. Machine learning (ML) methods, by integrating multidimensional clinical features and capturing complex non-linear interactions, offer a novel opportunity to build more robust and accurate prediction models. Compared with traditional statistical approaches, ML algorithms can simultaneously account for baseline characteristics and early treatment dynamics, enhancing prediction accuracy [11]. To date, no ML-based model has been reported specifically for predicting 48-week HBsAg clearance in IHC patients treated with Peg-IFN.

In this context, we conducted a multicenter study based on large-scale, prospective real-world cohorts from China. Our aims were: (1) to identify early predictors of HBsAg clearance following Peg-IFN therapy in IHC patients, (2) to construct and validate an early machine learning-based model enabling clinical decision-making, and (3) to develop a user-friendly clinical tool to assist personalized treatment decision-making.

Materials and methods

Study population

This retrospective analysis was based on data derived from two multicenter, prospective cohorts in China. The primary cohort was sourced from the STARHB Project (Starlight Project for Clinical Cure of Hepatitis B in China), a large-scale, real-world, prospective registry. The external validation cohort originated from a multicenter, prospective, randomized controlled trial. Both cohorts followed standardized protocols for baseline assessment, treatment follow-up, and outcome evaluation (see Supplementary Figs. 1–2 for study design and workflow).

Inclusion criteria were based on the 2019 Chinese Guidelines for the Prevention and Treatment of Chronic Hepatitis B [12] and the 2017 EASL Clinical Practice Guidelines [13]:(1) Age 18–60 years; (2) Persistent HBsAg positivity >6 months with a quantitative level < 1000 IU/mL; (3) HBeAg-negative status; (4) HBV DNA consistently < 2000 IU/mL; (5) ALT persistently within the normal range (< 40 U/L).Exclusion criteria were: (1) Incomplete 48-week follow-up (± 2 weeks window allowed); (2) Poor adherence (fewer than three follow-up visits or loss to follow-up); (3) Co-infection with HAV, HCV, HDV, or HIV; (4) Use of immunomodulatory agents during treatment; (5) Severe hematological abnormalities at baseline (neutrophil count < 1.5×10^9/L or platelet count < 90×10^9/L); (6) Autoimmune diseases or uncontrolled thyroid dysfunction.

The external validation cohort followed the same inclusion/exclusion criteria and visit schedule and enrolled patients from three tertiary hospitals in Beijing. Baseline characteristics are provided in Supplementary Table 1.

Treatment protocol

All patients received pegylated interferon alpha-2b (Peg-IFNα–2b)-based therapy. Those with undetectable baseline HBV DNA received Peg-IFNα–2b monotherapy, while patients with detectable HBV DNA were treated with Peg-IFNα–2b in combination with nucleos(t)ide analogues (NAs). According to previous studies [14], the overall 48-week HBsAg clearance rate did not differ significantly between the monotherapy and combination groups. Peg-IFNα–2b was administered subcutaneously at a standard dose of 1.5 μg/kg once weekly, with dose adjustments based on individual response and tolerability. First-line NAs included entecavir (ETV), tenofovir disoproxil fumarate (TDF), tenofovir alafenamide fumarate (TAF) [2, 13], or tenofovir amibufenamide (TMF), in accordance with current guidelines [15].

Follow-up and data collection

Patients underwent assessments at baseline (week 0) and every 12 weeks during treatment (± 7 days window). Data collected included demographics (age, sex, ethnicity, HBV transmission route), laboratory parameters (HBV serology, liver and renal function, complete blood count, HBV DNA, thyroid function and antibodies), and imaging (abdominal ultrasound and liver stiffness measurement, LSM).

The primary outcome was HBsAg clearance at week 48, defined as quantitative HBsAg < 0.05 IU/mL with HBV DNA < 20 IU/mL. Patients were classified into HBsAg clearance (CR) and non-clearance (NR) groups accordingly.

Data preprocessing

A three-tier quality control strategy was applied:

(1) Variables with > 20% missing values were excluded;

(2) Outliers in continuous variables were detected using 1.5× interquartile range (IQR) and treated with Winsorization (1st and 99th percentiles);

(3) Remaining missing values were imputed using random forest imputation (10 iterations, tree depth = 100).

Continuous variables were standardized using Z-score normalization based on training set statistics. Categorical variables were transformed using one-hot encoding.

Feature selection and collinearity control

A two-stage feature selection approach was used. First, multicollinearity was addressed using variance inflation factor (VIF > 5) and correlation coefficient (> 0.75) thresholds. Then, Least Absolute Shrinkage and Selection Operator (LASSO) regression (10-fold cross-validation with 1-SE criterion) and the Boruta algorithm (500 iterations, significance threshold $p < 0.01$) were applied. Features identified by both methods were retained. Bootstrap resampling (1000 iterations) was used to assess feature stability, and those with > 80% selection frequency were included in model construction.

Model development

Nine machine learning algorithms were implemented: logistic regression (LR), decision tree (DT), random forest (RF), gradient boosting (GB), extreme gradient boosting (XGBoost), Light Gradient Boosting Machine (LightGBM/LGB), support vector machine (SVM), multilayer perceptron (MLP), and naïve Bayes (NB). Bayesian optimization was used to fine-tune key hyperparameters, including number of trees (10–200), maximum tree depth (1–20), and minimum samples per leaf (2–20). The dataset was split using stratified random sampling into training (60%) and test (40%) sets (Supplementary Table 2). Model performance was evaluated using multiple metrics: area under the receiver operating characteristic curve (AUC), accuracy, precision, sensitivity, specificity, and F1-score. SHapley Additive exPlanations (SHAP) was used for model interpretation. Optimized hyperparameters for each model are shown in Supplementary Table 3.

Statistical analysis

Continuous variables with normal distribution were compared using the t-test and expressed as mean ± standard deviation. Non-normally distributed variables were compared using the Wilcoxon rank-sum test and presented as median (interquartile range). Categorical variables were compared using chi-square or Fisher's exact tests, as appropriate. All analyses were conducted using Python 3.12 and R 4.4.2, with $p < 0.05$ considered statistically significant.

Results

Baseline characteristics and treatment response

Baseline characteristics

As shown in Fig. 1, a total of 777 eligible IHC patients were included in the final analysis after rigorous screening. Of these, 528 (68.0%) were male, with a median age of 41 years (IQR: 35–49). The majority were of Han ethnicity (96.2%, 747/777). The median baseline HBsAg level was 121.36 IU/mL (IQR: 15.32–389.34.32.34), and the median HBV DNA level was 0 IU/mL (IQR: 0–150), indicating that most patients had undetectable HBV DNA, consistent with the low viral burden typical of IHC patients.After 48 weeks of Peg-IFN ± NA therapy, 232 patients (29.9%) achieved HBsAg clearance. As shown in Table 1, the clearance group had significantly lower baseline HBsAg levels [17.2 IU/mL (1.67–124.04.67.04) vs. 214.52 IU/mL (45.83–496.04.83.04), $P < 0.001$] and direct bilirubin levels [3.4 μmol/L (2.56–4.82) vs. 3.8 μmol/L (2.7–5.4), $P = 0.02$] than the non-clearance group. No significant differences were observed between the two groups in terms of sex, age, ethnicity, BMI, transmission route, or treatment regimen (all $P > 0.05$). Imaging data (liver ultrasound and liver stiffness measurement, LSM) had a high rate of missingness (see Supplementary Table 4). Among patients with available imaging data, there were no significant differences in LSM or controlled attenuation parameter (CAP) values between the CR and NR groups ($P > 0.05$ for both).

Week 12 on-treatment response

At week 12, patients in the HBsAg clearance group demonstrated stronger early treatment responses compared to those in the non-clearance group. Specifically, they had significantly lower serum HBsAg levels [median (IQR): 0.87 IU/mL (0.02–31.75) vs. 103 IU/mL (14.8–326.79.8.79), $P < 0.001$], greater HBsAg declines [0.94 $\log_{10}$ IU/mL (0.28–1.97) vs. 0.14 $\log_{10}$ IU/mL (– 0.07–0.57), $P < 0.001$], and higher aspartate aminotransferase

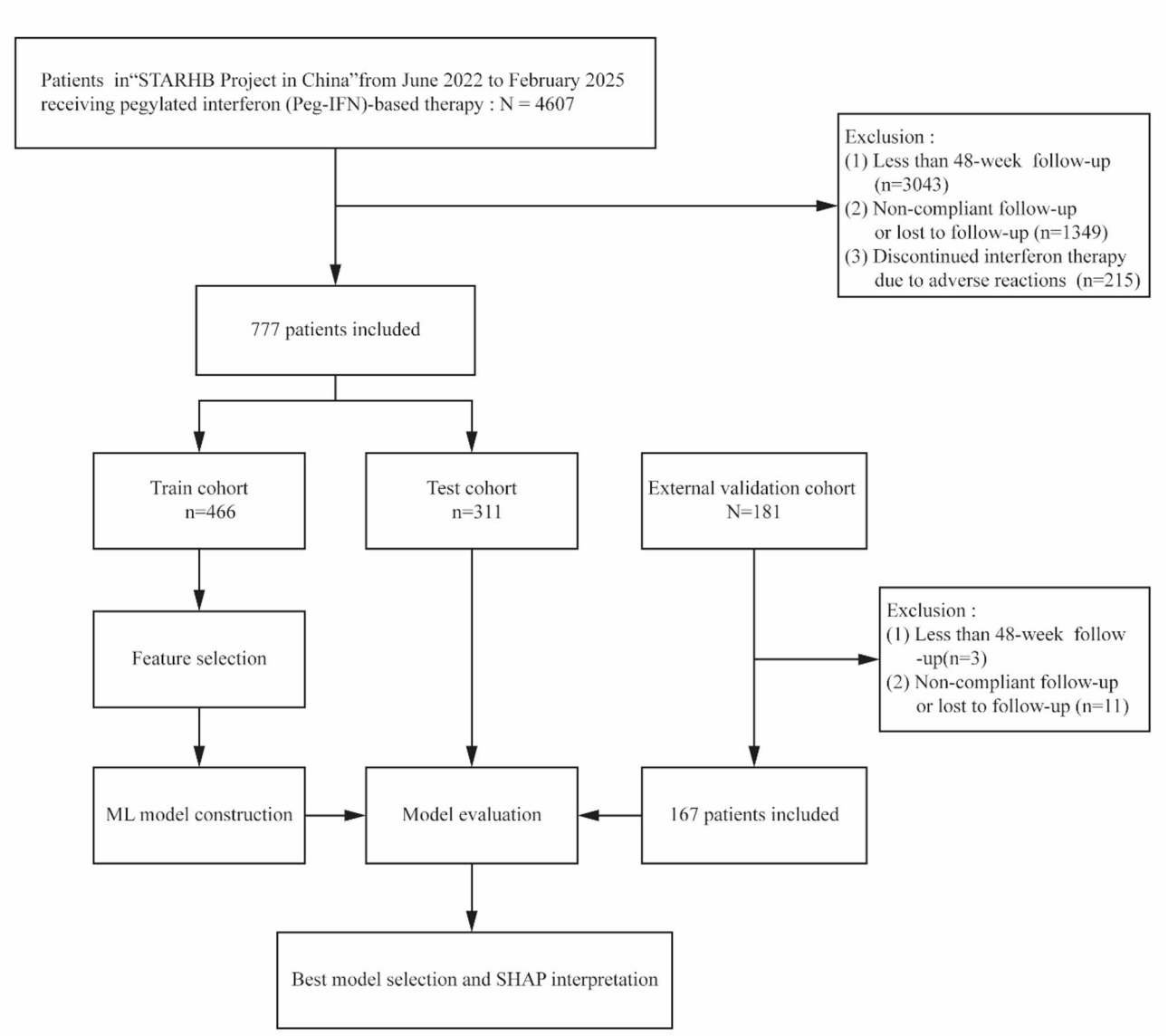

Fig. 1 Flowchart of patient selection and model development

(AST) levels [median (IQR): 54.8 U/L (39.6–81.0) vs. 48.0 U/L (36.0–68.6.0.6), $P = 0.002$].

External validation cohort

The external validation cohort included 167 patients, of whom 42 (25.1%) achieved HBsAg clearance at week 48. As shown in Supplementary Table 1, this cohort was well-matched to the primary cohort in key demographic characteristics, including sex distribution (58.7% male), median age (42 years), and clinical outcomes (HBsAg clearance rate: 25.1% vs. 29.9%, $P = 0.48$), supporting its representativeness for external validation purposes.

Feature selection

To identify key predictors associated with HBsAg clearance at week 48, we employed a dual-step feature selection process using both the Boruta algorithm and LASSO regression.

Boruta algorithm

The Boruta algorithm identified 13 important features ranked by their contribution to model performance (Fig. 2A). The top-ranked variables included: HBsAg level at week 24 (HBsAg 24w), HBsAg level at week 12 (HBsAg 12w), Ratio of alanine aminotransferase to HBsAg at week 12 (ALT12w/HBsAg 12w), Binary indicator for >1 log IU/mL decline in HBsAg at week 24

Table 1 Difference characteristics between HBsAg loss and non-HBsAg loss groups at 48 weeks

Variables	Overall (*n* = 777)	Non-HBsAg loss at 48 weeks (*n* = 545)	HBsAg loss at 48 weeks (*n* = 232)	*P*
Baseline characteristics				
Male, n (%)	528 (68)	376 (69)	152 (66)	0.39
Age, years	41 (35, 49)	42 (35, 49)	41 (34, 47)	0.13
Nation, n (%)				0.07
Han	747 (96)	519 (95)	228 (98)	
Others	30 (4)	26 (5)	4 (2)	
BMI, kg/m^2	23.14 (21.71, 24.22)	23.19 (21.77, 24.22)	22.94 (21.59, 24.22)	0.36
Transmission, n (%)				0.84
Non-vertical	699 (90)	489 (90)	210 (91)	
Vertical	78 (10)	56 (10)	22 (9)	
Treatment, n (%)				0.17
PEG-IFNα−2b + NAs	396 (51)	287 (53)	109 (47)	
PEG IFNα−2b	381 (49)	258 (47)	123 (53)	
HBVDNA, IU/mL	0 (0, 150)	0 (0, 164)	0 (0, 116.25)	0.68
HBsAg, IU/mL	121.36 (15.32, 389.34)	214.52 (45.83, 496.04)	17.2 (1.67, 124.04)	< 0.001
HBsAb, IU/L	0.14 (0, 1.05)	0.1 (0, 0.91)	0.16 (0, 1.36)	0.20
ALT, U/L	21 (16, 26.9)	21.4 (16, 27)	20.5 (15, 26)	0.11
AST, U/L	22 (18.56, 25)	22 (19, 25)	21.05 (18, 25)	0.10
TBIL, μmol/L	13.94 (10.3, 18.8)	14.1 (10.6, 19.24)	13.7 (9.95, 17.85)	0.08
TP, g/L	74.5 (71.6, 77.6)	74.4 (71.7, 77.4)	74.97 (71.4, 77.93)	0.60
ALB, g/L	46.1 (44.2, 47.9)	46 (44.3, 47.7)	46.45 (43.93, 48.4)	0.23
Cr, μmol/L	73.4 (61, 80.92)	73.76 (61, 80.21)	72.76 (60.81, 81.04)	0.83
DBIL, μmol/L	3.7 (2.7, 5.1)	3.8 (2.7, 5.4)	3.4 (2.56, 4.82)	0.02
HB, g/L	151 (140, 161)	152 (140, 161)	149.85 (140, 159.25)	0.13
WBC, ×10^9/L	5.47 (4.57, 6.53)	5.4 (4.5, 6.48)	5.53 (4.73, 6.64)	0.10
ANC, ×10^9/L	3.1 (2.42, 3.89)	3.07 (2.39, 3.86)	3.14 (2.62, 4.09)	0.17
PLT, ×10^9/L	204 (166, 243)	201 (165, 239)	208 (171.5, 249)	0.11
AFP, ng/ml	2.78 (2, 3.54)	2.85 (1.96, 3.64)	2.68 (2.08, 3.34)	0.10
Characteristics at week 12				
HBsAg, IU/mL	46.83 (2.3, 222.18)	103 (14.8, 326.79)	0.87 (0.02, 31.75)	< 0.001
HBsAg decline, log$_{10}$/IU/mL	0.26 (−0.02, 0.99)	0.14 (−0.07, 0.57)	0.94 (0.28, 1.97)	< 0.001
HBsAb, IU/L	0.5 (0, 2.32)	0.5 (0, 2)	0.68 (0, 4.7)	0.006
HBVDNA, IU/mL	0 (0, 14.86)	0 (0, 15.35)	0 (0, 11.59)	0.009
ALT, U/L	55 (40.4, 86)	53.41 (40.4, 81)	58.91 (40.83, 94.25)	0.09
ALT rise, U/L	34.38 (17, 62.32)	33.93 (16.2, 60.6)	35.35 (19.08, 75.5)	0.05
AST, U/L	50 (37.56, 73)	48 (36, 68.6)	54.8 (39.6, 81)	0.002
Characteristics at week 24				
HBsAg, IU/mL	9.15 (0.19, 94.89)	44.7 (3.83, 172.19)	0.01 (0.01, 0.77)	< 0.001
HBsAg decline, log10/IU/mL	0.81 (0.21, 1.96)	0.47 (0.12, 1.24)	2.19 (1.19, 3.22)	< 0.001
HBsAb, IU/L	0 (0, 6.98)	0 (0, 3.22)	4.28 (0.14, 23.37)	< 0.001
ALT, U/L	44.93 (32, 63.08)	44.31 (30.7, 64)	46 (34, 62)	0.33

HBsAg: Hepatitis B Surface Antigen, HBV DNA: Hepatitis B Virus Deoxyribonucleic Acid, ALT: Alanine Aminotransferase, AST: Aspartate Aminotransferase, TBIL: Total Bilirubin, TP: Total Protein, ALB: Albumin, Cr: Creatinine, DBIL: Direct Bilirubin, HB: Hemoglobin, WBC: White Blood Cell Count, ANC: Absolute Neutrophil Count, PLT: Platelet Count, AFP: Alpha-Fetoprotein, HBeAg: Hepatitis B e Antigen, COI: Cutoff Index, HBsAb: Hepatitis B Surface Antibody, PEG-IFNα−2b: Pegylated Interferon Alpha-2b, NAs: Nucleoside/Nucleotide Analogs, BMI: Body Mass Index, ULN: Upper Limit of Normal, ANC: Absolute Neutrophil Count, AFP: Alpha-Fetoprotein. Continuous variables with normal distribution are compared using t-tests, reported as mean ± SD. Non-normal continuous variables are compared using the Wilcoxon rank-sum test, reported as median (IQR). Categorical variables are expressed as percentages and compared using Fisher's exact tests. $P<0.05$ was considered statistically significant

(HBsAg 24w decline > 1log), Ratio of ALT at week 12 to baseline HBsAg (ALT12w/HBsAg baseline), Baseline HBsAg level (HBsAg baseline), Composite variable of HBsAg log decline and ALT fold-increase at week 12 ((Δlog HBsAg × ΔALT)12w), Anti-HBs level at week 24 (HBsAb 24w), Binary indicator for > 1 log IU/mL decline in HBsAg at week 12 (HBsAg 12w decline > 1log), HBV DNA level at week 12 (DNA 12w), HBsAb level at week

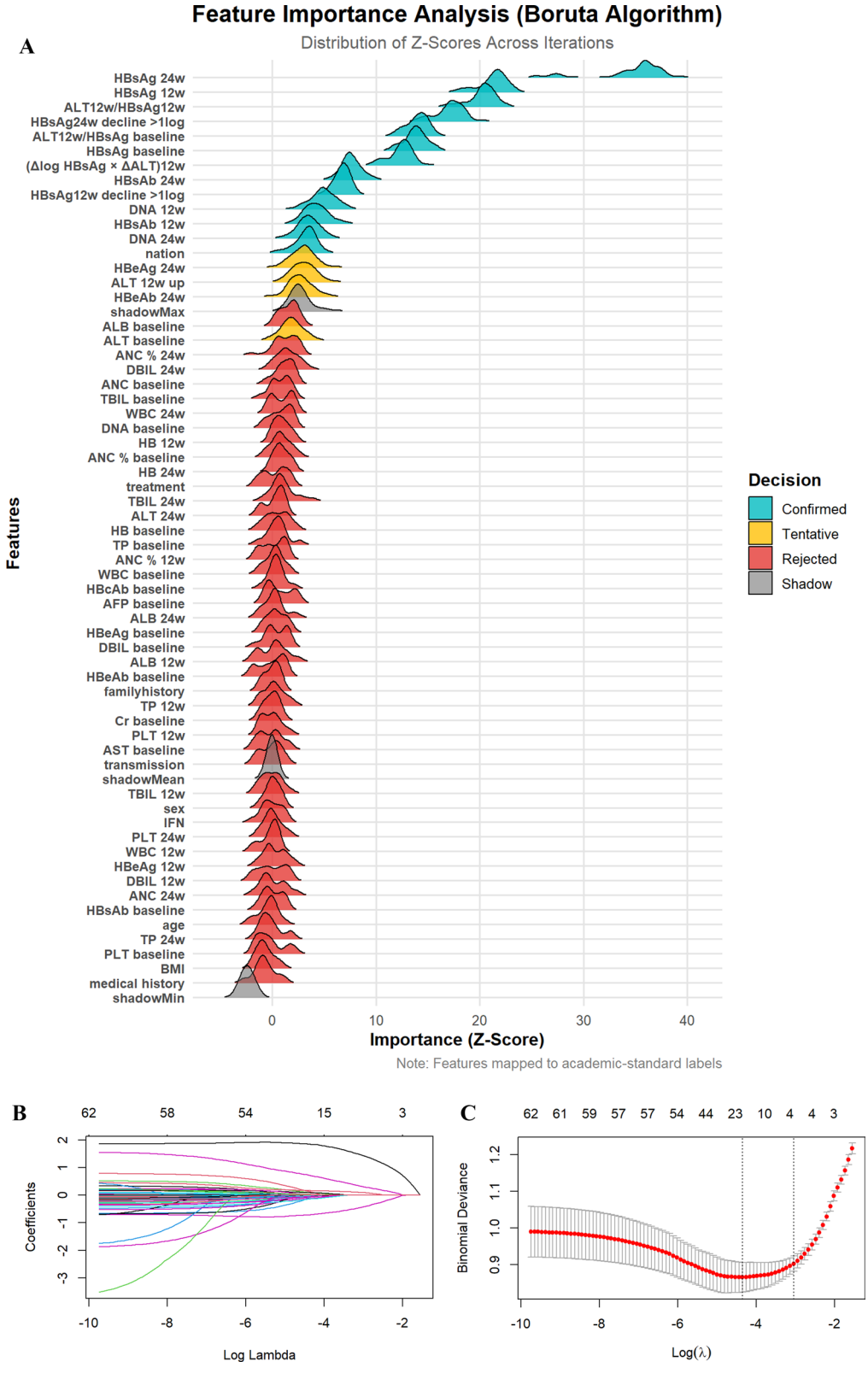

Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Feature selection by two algorithms. (**A**) Boruta algorithm feature selection. Confirmed (green): Key predictors; Rejected (red): Non-significant; Shadow (gray): Synthetic references; Tentative (orange): Uncertain features. (**B**) LASSO regression coefficient profiles of variables in the training dataset. (**C**) Selection of the optimal parameter (lambda) in the LASSO regression. HBsAg 24w: HBsAg level at week 24, HBsAg 12w: HBsAg level at week 12, ALT12w/HBsAg 12w: Ratio of ALT to HBsAg at week 12, HBsAg 24w decline > 1log: Whether HBsAg decline exceeds 1 log IU/mL at week 24, ALT12w/HBsAg baseline: Ratio of ALT at week 12 to baseline HBsAg, HBsAg baseline: Baseline HBsAg level, (Δlog HBsAg × ΔALT)12w: Product of log decline in HBsAg and increase in ALT at week 12, HBsAb 24w: HBsAb level at week 24, HBsAg 12w decline > 1log: Whether HBsAg decline exceeds 1 log IU/mL at week 12, DNA 12w: HBV DNA level at week 12, HBsAb 12w: HBsAb level at week 12, DNA 24w: HBV DNA level at week 24, nation: Nationality of the patient, ALT 12w up: Increase in ALT at week 12, HBeAg 24w: HBeAg level at week 24, HBeAb 24w: HBeAb level at week 24, shadowMax: Maximum shadow feature value (used for algorithm comparison), ALT baseline: Baseline ALT level, ANC % 24w: ANC percentage at week 24, ANC baseline: Baseline ANC, TBIL baseline: Baseline TBIL level, WBC 24w: WBC at week 24, DNA baseline: Baseline HBV DNA level, HB 12w: HB level at week 12, ANC % baseline: Baseline ANC percentage, treatment: Treatment status, ALT 24w: ALT level at week 24, TBIL 24w: TBIL level at week 24, ALT treatment: ALT treatment indicator, HB baseline: Baseline HB level, TP baseline: Baseline TP level, ANC % 12w: ANC percentage at week 12, WBC baseline: Baseline WBC, AFP baseline: Baseline AFP level, HBeAg baseline: Baseline HBeAg level, HBsAb baseline: Baseline HBsAb level, HBcAb baseline: HBcAb level at baseline, HBeAb baseline: Baseline HBeAb level, DBIL baseline: Baseline DBIL level, HB 24w: HB level at week 24, TP 12w: TP level at week 12, Cr baseline: Baseline Cr level, PLT 12w: PLT at week 12, AST baseline: Baseline AST level, transmission: Transmission mode, TBIL 12w: TBIL level at week 12, sex: Sex of the patient, PLT 24w: PLT at week 24, WBC 12w: WBC at week 12, DBIL 12w: DBIL level at week 12, ANC 12w: ANC at week 12, ANC 24w: ANC at week 24, TP 24w: TP level at week 24, ALB baseline: Baseline ALB level, DBIL 24w: DBIL level at week 24, ALB 12w: ALB level at week 12, ALB 24w: ALB level at week 24, age: Age of the patient, PLT baseline: Baseline PLT, BMI: BMI, IFN: IFN treatment status, medical history: Medical history status, familyhistory: Family History of Hepatitis B Infection, shadowMin: Minimum shadow feature value (used for algorithm comparison)

12 (HBsAb 12w), HBV DNA level at week 24 (DNA 24w), Ethnicity.These features were consistently confirmed across 500 iterations of the Boruta process.

LASSO regression

LASSO regression further refined feature selection by shrinking regression coefficients as the penalty parameter (λ) increased. As shown in Figs. 2B-C and 10-fold cross-validation identified the optimal λ value as 0.032. Under this setting, four key variables were retained: HBsAg baseline, HBsAg 12w decline > 1log, ALT12w/HBsAg 12w, HBsAg 24w decline > 1log.

Final predictor selection

By intersecting the Boruta and LASSO results, four common predictive variables were identified and retained for model development: Baseline HBsAg level,1 log decline in HBsAg at week 12 (yes/no), ALT12w/HBsAg 12w ratio,1 log decline in HBsAg at week 24 (yes/no).

Considering the need for early treatment decision-making, we compared model performance using all four variables versus a reduced three-variable model that excluded the week 24 HBsAg decline. Performance was comparable across training, test, and external validation datasets (Supplementary Tables 5–7). Notably, the three-variable random forest model achieved an AUC of 0.829 in the external validation cohort (vs. 0.839 for the four-variable model), with higher accuracy (0.731 vs. 0.683) and specificity (0.712 vs. 0.592). The three-variable model demonstrated non-inferior performance to its four-variable counterpart. Consequently, the three-variable model. was selected as the final prediction tool. This decision was driven by the primary clinical objective of enabling early (Week 12) risk stratification, thereby allowing for timely treatment adjustments without the need to wait for Week 24 data. (Supplementary Table 8)

Assessment of variable independence

To assess multicollinearity, we generated a correlation heatmap (Supplementary Fig. 3) and calculated variance inflation factors (VIFs). All pairwise correlation coefficients were < 0.7, and VIF values were < 3.5 (Supplementary Table 9), well below the conventional threshold of 5. These results support the absence of significant multicollinearity and the appropriateness of the selected predictors for model development.

Machine learning model construction and performance evaluation

Comparison of model performance

We evaluated the predictive performance of nine ML algorithms—including RF, GB, MLP, and others—using the training set ($n = 466$), internal test set ($n = 311$), and an external validation set ($n = 167$). Results are shown in Fig. 3; Tables 2, 3 and 4.

In the training set (Table 2), most models achieved an area under the curve (AUC) above 0.79. DT model yielded the highest AUC (0.866) and sensitivity (0.863), but its precision was relatively low (0.550), suggesting potential overfitting. In contrast, the RF model demonstrated the most balanced performance, with an AUC of 0.829, accuracy of 0.760, precision of 0.577, specificity of 0.774, and an F1 score of 0.643. Its performance exceeded that of traditional models such as LR and SVM. The GB and XGBoost models also performed well, with AUCs of 0.837 and 0.835, respectively—comparable to RF.

In the internal test set (Table 4), the RF model continued to outperform others, achieving an AUC of 0.827, accuracy of 0.794, precision of 0.659, specificity of 0.858, and an F1 score of 0.652. While GB and XGBoost also showed good AUCs (0.820 and 0.821, respectively), both had sensitivities below 0.70. Notably, SVM demonstrated high specificity (0.881) but low sensitivity (0.602), indicating a more conservative prediction tendency.

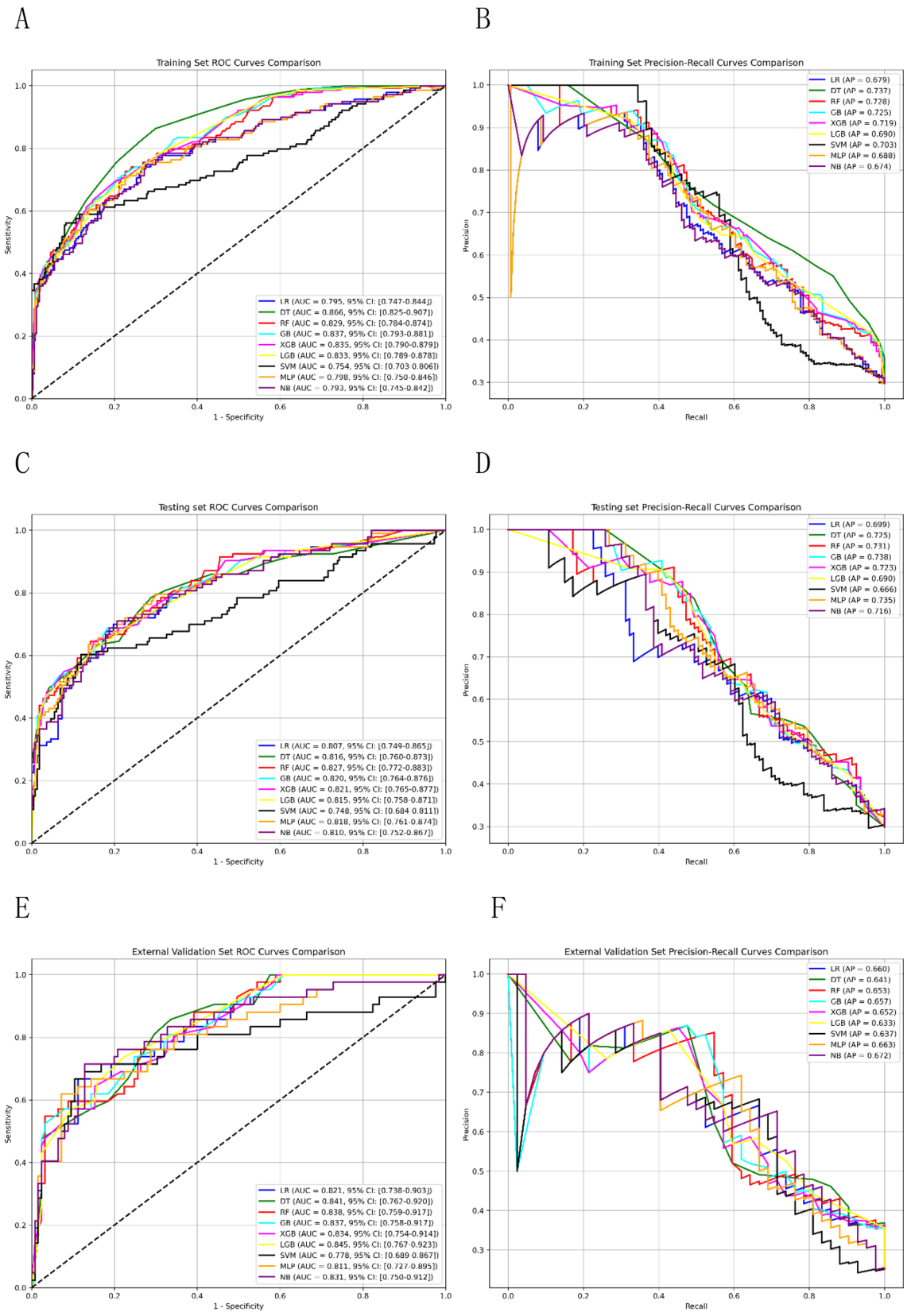

Fig. 3 Model performance evaluation: **A**-**B**. Training Set: ROC Curves (Left) and Precision-Recall Curves (Right). **C**-**D**. Internal Testing Set: ROC Curves (Left) and Precision-Recall Curves (Right). **E**-**F**. External Validation Set: ROC Curves (Left) and Precision-Recall Curves (Right)

Table 2 Performance comparison of different machine learning models in the training set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
LR	0.795	0.736	0.542	0.741	0.734	0.626
DT	0.866	0.749	0.550	0.863	0.700	0.672
RF	0.829	0.760	0.577	0.727	0.774	0.643
GB	0.837	0.766	0.590	0.705	0.792	0.643
XGB	0.835	0.766	0.591	0.698	0.795	0.640
LGB	0.833	0.723	0.524	0.777	0.700	0.626
SVM	0.754	0.811	0.743	0.561	0.917	0.639
MLP	0.798	0.751	0.566	0.712	0.768	0.631
NB	0.793	0.732	0.536	0.748	0.725	0.625

LR: Linear Regression, DT·Decision Tree, RF·Random Forest, GB: Gradient Boosting, XGB·eXtreme Gradient Boosting, LGB·Light Gradient Boosting Machine, SVM·Support Vector Machine, MLP·Multilayer Perceptron, NB: Naive Bayes

Table 3 Performance comparison of different machine learning models in the external validation set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
LR	0.821	0.832	0.667	0.667	0.888	0.667
DT	0.841	0.713	0.462	0.857	0.664	0.600
RF	0.838	0.862	0.852	0.548	0.968	0.667
GB	0.837	0.856	0.846	0.524	0.968	0.647
XGB	0.834	0.796	0.587	0.643	0.848	0.614
LGB	0.845	0.766	0.525	0.738	0.776	0.614
SVM	0.778	0.838	0.683	0.667	0.896	0.675
MLP	0.811	0.850	0.743	0.619	0.928	0.675
NB	0.831	0.832	0.652	0.714	0.872	0.682

LR: Linear Regression, DT·Decision Tree, RF·Random Forest, GB: Gradient Boosting, XGB·eXtreme Gradient Boosting, LGB·Light Gradient Boosting Machine, SVM·Support Vector Machine, MLP·Multilayer Perceptron, NB: Naive Bayes

Table 4 Performance comparison of different machine learning models in the testing set

Model	AUC	Accuracy	Precision	Sensitivity	Specificity	F1 score
LR	0.807	0.772	0.602	0.699	0.803	0.647
DT	0.816	0.733	0.537	0.785	0.711	0.638
RF	0.827	0.794	0.659	0.645	0.858	0.652
GB	0.820	0.778	0.615	0.688	0.817	0.650
XGB	0.821	0.788	0.645	0.645	0.849	0.645
LGB	0.815	0.788	0.655	0.613	0.862	0.633
SVM	0.748	0.797	0.683	0.602	0.881	0.640
MLP	0.818	0.730	0.532	0.796	0.702	0.638
NB	0.810	0.768	0.595	0.710	0.794	0.647

LR: Linear Regression, DT·Decision Tree, RF·Random Forest, GB: Gradient Boosting, XGB·eXtreme Gradient Boosting, LGB·Light Gradient Boosting Machine, SVM·Support Vector Machine, MLP·Multilayer Perceptron, NB: Naive Bayes

In the external validation set (Table 4), the RF model again exhibited superior performance, with the highest AUC of 0.838 and the highest overall accuracy (0.862). It also achieved a precision of 0.852 and an exceptionally high specificity of 0.968, with an F1 score of 0.667. In comparison, the XGBoost model showed a precision of only 0.587, and GB had a sensitivity of 0.524—both inferior to RF. Although the DT model had the highest sensitivity (0.857), its precision was low (0.462) and F1 score lowest (0.600), indicating a high rate of false positives.

As shown in Fig. 3, both the Receiver Operating Characteristic (ROC) curves and the precision-recall curves further confirmed the discrimination ability of each model across datasets. Taken together, the RF model demonstrated consistent and excellent performance across all datasets, balancing discrimination and generalizability.

Final model selection

Considering the performance across the training, test, and external validation datasets, the RF model was ultimately selected as the predictive model for this study. The advantages of the RF model include:

①Stable performance across all three datasets without evidence of overfitting;② High specificity (0.968 in the external validation set), aligning with the clinical principle of prioritizing the minimization of false positives to reduce unnecessary treatments;③ Strong interpretability, facilitating clinical understanding and application;④ Calibration curves demonstrating good agreement between predicted probabilities and observed outcomes (Brier score = 0.13), with decision curve analysis indicating high net benefit across a probability threshold range of 0.2–0.8 (Supplementary Fig. 4).

Model interpretation

To elucidate the predictive mechanisms of the RF model, we employed the SHAP framework for feature interpretability analysis (Fig. 4).

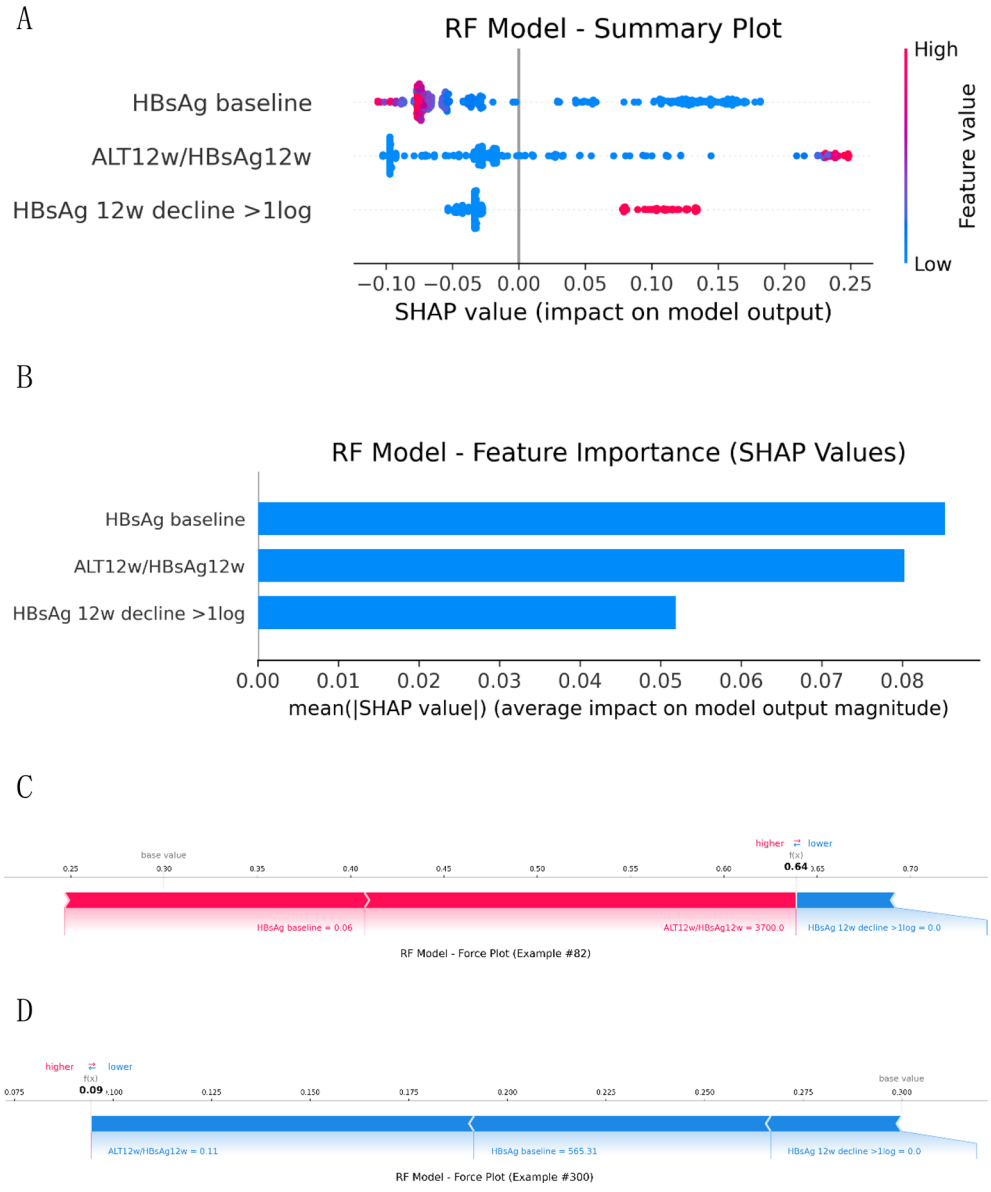

Fig. 4 SHAP interpretation of the random forest model: (**A**). SHAP summary plot shows the distribution of the impact of each feature on the model output for every individual patient in the dataset (**B**). Mean SHAP value bar plot displays the global importance of each feature, calculated as the mean absolute SHAP value across all patients (**C**-**D**). SHAP force plots for individual predictions: individual contributions from patients with HBsAg loss(**C**) and (**D**) without HBsAg loss. The baseline (model's average output) serves as a reference. Positive contributions (red) raise the predicted HBsAg clearance probability, while negative ones (blue) reduce it

Global feature importance

As shown in Figs. 4A-B, baseline HBsAg level (SHAP value = 0.085) emerged as the most critical predictor, exhibiting the highest mean absolute SHAP value and contributing most significantly to model predictions. This was followed by ALT12w/HBsAg12w (SHAP value = 0.080) and HBsAg12w decline > 1log (SHAP value = 0.052).

SHAP risk direction analysis revealed that lower baseline HBsAg levels (blue point clusters) and a greater HBsAg decline > 1 log IU/mL at week 12 (red points skewed right) were strongly associated with an increased probability of HBsAg clearance, consistent with established clinical knowledge.

Individualized prediction interpretation

To further elucidate decision-making logic of the model, local explanation plots were generated for two representative samples (Figs. 4C-D).

For sample 82 (Fig. 4C), a low baseline HBsAg level (0.06 IU/mL) and a high ALT12w/HBsAg12w ratio provided positive contributions to the prediction. Despite the absence of an HBsAg decline > 1 log IU/mL at week 12 (HBsAg12w decline > 1log = 0), the model predicted a clearance probability of 0.64.

In contrast, for sample 300 (Fig. 4D), a high baseline HBsAg level (565.31 IU/mL), a low ALT12w/HBsAg12w ratio (0.11), and an HBsAg decline < 1 log IU/mL at week 12 (HBsAg12w decline > 1log = 0) all contributed negatively, resulting in a predicted clearance probability of only 0.09. These findings indicate that the model's decision-making aligns closely with viral dynamics, offering robust clinical interpretability.

Clinical translation and decision support tools

To bridge the predictive model with clinical practice, we developed and validated multiple decision-support tools. Initially, patients were stratified into low-, medium-, and high-risk groups based on the RF model's predicted probabilities, which revealed significantly different HBsAg clearance rates across all datasets ($P < 0.001$, Supplementary Fig. 5). We subsequently created an online calculator to operationalize the model, providing individualized clearance probabilities (0–100%) from patient-specific inputs (baseline HBsAg, week 12 HBsAg, and week 12 ALT). This tool is accessible at: https://ihcmodel-atvwnnehryrkimrl8wyhy4.streamlit.app/ (Fig. 5).

To offer a more accessible, model-free alternative, we also derived a futility rule and a simplified clinical scoring system. A low-probability subgroup, defined by the 25th percentile of the training set predictions (probability < 0.26), exhibited a very low HBsAg clearance rate of 6.5% (9/138) in the external validation cohort. Furthermore, we constructed a binary scoring system based on

Fig. 5 Web-based calculator interface for clinical application

Table 5 Performance of the clinical futility scoring system in the external validation cohort ($n = 167$)

Probability category	Total score	Patients, *n* (%)	HBsAg clearance, *n* (%)	Clearance rate	NPV
Low	< 2	58 (34.7%)	4 (6.9%)	6.9%	93.1%
Intermediate/High	≥ 2	109 (65.3%)	38 (34.9%)	34.9%	-

two key week-12 predictors. Patients were assigned 2 points for an ALT12w/HBsAg12w ratio ≥ 30(see Supplementary Fig. 6), and an additional 2 points for an HBsAg decline ≥ 1log$_{10}$ IU/mL at week 12, resulting in a total score ranging from 0 to 4. Patients with a total score < 2 were classified as "Low Probability" responders. As shown in Table 5, this rule effectively identified a patient subgroup with a minimal chance of success (clearance rate:6.9%), yielding a high negative predictive value of 93.1%. This provides a pragmatic tool for early futility assessment.

Discussion

This study, based on a large-scale, multicenter cohort of IHC patients, developed and validated a ML model to predict HBsAg clearance at 48 weeks following Peg-IFN therapy. By systematically comparing nine mainstream ML algorithms, including RF, GB, and XGBoost, we identified RF as the optimal predictive model. The model demonstrated robust generalizability in the external validation cohort (AUC = 0.838) and exceptional specificity (0.968), providing a reliable tool for personalized treatment decision-making in clinical practice.

Three key predictors were identified: baseline HBsAg level, HBsAg decline > 1 log IU/mL at week 12, and the ALT12w/HBsAg12w ratio. The first two predictors have been well-validated in prior studies. A meta-analysis by Zhang et al. [8]encompassing 102 studies confirmed baseline HBsAg level as the most significant predictor of HBsAg clearance, while Chen et al. [9] highlighted the predictive accuracy of early HBsAg decline.

This study is the first to propose the ALT12w/HBsAg12w ratio as a novel dynamic predictor, reflecting the interplay between hepatic inflammation and viral suppression. SHAP analysis underscored its substantial contribution to the model, suggesting a critical role in HBsAg clearance. Previous studies by Wu et al. [10], Yang et al. [16], and our team [14] have supported the positive association between early ALT elevation and HBsAg clearance.

From an immunological perspective, low HBsAg levels facilitate the restoration of T-cell function, enhancing antiviral immune responses [17] Single-cell analysis by Narmada et al. [18] revealed an adaptive immune activation pattern centered on CD4 + cytotoxic T lymphocytes in functionally cured individuals. Elevated ALT likely reflects interferon-induced immune-mediated hepatocellular injury, indicating active host immune clearance of infected hepatocytes [19, 20].Thus, baseline HBsAg, week 12 HBsAg decline >1 log IU/mL, and the ALT12w/HBsAg12w ratio collectively represent a biological framework for HBsAg clearance, capturing initial viral load, treatment response, and immune activation.

Compared to traditional statistical methods, ML algorithms, particularly ensemble models like RF, excel in handling high-dimensional variables, non-linear relationships, and complex interactions [11, 21, 22]. By combining LASSO regression and the Boruta algorithm, we screened 61 candidate variables to eliminate redundancy and noise, enhancing model consistency and generalizability across datasets.

SHAP analysis quantified each feature's contribution to individual predictions, enabling transparent interpretation of the "black-box" RF model. This interpretability enhances clinical acceptability, allowing physicians to understand the rationale behind predictions and integrate them into decision-making. The web-based prediction tool further facilitates real-time individualized predictions, enabling clinicians to identify patients with high clearance probabilities, optimize treatment strategies, and improve therapeutic efficiency.

The dynamic pattern of HBsAg decline observed in this study supports the biological plausibility of our predictive variables. Patients who achieved HBsAg clearance had lower baseline HBsAg levels and showed a rapid logarithmic decline within the first 12–24 weeks. These temporal trends (Supplementary Table 10) indicate that clearers maintained consistently lower HBsAg levels and greater declines than non-clearers, validating the dynamic predictors included in our model. Although the 24-week HBsAg decline also showed predictive value (AUC increased from 0.827 to 0.866 after its inclusion), the improvement was not statistically significant (DeLong's test, $P = 0.60$). Notably, current chronic hepatitis B guidelines and recent expert consensus on functional clinical cure emphasize the prognostic importance of week-24 markers [13, 15, 23]; therefore, we present the week-24 data in the Results and Supplementary Tables for transparency and comparison. For clinical practicality and early treatment guidance, we prioritized week-12 parameters to construct an early prediction model, allowing timely identification of non-responders without waiting for additional follow-up data.

The clinical utility of this model is demonstrated by its translation into multiple actionable tools for mid-treatment decision-making at week 12. First, we developed an online calculator to provide individualized clearance probabilities, proposing thresholds of ≥ 0.7 (high-probability) and < 0.3(low-probability) to guide therapy. For high-probability patients, extending Peg-IFN may be beneficial, whereas for low-probability patients, combining NAs or exploring alternatives could avoid ineffective treatment.To further enhance accessibility, we derived a simplified clinical scoring system from the model. This tool, requiring only the week 12 ALT/HBsAg ratio and HBsAg decline, successfully identified a substantial patient subgroup (34.7%) with a minimal chance of clearance (6.9%). The high negative predictive value (93.1%) empowers clinicians to confidently consider discontinuing Peg-IFN in these "low-probability responders" as early as week 12. This approach directly addresses the clinical dilemma of managing non-responders. By stopping ineffective therapy early, we can spare patients from unnecessary side effects and reduce costs. Together, these tools bridge the gap between predictive modeling and bedside practice, paving the way for more personalized and efficient management strategies.

This study has several limitations. First, as a retrospective analysis, despite strict inclusion criteria, selection bias cannot be entirely excluded. The lack of significant difference in HBsAg clearance rates between monotherapy and combination therapy groups (P = 0.17), coupled with low median HBV DNA levels (0 IU/mL, often below the detection limit), suggests a limited impact of viral load on treatment outcomes. However, potential biases warrant further validation in larger populations. Second, the study primarily included Han Chinese patients, with external validation limited to three centers in Northern China, necessitating broader validation across diverse ethnicities and regions. Third, HBV genotype was not included in the model, despite evidence suggesting that genotype B patients may achieve HBsAg clearance more readily than genotype C patients [24]. Fourth, emerging biomarkers such as hepatitis B core-related antigen (HBcrAg), HBV RNA, and host genetic polymorphisms were not analyzed, but their inclusion in future studies could enhance predictive power. Finally, the absence of host immune profile data limited our ability to explore the heterogeneity of individual immune responses.

Among the IHC patients in this study, the HBsAg clearance rate was 17.5% at 24 weeks and 29.8% at 48 weeks. This suggests that while both week 12 and week 24 dynamic predictors can forecast 48-week clearance, their predictive performance is comparable, likely due to the modest difference in cumulative clearance rates between these time points. This observation underscores the need for predictive models targeting rapid HBsAg clearance (within 24 weeks), which may indicate stronger immune activation and viral suppression. Identifying such patients early could optimize Peg-IFN treatment duration and strategies. Future studies by our team will focus on identifying clinical and immunological factors associated with rapid HBsAg clearance within 24 weeks, further refining the proposed predictive framework.

Future research should integrate HBV genotype, HBcrAg, HBV RNA, and human leukocyte antigen (HLA) profiles to develop multimodal predictive models. Prospective randomized controlled trials are needed to validate the clinical utility of model-guided treatment decisions. Ultimately, these efforts aim to advance precision and individualized treatment strategies for IHC patients, contributing to the World Health Organization's goal of eliminating viral hepatitis by 2030.

Conclusion

This study developed and externally validated a machine learning model to predict HBsAg clearance in inactive HBsAg carriers undergoing Peg-IFN therapy. Baseline HBsAg level, week 12 HBsAg decline, and the week 12 ALT/HBsAg ratio were identified as the most relevant predictors. The Random Forest model achieved stable performance and good interpretability. The proposed online calculator may assist clinicians in early response evaluation and individualized treatment planning; however, its clinical utility requires further validation in prospective, multiethnic cohorts before widespread application.

Abbreviations

HBV	Hepatitis B virus
HBsAg	Hepatitis B surface antigen
HCC	Hepatocellular carcinoma
IHC	Inactive HBsAg carrier
Peg-IFN	Pegylated interferon
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ML	Machine learning
Peg-IFNα-2b	Pegylated interferon alpha-2b
NAs	Nucleos(t)ide analogues
ETV	Entecavir
TDF	Tenofovir disoproxil fumarate
TAF	Tenofovir alafenamide fumarate
TMF	Tenofovir amibufenamide
LSM	Liver stiffness measurement
CR	HBsAg clearance responder
NR	Non-responder
LASSO	Least absolute shrinkage and selection operator
VIF	Variance inflation factor
RF	Random forest
GB	Gradient boosting
XGBoost	Extreme gradient boosting
LightGBM / LGB	Light gradient boosting machine
SVM	Support vector machine
MLP	Multilayer perceptron
NB	Naïve bayes
SHAP	SHapley additive explanations
AUC	Area under the receiver operating characteristic Curve
SD	Standard deviation
IQR	Interquartile range
BMI	Body mass index
HBeAg	Hepatitis B e antigen
HBsAb	Hepatitis B surface antibody
HBcrAg	Hepatitis B core-related antigen
HLA	Human leukocyte antigen
ROC	Receiver operating characteristic
CAP	Controlled attenuation parameter
TP	Total protein
ALB	Albumin
Cr	Creatinine
DBIL	Direct bilirubin
TBIL	Total bilirubin
HB	Hemoglobin
WBC	White blood cell count
ANC	Absolute neutrophil count
PLT	Platelet count
AFP	Alpha-fetoprotein
COI	Cutoff index
ULN	Upper limit of normal
LR	Logistic regression
DT	Decision tree
DNA	HBV DNA (Hepatitis B Virus Deoxyribonucleic Acid)
ALT12w/HBsAg12w	Ratio of ALT to HBsAg at week 12
HBsAg 12w decline >1log	HBsAg decline >1 log IU/mL at week 12 (binary)
HBsAg 24w decline >1log	HBsAg decline >1 log IU/mL at week 24 (binary)
ALT12w/HBsAg baseline	Ratio of ALT at week 12 to baseline HBsAg
(Δlog HBsAg × ΔALT)12w	Product of log HBsAg decline and ALT fold change at week 12
DNA 12w / 24w / baseline	HBV DNA levels at corresponding time points
HBsAb 12w / 24w / baseline	Anti-HBs levels at corresponding time points
HBeAb / HBcAb	Hepatitis B e/core Antibody
PLT 12w / 24w / baseline	Platelet count at corresponding time points
TP 12w / 24w / baseline	Total protein at corresponding time points
familyhistory	Family History of Hepatitis B Infection

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12985-025-03012-1.

Supplementary Material 1

Acknowledgements

We acknowledge the contributions of all participating centers in the STARHB Project and the multicenter randomized controlled trial for their support in patient recruitment, data collection, and follow-up. We thank all study participants for their cooperation. This study was supported by National Key Research and Development Program of Ministry of Science and Technology (2023YFC2308100): Dr. Sujun Zheng; Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (Academic Leader-02-14): Dr. Sujun Zheng; Natural Science Foundation of Beijing Municipality (7222093): Dr. Sujun Zheng; Capital Clinical Diagnostic Techniques and Translational Application Projects (Z211100002921059): Prof. Xinyue Chen, provided financial support for study design, data collection, analysis, interpretation, manuscript writing, and the decision to submit the manuscript for publication. No individuals who do not meet authorship criteria contributed to this study.

Author contributions

Jianxia Dong and Shan Ren contributed equally to this work. Jianxia Dong designed the study, analyzed the data, and drafted the manuscript. Shan Ren contributed to data analysis and manuscript writing. Jing Zhao, Pengxuan Wu and Haitian Yu revised the manuscript. Yao Xie, Junliang Fu, Xiaorong Mao, Zhiliang Gao, Bingliang Lin, Qingfa Ruan, Yongfang Jiang, Xiulan Xue, Yueyong Zhu, Haidong Zhao, and Haifang Cao were responsible for patient recruitment and data provision. Xinyue Chen and Sujun Zheng supervised the study, provided critical revisions, and obtained funding. All authors read and approved the final manuscript.

Funding

This study was supported by National Key Research and Development Program of Ministry of Science and Technology (2023YFC2308100): Dr. Sujun Zheng; Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan (Academic Leader-02-14): Dr. Sujun Zheng; Natural Science Foundation of Beijing Municipality (7222093): Dr. Sujun Zheng; Capital Clinical Diagnostic Techniques and Translational Application Projects (Z211100002921059): Prof. Xinyue Chen.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all participants prior to data collection. As this study involved a retrospective analysis of prospectively collected clinical data, the use of patient information was reviewed and

approved by the institutional review boards of the participating centers in accordance with the ethical standards of the Declaration of Helsinki. Ethical approval was obtained from the institutional review boards of all participating centers (Primary cohort: 2022-067-K; External cohort: 2022-047-K).

Consent for publication
We give our consent for information about our relative (circle as appropriate) to be published in Virology Journal.

Competing interests
The authors declare no competing interests.

Author details
[1]The First Unit, Department of Hepatology, Beijing Youan Hospital, Capital Medical University, 8 Xitoutiao, Youan men wai, Fengtai District, Beijing 100069, China
[2]Beijing Ditan Hospital, Capital Medical University, Beijing, China
[3]Senior Department of Infectious Diseases, the Fifth Medical Center of PLA General Hospital, National Clinical Research Center for Infectious Diseases, Beijing, China
[4]The First Hospital of Lanzhou University, Lanzhou, China
[5]The Third Affiliated Hospital of Sun Yat-sen university, Guangzhou, China
[6]Xiamen Hospital of Traditional Chinese Medicine, Xiamen, China
[7]The Second Xiangya Hospital of Central South University, Changsha, China
[8]The First Affiliated Hospital of Xiamen University, Xiamen, China
[9]The First Affiliated Hospital of Fujian Medical University, Fuzhou, China
[10]Xiamen Chang Gung Hospital, Xiamen, China
[11]Qinghai Provincial Fourth People's Hospital, Xining, China

Received: 3 August 2025 / Accepted: 27 October 2025
Published online: 23 November 2025

References

1. Organization WH. WHO sounds alarm on viral hepatitis infections claiming 3500 lives each day. 2024. https://www.who.int/news/item/09-04-2024-who-sounds-alarm-on-viral-hepatitis-infections-claiming-3500-lives-each-day
2. Terrault NA, Lok ASF, McMahon BJ, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. Hepatology Apr. 2018;67(4):1560–99. https://doi.org/10.1002/hep.29800.
3. Taida T, Arai M, Kanda T, et al. The prognosis of hepatitis B inactive carriers in Japan: a multicenter prospective study. J Gastroenterol. 2017;52(1):113–22. https://doi.org/10.1007/s00535-016-1229-6.
4. Diallo I, Ndiaye B, Fall CA, et al. Inactive hepatitis B carriers: outcomes of patients followed at Hôpital principal de Dakar, Senegal. Pan Afr Med J. 2018;31:49. https://doi.org/10.11604/pamj.2018.31.49.16296.
5. Song A, Lin X, Lu J, et al. Pegylated interferon treatment for the effective clearance of hepatitis B surface antigen in inactive HBsAg carriers: a meta-analysis. Front Immunol. 2021;12:779347. https://doi.org/10.3389/fimmu.2021.779347.
6. Wu F, Lu R, Liu Y, et al. Efficacy and safety of peginterferon alpha monotherapy in Chinese inactive chronic hepatitis B virus carriers. Liver Int. 2021;41(9):2032–45. https://doi.org/10.1111/liv.14897.
7. Huang Y, Qi M, Liao C, et al. Analysis of the efficacy and safety of pegylated interferon-α2b treatment in inactive hepatitis B surface antigen carriers. Infect Dis Ther. 2021;10(4):2323–31. https://doi.org/10.1007/s40121-021-00511-w.
8. Zhang PX, Tang QQ, Zhu J, Deng WY, Zhang ZH. Predictive models for functional cure in patients with CHB receiving PEG-IFN therapy based on HBsAg quantification through meta-analysis. Hepatol Int. 2024;18(4):1110–21. https://doi.org/10.1007/s12072-024-10666-6.
9. Chen Q, Zhang J, Huang J, et al. Early serum HBsAg drop is a strong predictor of hbeag seroconversion and HBsAg loss to pegylated interferon Alfa-2a in chronic hepatitis B patients with prior Nucleos(t)ide analogue exposure. Med Sci Monit Jun. 2019;23:25:4665–74. https://doi.org/10.12659/msm.916441.
10. Wu FP, Yang Y, Li M. Add-on pegylated interferon augments hepatitis B surface antigen clearance vs continuous nucleos(t)ide analog monotherapy in Chinese patients with chronic hepatitis B and hepatitis B surface antigen ≤ 1500 IU/mL: an observational study. World J Gastroenterol. 2020;7(13):1525–39. https://doi.org/10.3748/wjg.v26.i13.1525.
11. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. N Engl J Med. 2023;30(13):1201–8. https://doi.org/10.1056/NEJMra2302038.
12. [The guidelines of prevention and treatment for chronic hepatitis B. (2019 version)]. *Zhonghua Gan Zang Bing Za Zhi*. Dec 20 2019;27(12):938–961. https://doi.org/10.3760/cma.j.issn.1007-3418.2019.12.007
13. EASL. 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J Hepatol*. Aug 2017;67(2):370–398. https://doi.org/10.1016/j.jhep.2017.03.021
14. Cao Z, Liu Y, Ma L, et al. A potent hepatitis B surface antigen response in subjects with inactive hepatitis B surface antigen carrier treated with pegylated-interferon alpha. Hepatology. 2017;66(4):1058–66. https://doi.org/10.1002/hep.29213.
15. You H, Wang F, Li T, et al. Guidelines for the prevention and treatment of chronic hepatitis B (version 2022). J Clin Transl Hepatol Nov. 2023;28(6):1425–42. https://doi.org/10.14218/jcth.2023.00320.
16. Yang X, Zhang K, Xu Q, et al. Interferon add-on therapy increased clinical cure significantly for interferon-experienced chronic hepatitis B patients with low HBsAg. Front Immunol. 2022;13:997608. https://doi.org/10.3389/fimmu.2022.997608.
17. Tout I, Loureiro D, Mansouri A, Soumelis V, Boyer N, Asselah T. Hepatitis B surface antigen seroclearance: immune mechanisms, clinical impact, importance for drug development. J Hepatol. 2020;73(2):409–22. https://doi.org/10.1016/j.jhep.2020.04.013.
18. Narmada BC, Khakpoor A, Shirgaonkar N, et al. Single-cell landscape of functionally cured chronic hepatitis B patients reveals activation of innate and altered CD4-CTL-driven adaptive immunity. J Hepatol. 2024;81(1):42–61. https://doi.org/10.1016/j.jhep.2024.02.017.
19. Wu L, Deng H, Feng X, et al. Interferon-γ(+) Th1 activates intrahepatic resident memory T cells to promote HBsAg loss by inducing M1 macrophage polarization. J Med Virol. 2024;96(5):e29627. https://doi.org/10.1002/jmv.29627.
20. Wang H, Luo H, Wan X, et al. TNF-α/IFN-γ profile of HBV-specific CD4 T cells is associated with liver damage and viral clearance in chronic HBV infection. J Hepatol. 2020;72(1):45–56. https://doi.org/10.1016/j.jhep.2019.08.024.
21. Deo RC. Machine learning in medicine. Circulation. 2015;17(20):1920–30. https://doi.org/10.1161/circulationaha.115.001593.
22. Deo RC. Machine learning in medicine. Circulation. 2020;20(16):1521–3. https://doi.org/10.1161/circulationaha.120.050583.
23. Wu D, Kao JH, Piratvisuth T, et al. Update on the treatment navigation for functional cure of chronic hepatitis B: expert consensus 2.0. Clin Mol Hepatol Feb. 2025;31(Suppl):S134–64. https://doi.org/10.3350/cmh.2024.0780.
24. Jeng WJ, Papatheodoridis GV, Lok ASF. Hepatitis B. Lancet. 2023(10381). https://doi.org/10.1016/s0140-6736(22)01468-4.

Publisher's note