Development of a machine learning-based model to predict hepatic inflammation in chronic hepatitis B patients with concurrent hepatic steatosis: a cohort study

Fajuan Rui,[a,b,c,y] Yee Hui Yeo,[d,y] Liang Xu,[e,f,g,y] Qi Zheng,[h] Xiaoming Xu,[a,b,c] Wenjing Ni,[a,b,c] Youwen Tan,[i] Qing-Lei Zeng,[j] Zebao He,[k] Xiaorong Tian,[l,m] Qi Xue,[n] Yuanwang Qiu,[o] Chuanwu Zhu,[p] Weimao Ding,[q] Jian Wang,[b,c] Rui Huang,[b,c] Yayun Xu,[r] Yunliang Chen,[l,m] Junqing Fan,[l,m] Zhiwen Fan,[s] Xiaolong Qi,[t] Daniel Q. Huang,[u,v] Qing Xie,[w,***] Junping Shi,[x,**] Chao Wu,[a,b,c,z] and Jie Li[b,a,c,*]

[a]Department of Infectious Diseases, Nanjing Drum Tower Hospital Clinical College of Nanjing University of Chinese Medicine, Nanjing, Jiangsu, China
[b]Department of Infectious Disease, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, Jiangsu, China
[c]Institute of Viruses and Infectious Diseases, Nanjing University, Nanjing, Jiangsu, China
[d]Karsh Division of Gastroenterology and Hepatology, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA
[e]Clinical School of the Second People's Hospital, Tianjin Medical University, Tianjin, China
[f]Department of Hepatology, Tianjin Second People's Hospital, Tianjin, China
[g]Tianjin Research Institute of Liver Diseases, Tianjin, China
[h]Department of Hepatology, Hepatology Research Institute, The First Affiliated Hospital, Fujian Medical University, Fuzhou, Fujian, China
[i]Department of Hepatology, The Third Hospital of Zhenjiang Affiliated Jiangsu University, Zhenjiang, Jiangsu, China
[j]Department of Infectious Diseases and Hepatology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China
[k]Department of Infectious Diseases, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou, Zhejiang, China
[l]School of Computer Science, China University of Geosciences, Wuhan, Hubei, China
[m]Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan, Hubei, China
[n]Department of Infectious Disease, Shandong Provincial Hospital Affiliated to Shandong Frist Medical University, Ji'nan, Shandong, China
[o]Department of Infectious Diseases, The Fifth People's Hospital of Wuxi, Wuxi, Jiangsu, China
[p]Department of Infectious Diseases, The Affiliated Infectious Diseases Hospital of Soochow University, Suzhou, Jiangsu, China
[q]Department of Hepatology, Huai'an No.4 People's Hospital, Huai'an, Jiangsu, China
[r]Department of Infectious Disease, Shandong Provincial Hospital, Shandong University, Ji'nan, Shandong, China
[s]Department of Pathology, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, Jiangsu, China
[t]Center of Portal Hypertension, Department of Radiology, Zhongda Hospital, Medical of School, Southeast University, Nanjing, Jiangsu, China
[u]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[v]Division of Gastroenterology and Hepatology, Department of Medicine, National University Health System, Singapore
[w]Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China
[x]Department of Infectious & Hepatology Diseases, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, Zhejiang, China

Summary

Background With increasingly prevalent coexistence of chronic hepatitis B (CHB) and hepatic steatosis (HS), simple, non-invasive diagnostic methods to accurately assess the severity of hepatic inflammation are needed. We aimed to build a machine learning (ML) based model to detect hepatic inflammation in patients with CHB and concurrent HS.

Methods We conducted a multicenter, retrospective cohort study in China. Treatment-naive CHB patients with biopsy-proven HS between April 2004 and September 2022 were included. The optimal features for model development were selected by SHapley Additive explanations, and an ML algorithm with the best accuracy to diagnose moderate to severe hepatic inflammation (Scheuer's system ≥ G3) was determined and assessed by decision curve analysis (DCA) and calibration curve. This study is registered with ClinicalTrials.gov (NCT05766449).

eClinicalMedicine
2024;68: 102419
Published Online xxx
https://doi.org/10.1016/j.eclinm.2023.102419

*Corresponding author. Department of Infectious Diseases, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, Jiangsu, China.
**Corresponding author. Department of Infectious & Hepatology Diseases, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, Zhejiang, China.
***Corresponding author. Department of Infectious Diseases, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China.
E-mail addresses: lijier@nju.edu.cn (J. Li), 20131004@hznu.edu.cn (J. Shi), xieqingrjh@163.com (Q. Xie).
[y]These authors contributed equally.
[z]Senior author.

Findings From a pool of 1,787 treatment-naive patients with CHB and HS across eleven hospitals, 689 patients from nine of these hospitals were chosen for the development of the diagnostic model. The remaining two hospitals contributed to two independent external validation cohorts, comprising 509 patients in validation cohort 1 and 589 in validation cohort 2. Eleven features regarding inflammation, hepatic and metabolic functions were identified. The gradient boosting classifier (GBC) model showed the best performance in predicting moderate to severe hepatic inflammation, with an area under the receiver operating characteristic curve (AUROC) of 0.86 (95% CI 0.83–0.88) in the training cohort, and 0.89 (95% CI 0.86–0.92), 0.76 (95% CI 0.73–0.80) in the first and second external validation cohorts, respectively. A publicly accessible web tool was generated for the model.

Interpretation Using simple parameters, the GBC model predicted hepatic inflammation in CHB patients with concurrent HS. It holds promise for guiding clinical management and improving patient outcomes.

Funding This research was supported by the National Natural Science Foundation of China (No. 82170609, 81970545), Natural Science Foundation of Shandong Province (Major Project) (No. ZR2020KH006), Natural Science Foundation of Jiangsu Province (No.BK20231118), Tianjin Key Medical Discipline (Specialty), Construction Project, TJYXZDXK-059B, Tianjin Health Science and Technology Project key discipline special, TJWJ2022XK034, and Research project of Chinese traditional medicine and Chinese traditional medicine combined with Western medicine of Tianjin municipal health and Family Planning Commission (2021022).

Keywords: Chronic hepatitis B; Hepatic steatosis; Inflammation; Machine learning; Diagnostic model

Research in context

Evidence before this study

We searched PubMed for publications without language restriction, published before September 30th, 2023, using search terms "chronic hepatitis B" and "hepatic steatosis" and "inflammation", with search terms found in abstract, title or MESH headings. Although there have been studies on chronic hepatitis B (CHB) with concurrent hepatic steatosis (HS), no diagnostic model for hepatic inflammation in CHB patients with concurrent HS has been established.

Added value of this study

In this study, we used machine learning algorithms to construct a diagnostic model for hepatic inflammation, by deriving data from 1,787 patients with liver biopsy-confirmed CHB with concurrent HS in eleven hospitals across China. Our results showed that the gradient boosting classifier (GBC) model including eleven features [aspartate transaminase, prothrombin time, platelet, albumin, gamma-glutamyl transferase, hepatitis B e antigen (HBeAg)-positive, hepatitis B surface antigen, white blood cell, international normalized ratio, body mass index and total bilirubin] had the highest accuracy. The area under the receiver operating characteristic curve (AUROC) of the GBC model was 0.86 (95% CI 0.83–0.88), 0.89 (95% CI 0.86–0.92) and 0.76 (95% CI 0.73–0.80) in the training, external validation 1 and external validation 2 cohorts, respectively. Further, subgroup analysis was employed to determine the performance of the GBC model in subgroups stratified by age, sex, body mass index, HBeAg status, HBV DNA level, and presence of diabetes.

Implications of all the available evidence

The GBC model established in our study has the potential to be a valuable diagnostic tool for identifying hepatic inflammation in treatment-naive CHB patients with concurrent HS. It can help clinicians assess the severity of liver inflammation. To facilitate the use of the GBC model, a free website was created: https://py3.reallife-liver.com/.

Introduction

Chronic hepatitis B (CHB) is the most prevalent viral liver disease, with an estimated global prevalence of 4.1% in 2019, accounting for over 300 million infections.[1–4] Meanwhile, the global prevalence of fatty liver reached 30% in 2019 due to improved living standards, changes in lifestyle, and dietary habits.[5] Notably, in regions like Asia,[6] where CHB is endemic, the coexistence of fatty liver and hepatitis B virus (HBV) infection is likely to be common. Our previous meta-analysis showed that the prevalence of hepatic steatosis (HS) among patients with CHB was 32.83%.[7] Both CHB and HS are predominant contributors to chronic liver injury, increasing the risk of cirrhosis and hepatocellular carcinoma (HCC).[8,9] Furthermore, studies have confirmed that the presence of steatohepatitis significantly increases the risk of liver-related complications and mortality in CHB patients.[10,11]

Chronic hepatic inflammation is a key driving force for hepatic fibrosis, cirrhosis, and even HCC.[12–17]

Guidelines recommend that moderate to severe inflammation confirmed by liver biopsy is an important indication for antiviral therapy.[17–20] Timely and accurate diagnosis of hepatic inflammation is crucial for guiding clinical management and improving the prognosis of patients with chronic liver diseases (CLDs). Although liver biopsy has long been considered the gold standard for diagnosing hepatic inflammation, it suffers from several drawbacks, such as invasiveness, high cost, and the risk of procedure-related complications. As a result, liver biopsy is not feasible for all patients with CHB. Therefore, non-invasive tests (NITs) based on serum indicators and imaging methods have gained favor in clinical practice.

Currently, several NITs have been developed for identifying hepatic inflammation in different contexts. Zhang et al.[21] established a nomogram specifically for CHB patients with normal alanine transaminase (ALT), which was based on hepatitis B e antigen (HBeAg), aspartate transaminase (AST), and platelet (PLT), with an area under the receiver operating characteristic curve (AUROC) of 0.751. Similarly, Chen et al.[22] established a non-invasive diagnostic model for CHB patients with ALT ≤2 times the upper limit of normal, utilizing AST, gamma-glutamyl transferase (GGT), anti-hepatitis B virus core antibody (anti-HBc), and prothrombin time (PT). This model demonstrated an AUROC of 0.767.

Machine learning (ML) uses artificial intelligence to generate predictive models more efficiently than traditional methods by uncovering hidden patterns within large, complex datasets. An increasing number of studies have applied ML techniques to facilitate disease progression prediction in chronic viral hepatitis, nonalcoholic fatty liver disease (NAFLD), cirrhosis complications, and other CLDs, making ML-based models become valuable tools for implementation in clinical practice.[23–26] In this study, we aimed to develop a ML-based diagnostic model for moderate to severe hepatic inflammation (Scheuer's system ≥ G3) in CHB patients with concurrent HS to provide a non-invasive and cost-effective alternative to liver biopsy, enabling early detection of hepatic inflammation in these patients.

Methods

Study population

This multicenter, retrospective study included treatment-naive CHB patients with concurrent HS with liver biopsy between April 2004 and September 2022 from eleven Chinese hospitals. Patients were divided into a training cohort (Nanjing Drum Tower Hospital; The First Affiliated Hospital of Fujian Medical University; The Affiliated Hospital of Hangzhou Normal University; The First Affiliated Hospital of Zhengzhou University; Taizhou Hospital of Zhejiang Province affiliated to Wenzhou Medical University; The Fifth People's Hospital of Wuxi; The Affiliated Infectious Diseases Hospital of Soochow University; Huai'an No.4 People's Hospital; The Third Hospital of Zhenjiang Affiliated Jiangsu University) and two independent external validation cohorts (Tianjin Second People's Hospital [validation cohort 1] and Shanghai Ruijin Hospital [validation cohort 2]). This study is registered with ClinicalTrials.gov (NCT05766449).

Ethics statement

This study has received approval from the Institutional Ethics Review Board of all the involved hospitals. As this was a retrospective study, the requirement for informed consent was waived.

Inclusion and exclusion criteria

Adult participants with CHB [defined as a positive hepatitis B surface antigen (HBsAg) for at least 6 months] who were treatment naive before liver biopsy, and were concurrent with HS [steatosis >5% on histological assessment (training cohort and validation cohort 1) or ultrasound assessment (validation cohort 2)] were included. The exclusion criteria were as follows: (1) other viral co-infections such as hepatitis C virus, hepatitis D virus, hepatitis E virus, and human immunodeficiency virus); (2) coexistence with alcohol-associated fatty liver disease, autoimmune hepatitis, and other CLDs; (3) concurrent with HCC and other malignancies; (4) complicated with various end-stage diseases, and systemic inflammation such as ongoing infection, respiratory failure, decompensated heart failure, and flare of autoimmune diseases; (5) insufficient clinical data in the medical file.

Data collection

Before initiating the data collection process, training sessions on the data extraction form were conducted for all involved staff in participating centers. Medical records of eligible patients were meticulously reviewed. Clinical and laboratory data, including demographic characteristics, routine blood tests, liver biochemistry, coagulation, and virology data, were systematically extracted into standardized forms. To ensure data relevancy and consistency, only laboratory examinations conducted within a 14-day window before the liver biopsy were considered. A second reviewer randomly assessed and verified 30% of the extracted data. Given the variations in units across participating centers, the units of all laboratory measurements were harmonized. Extreme outliers, either significantly large or small values, were flagged for review. Such outliers were re-evaluated by the site's principal investigator or the designated physician-in-charge to ascertain data validity and to rule out data entry errors.

Liver biopsy

Ultrasound-guided liver biopsies were performed with 16-gauge biopsy needles to obtain hepatic specimens

with at least 1 cm and six portal tracts. Pathological examinations were performed by two independent board-certified pathologists from the respective centers who were blinded to the clinical data. Discrepancies were resolved through discussion, reaching a consensus. Inter-observer variability was assessed to evaluate the agreement between the pathologists (Supplemental Method). The obtained specimens were assessed using Scheuer's[27] criteria to evaluate the grade of liver inflammation (0: none; G1: mild inflammation; G2: mild to moderate inflammation; G3: moderate inflammation; and G4: severe inflammation.[28] Moreover, the Scheuer and Brunt criteria were utilized respectively to assess fibrosis staging (0–4) and the degree of steatosis (0–3). Scheuer staging system is extensively used by pathologists in each research center as a principal method for staging. In this system, grades G3 to G4 indicate moderate to severe inflammation.[28]

Overview of ML models

We developed and compared seven models, including gradient boosting classifier (GBC), random forest (RF), eXtreme Gradient Boosting (XGB), adaptive boosting (ADB), gaussian naive bayes (GNB), logistic regression (LR) and K-nearest neighbors (KNN) to predict moderate to severe hepatic inflammation. GBC builds a strong classifier by iteratively training a weak classifier. Each iteration adjusts the weight of the sample so that previously misclassified samples receive more attention.[29] RF is an ensemble learning model that is based on multiple single decision trees.[30] By aggregating the results of these decision trees into one result, it compensates for the disadvantage of overfitting in single decision tree and reduces variance.[31,32] XGB is an ensemble learning model based on the gradient boosting algorithm.[33] It builds a strong classifier by iteratively training multiple weak classifiers and optimizing the loss function using gradient descent. XGB performs well when processing structured data and large-scale data sets, with efficiency and accuracy. ADB adjusts the weight of the sample based on the classification error rate of the sample, so that samples with higher classification error rates receive more attention, and the final classification result is determined through weighted voting.[34,35] GNB is a probabilistic classifier based on 'bayes' theorem. It assumes that features are independent and uses Gaussian distribution to model the probability distribution of continuous features.[36] GNB is simple, fast, and works well for classification problems with high-dimensional data and small samples. LR is a learning algorithm with a logistic function as its core. It assesses the relationship between dependent variables and one or more independent variables, deriving classification probabilities using logical functions.[37] KNN is an instance-based learning algorithm. It classifies a new sample into the most common category among its K nearest neighbors by measuring the distance between samples. KNN is simple and easy to understand, and is suitable for problems with multi-category classification and non-linear decision boundaries.

Primary objective

The primary objective of this study was to develop a ML-based model for diagnosing moderate to severe hepatic inflammation in CHB patients who have concurrent HS.

Statistical analysis

Patient data were presented as continuous or categorical variables. Kolmogorov-smirnov test was used to assess whether the data followed a normal distribution. For normally-distributed continuous variables, the data were described as mean ± standard deviation and compared using the t-test. If continuous variables did not conform to a normal distribution, the Mann–Whitney U test was used, and results were presented as median (interquartile range). Categorical data were presented as numbers and frequencies, and either the Chi-square test or Fisher's exact test was used for comparisons. All statistical tests were two-sided, with *P*-values <0.05 indicating statistical significance. SPSS software version 25.0 (SPSS Inc., Chicago, IL) was used in this study.

In the training queue, over 35% of missing parameters were excluded from the analysis. 23 variables of the training cohort, including sex, age, body mass index (BMI), diabetes mellitus, hypertension, white blood cell (WBC), neutrophils, hemoglobin (Hb), PLT, total bilirubin (Tbil), albumin (ALB), ALT, AST, alkaline phosphatase (ALP), GGT, uric acid (UA), total cholesterol (TC), triglyceride (TG), PT, international normalized ratio (INR), HBsAg, HBeAg-positive, and HBV DNA $\geq 10^5$ IU/mL, were included in seven machine algorithms to generate a pre-model for diagnosing moderate to severe hepatic inflammation. To avoid overfitting, five-fold cross-validations were used in the ML model-building process.[38] The SHapley Additive exPlanations score[39] was calculated to assess the importance of each clinical feature and provide a quantitative description of the overall relationship between moderate to severe hepatic inflammation and all 23 features according to the pre-model. This process requires sequentially integrating features, starting with the most important feature, and gradually adding the next feature in order of importance. Through this method, we conducted 23 feature inclusion iterations and ultimately selected the number of features that generated the highest AUROC in these iterations. The optimal cutoff of GBC model was estimated using the Youden index. The diagnosis performances for the ML models were assessed by the AUROC, sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV).[40] The Delong test[41] was used to compare the

differences between various AUROCs. Furthermore, decision curve analysis (DCA)[42,43] and calibration[44] were used to assess the diagnostic performance of the model. Missing data were addressed by multiple imputations.[45] Python programming software (version 3.8.0) was used in this study. Multiple imputation was performed on the original data set using miceforest. We utilized stratified random sampling to segment the dataset, a process executed using the StratifiedKFold function from the sklearn library. This approach allowed for a rigorous comparative analysis of various machine learning models, including GBC, RF, XGB, ADB, GNB, LR and KNN.

Finally, the web-based tool for moderate to severe hepatic inflammation was developed with open access at https://py3.reallife-liver.com/.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Patient characteristics

A total of 2,455 CHB patients with concurrent HS who underwent liver biopsies were identified. Among them, 668 patients who failed to meet the inclusion criteria were excluded. Finally, 1,787 patients were included for analysis (Fig. 1).

The baseline characteristics of the training cohort (n = 689), validation cohort 1 (n = 509), and validation cohort 2 (n = 589) are presented in Table 1. In the training cohort, the median age was 39.0, and the median BMI was 24.8 kg/m^2. Among them, 78.1% were male, and 9.0% had diabetes. In validation cohort 1, the median age was 36.0 years, and the median BMI was 26.3 kg/m^2, with 73.1% being male and 6.9% having diabetes. For validation cohort 2, the median age mirrored that of the training cohort, at 39.0 years old. The median BMI was 24.2 kg/m^2, 75.2% of them were male and only 2.7% had diabetes.

Based on liver biopsy, there were 157 (22.8%), 31 (6.1%), and 93 (15.8%) had moderate to severe hepatic inflammation in the training cohort, the validation cohort 1, and the validation cohort 2, respectively. The comparison of population characteristics between patients with moderate to severe hepatic inflammation and those without moderate to severe hepatic inflammation is presented in Supplementary Table S1.

Diagnostic performance of moderate to severe hepatic inflammation in the final GBC model

Twenty-three variables from the training cohort were integrated into ML algorithms to create pre-models (Supplementary Fig. S1). Following this, the SHapley Additive exPlanations method was applied to evaluate the significance of each variable. Ultimately, eleven variables from the training cohort (AST, PT, PLT, ALB, GGT, HBeAg-positive, HBsAg, WBC, INR, BMI, and Tbil) were incorporated into the final GBC model for identifying moderate to severe inflammation (Fig. 2A). Furthermore, the variables included in the final RF (including AST, ALB, PT, PLT, GGT, INR, ALT, BMI, HBeAg-positive, and HBsAg), XGB (including AST, PT,

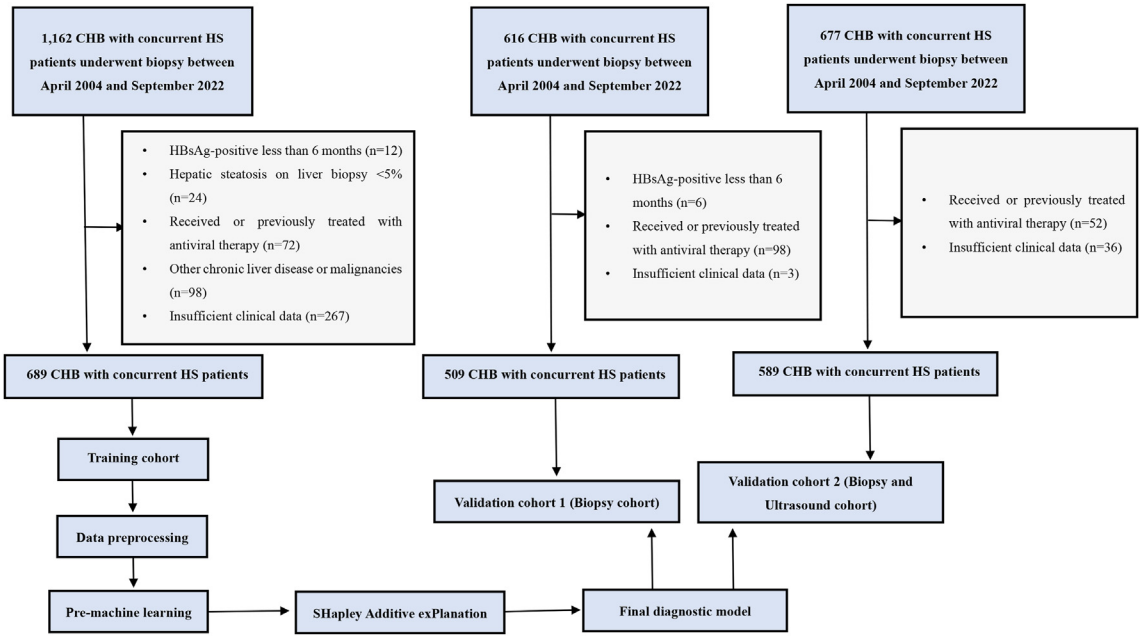


***Fig. 1:* Flow diagram of the study population**. CHB, Chronic hepatitis B; HS, hepatic steatosis; HBsAg, hepatitis B surface antigen.

Characteristics	Training cohort (n = 689)	Validation cohort 1 (n = 509)	Validation cohort 2 (n = 589)
Male, n/total (%)	538/689 (78.1)	372/509 (73.1)	443/589 (75.2)
Age (years)	39.0 (32.5–48.0)	36.0 (30.0–45.0)	39.0 (32.0–48.0)
Age <40 years, n/total (%)	346/689 (50.2)	314/509 (61.7)	299/589 (50.8)
Age ≥40 years, n/total (%)	343/689 (49.8)	195/509 (38.3)	290/589 (49.2)
Body mass index (kg/m^2)	24.8 (22.8–26.8)	26.3 (24.2–28.7)	24.2 (22.2–26.4)
Body mass index <25 kg/m^2, n/total (%)	315/606 (52.0)	181/509 (35.6)	336/568 (59.2)
Body mass index ≥25 kg/m^2, n/total (%)	291/606 (48.0)	328/509 (64.4)	232/568 (40.8)
Diabetes mellitus, n/total (%)	62/689 (9.0)	35/509 (6.9)	16/589 (2.7)
Hypertension, n/total (%)	56/689 (8.1)	68/509 (13.4)	24/589 (4.1)
White blood cell (10^9/L)	5.4 (4.7–6.4)	5.6 (4.7–6.5)	5.6 (4.9–6.7)
Neutrophils (10^9/L)	3.0 (2.4–3.7)	3.1 (2.5–4.0)	3.3 (2.6–4.0)
Hemoglobin (g/L)	150.0 (140.0–159.0)	154.0 (141.0–162.0)	152.0 (140.0–161.0)
Platelet (10^9/L)	186.0 (152.0–222.5)	207.0 (169.0–247.0)	174.0 (142.0–212.0)
Total bilirubin (μmol/L)	14.1 (11.1–18.4)	14.5 (11.4–18.2)	15.7 (12.0–20.4)
Albumin (g/L)	43.9 (41.0–46.1)	45.8 (43.0–48.4)	43.0 (41.0–46.0)
Alanine transaminase (U/L)	48.3 (31.0–86.0)	44.0 (26.1–77.0)	44.0 (29.5–67.0)
Aspartate transaminase (U/L)	33.0 (24.0–49.0)	28.5 (22.0–46.0)	33.0 (26.0–48.0)
Alkaline phosphatase (U/L)	80.3 (62.0–101.0)	67.5 (55.0–83.0)	71.0 (59.0–85.0)
Gamma-glutamyl transferase (U/L)	32.0 (22.0–52.0)	35.0 (23.0–55.0)	29.0 (20.0–44.0)
Uric acid (μmol/L)	338.4 (284.0–394.2)	337.0 (279.0–398.0)	N.A
Total cholesterol (mmol/L)	4.5 (3.9–5.1)	4.6 (4.1–5.2)	N.A
Triglyceride (mmol/L)	1.3 (0.9–1.8)	1.2 (0.9–1.7)	N.A
Prothrombin time (s)	12.4 (11.4–13.2)	12.5 (11–13.2)	11.8 (11.1–12.8)
International normalized ratio	1.0 (1.0–1.1)	1.0 (0.9–1.0)	1.0 (0.9–1.0)
HBsAg (IU/mL)	1971.6 (370.3–6894.0)	4112.5 (1486.1–15544.5)	3024.4 (992.1–11651.2)
HBeAg (S/CO)	0.6 (0.3–291.7)	1.5 (0.4–1205.5)	6.4 (0.4–978.8)
HBeAg-Positive, n/total (%)	293/654 (44.8)	223/441 (50.6)	309/589 (52.5)
HBV DNA (log10 IU/mL)	5.2 (3.4–7.4)	4.5 (2.5–7.9)	5.3 (4.0–7.0)
HBV DNA <10^5 IU/mL, n/total (%)	354/689 (51.4)	276/509 (54.2)	266/589 (45.2)
HBV DNA ≥10^5 IU/mL, n/total (%)	335/689 (48.6)	233/509 (45.8)	323/589 (54.8)

Note: HBsAg, hepatitis B surface antigen; HBeAg, hepatitis B e antigen. Values presented as n (%), median (interquartile range).

Table 1: **Baseline characteristics of the cohort.**

PLT, ALB, GGT, HBsAg, BMI, HBeAg-positive, INR, Tbil, TC, and TG), ADB (including AST, PLT, PT, GGT, ALB, HBsAg, INR, BMI, HBeAg-positive, neutrophils, and Tbil), GNB (including AST, ALT, Hb, ALB, PT, and PLT), LR (including AST, ALT, PLT, Hb, HBsAg, Age, and ALB), and KNN (including HBsAg, ALT, PLT, UA, AST, and GGT) diagnostic models were also shown in Fig. 2. Physicians could calculate the GBC model freely at https://py3.reallife-liver.com/.

For diagnosing moderate to severe hepatic inflammation, the AUROCs of seven final ML models (GBC, RF, XGB, ADB, GBC, LR, and KNN) ranged from 0.62 to 0.86 in the training cohort. In the validation cohort 1 and 2, the AUROCs ranged from 0.61 to 0.89 and from 0.72 to 0.78, respectively (Fig. 3). Due to the absence of UA, TC and TG indicators in the validation cohort 2, only the performance of GBC, RF, ADB, GNB, and LR models were calculated. The GBC model obtained the optimal AUROC in the training cohort and validation cohort 1, but had lower performance in validation cohort 2.

The optimal cutoff of the final GBC model for diagnosing moderate to severe hepatic inflammation was 0.218. At this cutoff, the corresponding accuracy, sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) of the GBC model in the training cohort were 0.80 (95% CI 0.77–0.83), 0.75 (95% CI 0.72–0.78), 0.81 (95% CI 0.78–0.84), 0.53 (95% CI 0.50–0.57) and 0.92 (95% CI 0.90–0.94) (Table 2). In the validation cohort 1, the GBC model achieved a high SE (0.85 [95% CI 0.82–0.88]). However, the SE of the GBC model in validation cohort 2 showed a lower value of 0.58 (95% CI 0.54–0.62) corresponding to the optimal cutoff (Table 2).

Performance of the GBC model for moderate to severe hepatic inflammation diagnosis under different subgroups

Subgroup analysis can provide a deeper understanding of the diagnostic performance of the GBC model in specific patient populations. Therefore, the performance of the GBC model in diagnosing moderate to severe

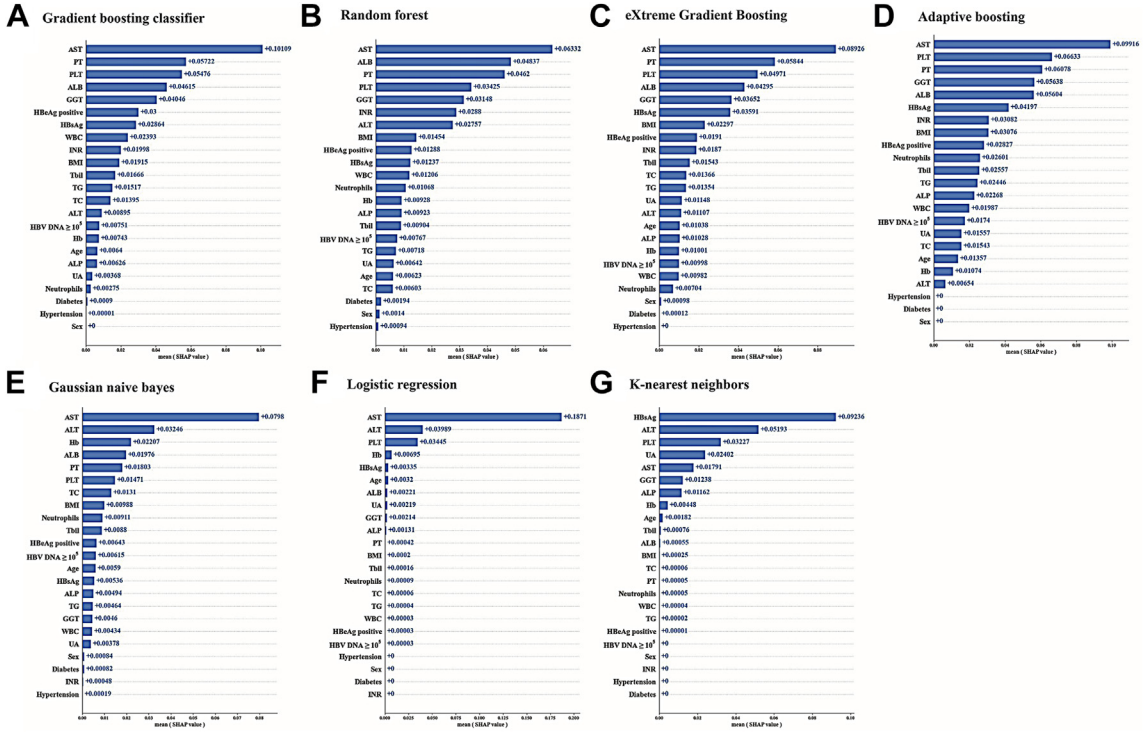
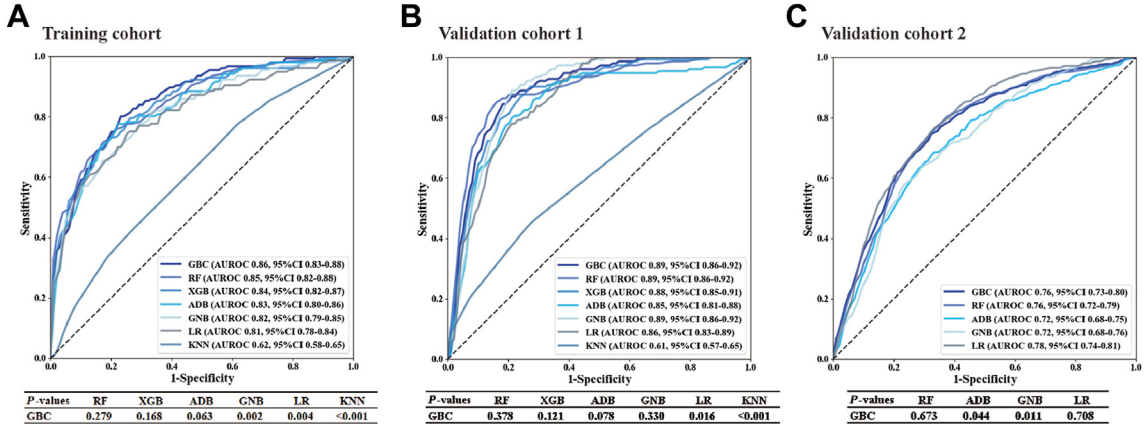


Fig. 2: **SHapley Additive exPlanations plot: the impact of clinical features for diagnosing moderate to severe hepatic inflammation in the gradient boosting classifier, random forest, eXtreme Gradient Boosting, adaptive boosting, gaussian naive bayes, logistic regression, and K-nearest neighbors**. A. gradient boosting classifier, B. random forest, C. eXtreme Gradient Boosting, D. adaptive boosting, E. gaussian naive bayes, F. logistic regression, G. K-nearest neighbors. Note: BMI, body mass index; WBC, white blood cell; Hb, hemoglobin; PLT, platelet; Tbil, total bilirubin; ALB, albumin; ALT, alanine transaminase; AST, aspartate transaminase; ALP, alkaline phosphatase; GGT, gamma-glutamyl transferase; UA, uric acid; TC, total cholesterol; TG, triglyceride; PT, prothrombin time; INR, international normalized ratio, HBsAg, hepatitis B surface antigen; HBeAg, hepatitis B e antigen.



P-values	RF	XGB	ADB	GNB	LR	KNN
GBC	0.279	0.168	0.063	0.002	0.004	<0.001

P-values	RF	XGB	ADB	GNB	LR	KNN
GBC	0.378	0.121	0.078	0.330	0.016	<0.001

P-values	RF	ADB	GNB	LR
GBC	0.673	0.044	0.011	0.708

Fig. 3: **Diagnostic performance of the final machine learning models for moderate to severe hepatic inflammation in training cohort and external validation cohorts**. A. Training cohort, B. Validation cohort 1, C. Validation cohort 2. Note: GBC, gradient boosting classifier; RF, random forest; XGB, eXtreme Gradient Boosting; ADB, adaptive boosting; GNB, gaussian naive bayes; LR, logistic regression; KNN, K-nearest neighbors.

	Accuracy	Area under receiver operating characteristic curve	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Training cohort						
GBC	0.80 (0.77–0.83)	0.86 (0.83–0.88)	0.75 (0.72–0.78)	0.81 (0.78–0.84)	0.53 (0.50–0.57)	0.92 (0.90–0.94)
RF	0.80 (0.77–0.83)	0.85 (0.82–0.88)	0.69 (0.66–0.73)	0.83 (0.80–0.86)	0.55 (0.52–0.59)	0.90 (0.88–0.93)
XGB	0.79 (0.76–0.82)	0.84 (0.82–0.87)	0.71 (0.68–0.75)	0.81 (0.78–0.84)	0.53 (0.49–0.56)	0.91 (0.89–0.93)
ADB	0.76 (0.73–0.80)	0.83 (0.80–0.86)	0.73 (0.70–0.76)	0.77 (0.74–0.80)	0.49 (0.45–0.53)	0.91 (0.89–0.93)
GNB	0.79 (0.76–0.82)	0.82 (0.79–0.85)	0.65 (0.62–0.69)	0.83 (0.80–0.86)	0.54 (0.50–0.57)	0.89 (0.87–0.92)
LR	0.76 (0.73–0.79)	0.81 (0.78–0.84)	0.70 (0.67–0.73)	0.78 (0.75–0.81)	0.49 (0.46–0.53)	0.90 (0.87–0.92)
KNN	0.67 (0.63–0.70)	0.62 (0.58–0.65)	0.42 (0.38–0.46)	0.74 (0.71–0.78)	0.33 (0.29–0.36)	0.81 (0.78–0.84)
Validation cohort 1						
GBC	0.79 (0.76–0.83)	0.89 (0.86–0.92)	0.85 (0.82–0.88)	0.79 (0.76–0.83)	0.21 (0.17–0.25)	0.99 (0.98–1.00)
RF	0.84 (0.81–0.87)	0.89 (0.86–0.92)	0.85 (0.82–0.88)	0.84 (0.80–0.87)	0.25 (0.22–0.29)	0.99 (0.98–1.00)
XGB	0.79 (0.76–0.83)	0.88 (0.85–0.91)	0.83 (0.79–0.86)	0.79 (0.76–0.83)	0.21 (0.17–0.24)	0.99 (0.98–1.00)
ADB	0.77 (0.74–0.81)	0.85 (0.81–0.88)	0.81 (0.77–0.84)	0.77 (0.73–0.81)	0.19 (0.15–0.22)	0.98 (0.97–0.99)
GNB	0.83 (0.80–0.86)	0.89 (0.86–0.92)	0.75 (0.72–0.79)	0.84 (0.80–0.87)	0.23 (0.20–0.27)	0.98 (0.97–0.99)
LR	0.81 (0.77–0.84)	0.86 (0.83–0.89)	0.75 (0.71–0.79)	0.81 (0.78–0.84)	0.20 (0.17–0.24)	0.98 (0.97–0.99)
KNN	0.70 (0.66–0.74)	0.61 (0.57–0.65)	0.46 (0.41–0.50)	0.72 (0.68–0.76)	0.10 (0.07–0.12)	0.95 (0.94–0.97)
Validation cohort 2						
GBC	0.77 (0.74–0.81)	0.76 (0.73–0.80)	0.58 (0.54–0.62)	0.81 (0.78–0.84)	0.36 (0.32–0.40)	0.91 (0.89–0.93)
RF	0.78 (0.74–0.81)	0.76 (0.72–0.79)	0.47 (0.43–0.51)	0.83 (0.80–0.86)	0.35 (0.31–0.39)	0.89 (0.87–0.92)
ADB	0.72 (0.68–0.75)	0.72 (0.68–0.75)	0.60 (0.56–0.64)	0.74 (0.70–0.77)	0.30 (0.26–0.34)	0.91 (0.89–0.93)
GNB	0.76 (0.73–0.80)	0.72 (0.68–0.76)	0.42 (0.38–0.45)	0.83 (0.80–0.86)	0.31 (0.27–0.35)	0.88 (0.86–0.91)
LR	0.70 (0.67–0.74)	0.78 (0.74–0.81)	0.73 (0.70–0.77)	0.70 (0.66–0.74)	0.31 (0.28–0.35)	0.93 (0.91–0.95)

Note: GBC, gradient boosting classifier; RF, random forest; XGB, eXtreme Gradient Boosting; ADB, adaptive boosting; GNB, gaussian naive bayes; LR, logistic regression; KNN, K-nearest neighbors.

Table 2: **Diagnostic performance of the machine learning models for moderate to severe hepatic inflammation.**

hepatic inflammation was investigated in different subgroups (age, sex, BMI, HBeAg status, HBV DNA level, and presence of diabetes). In both the training cohort and the validation cohorts, the GBC model correspondingly had the highest AUROCs of 0.87 (95% CI 0.84–0.91), 0.90 (95% CI 0.86–0.94), and 0.79 (95% CI 0.74–0.83) in patients with HBV DNA $\geq 10^5$ IU/mL (Supplementary Table S2).

Evaluation of the GNB model performance for moderate to severe hepatic inflammation diagnosis by DCA and calibration curve

The DCA for the GBC model demonstrated a consistent net benefit in the training cohort and two validation cohorts over a range of threshold probabilities. In all three cohorts, the GBC model outperformed the 'treat none' strategy, indicating that it had practical utility in decision-making (Fig. 4). The calibration curve showed that the GBC model achieved good diagnostic performance in the training cohort, but underestimated the risk of moderate to severe hepatic inflammation in the validation cohort 1 (Supplementary Fig. S2).

Discussion

In this study, we developed and validated a ML-based model to predict moderate to severe hepatic inflammation for treatment-naive CHB patients with concurrent HS. The diagnostic model exhibited accuracy in both the training and validation cohorts and remained stable across different subgroups. In terms of predictive values, the GBC model generally demonstrates a high NPV, with most values exceeding 0.9. This indicated the model's accuracy and stability in excluding cases of moderate to severe hepatic inflammation. Although the PPV in the external validation cohort 1 was lower, this was likely attributed to the lower prevalence of individuals with moderate to severe hepatic inflammation in this group. Further, DCA and calibration curves were used to evaluate the net benefit and diagnostic performance of the GBC model. Our study showed that the GBC model has the potential to identify moderate to severe hepatic inflammation in CHB patients with concurrent HS, which may assist clinicians in assessing disease severity for further examinations and aiding decision for treatment.

Chronic hepatic inflammation plays a pivotal role in driving hepatic adverse events. A cohort study conducted in North America involving 420 HBsAg-positive patients found that patients combined with steatohepatitis had a significantly higher risk of advanced fibrosis, compared with patients without fatty liver disease or steatosis alone.[11] The initiation of hepatic fibrosis typically stems from inflammation, and pro-inflammatory factors could activate hepatic stellate cells from quiescent status to a fibrogenic phenotype,

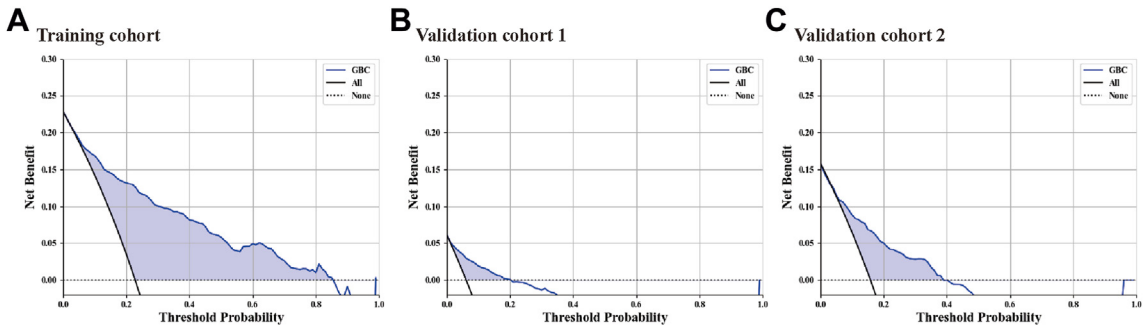


Fig. 4: **Net benefits of the GBC model by decision curve analysis for diagnosing moderate to severe hepatic inflammation**. A. Training cohort, B. Validation cohort 1, C. Validation cohort 2. Note: GBC, gradient boosting classifier.

thereby exacerbating the disease severity.[46] At the same time, the epidemic of fatty liver disease makes the combination of CHB and HS a common clinical phenomenon, potentially accelerating the progression of liver diseases. Therefore, early and accurate identification of inflammation, along with timely intervention, becomes essential for improving the prognosis of CHB patients with concurrent HS. However, most existing non-invasive scoring models based on serum indicators have shown poor performance. Other scoring models established on serum biomarkers combined with imaging methods can improve diagnostic accuracy to some extent, but they greatly increase the cost of examinations. Furthermore, current non-invasive scoring models primarily target either CHB or HS patients, thereby constraining their applicability to patients with both conditions.[21,22,47–50]

With the advancement of ML, clinicians are now able to turn high volumes of data into feasible models to improve their competence in diagnosing disease. A growing number of studies have developed ML models to assess hepatic pathological changes in patients with CLDs.[51–54] For instance, Wang et al. developed a dual-task convolutional neural network model based on ultrasonic shear wave elastography for CHB patients, reaching an AUROC of about 0.8 in staging hepatic inflammation.[51] Zhou et al. built a diagnostic model based on the random forest-backward feature elimination algorithm,[54] which achieved an AUROC of over 0.8 in grading the severity of hepatic inflammation in 650 CHB patients. However, these studies only recruited patients with a single liver disease, which cannot meet the clinical needs of patients who were concomitant with two important CLDs. Compared with traditional methods, ML algorithms have demonstrated sound performance in disease diagnosis. Therefore, we established and validated the GBC model tailored for CHB patients with concurrent HS. The GBC model, based on serum indicators for the diagnosis of moderate to severe hepatic inflammation, reached an AUROC of 0.86, showing similar accuracy to expensive and less common imaging models.[48–50]

Our study identified eleven clinical features, including AST, PT, PLT, ALB, GGT, HBeAg-positive, HBsAg, WBC, INR, BMI, and Tbil, that were significantly correlated with moderate to severe hepatic inflammation. The correlation of AST and GGT with hepatic inflammation has been well recognized, and they are commonly used in model development.[47,54,55] Moreover, researchers have also confirmed that ALB, PLT, and INR were significantly associated with the severity of hepatic inflammation.[21,22,56,57] The likely reason for this is that severe damage to liver cell mass and hepatic inflammation can affect the production of thrombopoietin and ALB.[58–60] PT is another important indicator for liver synthesis and reserve function.[22,61] It reflects the severity and prognosis of liver disease, which makes it an essential variable in the establishment of diagnostic models.[22,56]

BMI is one of the important indicators in NAFLD Fibrosis score (NFS), which can be used to evaluate the severity of fibrosis in HS patients.[62] Our study further found that BMI was an important indicator for assessing hepatic inflammation in CHB patients with concurrent HS. In addition, virological-related indicators, such as HBsAg level and HBeAg status, are also associated with the degree of hepatic inflammation. Li et al. previously constructed a diagnostic model for hepatic inflammation in CHB patients based on HBsAg levels and other serological indicators,[47] while Zhang et al. also combined HBeAg with other indicators to establish a diagnostic model for hepatic inflammation in HBeAg-positive CHB patients.[21]

In addition, our study assessed the performance of the GBC model in diverse subgroups to detect any potential biases present within the datasets. The consistent results across different subgroups proved the generalizability and stability of the GBC model. Notably, in subpopulation living with diabetes, an AUROC of over 0.8 was achieved in both the training cohort and external validation cohort 1, and an AUROC of 0.76 was also

achieved in the external validation cohort 2. Diabetes, primarily characterized by hyperglycemia, can trigger inflammatory responses in liver cells, leading to severe damage caused by mitochondrial oxidative stress, endoplasmic reticulum stress, and reduced lysosomal autophagy.[63,64] Therefore, it is essential to enhance the screening and management of hepatic inflammation in patients with CHB and HS who also have diabetes. Importantly, the diagnostic efficacy of the GBC model was validated in HS patients diagnosed by ultrasound, with an AUROC above 0.75. Therefore, the GBC model is applicable in CHB patients with concurrent HS diagnosed by ultrasound. In cases where liver biopsy cannot be performed, the GBC model developed in this study could be a supportive diagnostic tool for hepatic inflammation. Further validation in other populations is needed to determine its generalizability. Finally, we established a free website to facilitate the use of clinicians.

While our study showed promising results, it is important to acknowledge its limitations. First, the GBC model was developed using data mainly from Chinese patients, necessitating further validation in diverse racial groups to ensure their generalizability. Secondly, the GBC model was specifically developed based on treatment-naive CHB patients, and their diagnostic efficacy for CHB patients with concurrent HS who are undergoing antiviral therapy needs to be further evaluated. Thirdly, the difference between G2 and G3 is minor, and both inter-observer and intra-observer variations objectively exist, even among expert pathologists. However, the analysis for inter-observer variability showed good agreement. Future prospective studies with central reporting of the pathological report are needed to validate our model. Finally, the retrospective nature of data collection from multiple centers led to instances of missing data. Although this limitation may be mitigated by strict inclusion and exclusion criteria and the large sample size, prospective international multicenter studies are required for further validation of the GBC model's performance.

In conclusion, our study developed the GBC model for accurately predicting moderate to severe hepatic inflammation in CHB patients with concurrent HS. The GBC model achieved stable diagnostic performance in different subgroups and two independent external validation cohorts. Furthermore, DCA and calibration curve analysis confirmed the feasibility of the GBC model in diagnosing the severity of hepatic inflammation. With the high accuracy and reliability, the GBC model may be a useful, non-invasive tool to assess the severity of hepatic inflammation in treatment-naive CHB patients with concurrent HS.

Contributors

Study concept and design: Jie Li, Junping Shi, Xiaolong Qi and Yee Hui Yeo; Acquisition of data and technique support: Fajuan Rui, Wenjing Ni, Xiaoming Xu, Liang Xu, Yee Hui Yeo, Youwen Tan, Qi Zheng, Xiaorong Tian, Qing-Lei Zeng, Zebao He, Yuanwang Qiu, Chuanwu Zhu, Weimao Ding, Jian Wang, Rui Huang, Yayun Xu, Qi Xue, Yunliang Chen, Junqing Fan, Zhiwen Fan, Chao Wu; Drafting manuscript: Fajuan Rui, Xiaoming Xu, and Wenjing Ni; Critical revision of the manuscript: Jie Li, Junping Shi, Xiaolong Qi, Yee Hui Yeo, Daniel Q. Huang and Qing Xie. Access and verification of the underlying data: Jie Li and Junping Shi. All authors read and approved the final manuscript.

Data sharing statement

We are unable to provide access to our dataset for privacy reasons. The protocol and statistical analysis methods used in the study can be requested directly from the corresponding author after approval.

Declaration of interests

All authors have no conflict of interest related to this publication.

Acknowledgements

This research was supported by the National Natural Science Fund (No. 82170609, 81970545), Natural Science Foundation of Shandong Province (Major Project), (No. ZR2020KH006), Natural Science Foundation of Jiangsu Province (No.BK20231118), Tianjin Key Medical Discipline (Specialty), Construction Project, TJYXZDXK-059B, Tianjin Health Science and Technology Project key discipline special, TJWJ2022XK034, and Research project of Chinese traditional medicine and Chinese traditional medicine combined with Western medicine of Tianjin municipal health and Family Planning Commission (2021022).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.eclinm.2023.102419.

References

1 Polaris Observatory C. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. *Lancet Gastroenterol Hepatol*. 2018;3(6):383–403.
2 Mao X, Cheung KS, Peng C, et al. Steatosis, HBV-related HCC, cirrhosis, and HBsAg seroclearance: a systematic review and meta-analysis. *Hepatology*. 2023;77(5):1735–1745.
3 Wang J, Liu J, Liu Y, et al. Clinical features of chronic hepatitis B patients with lean nonalcoholic fatty liver disease. *Hepatol Res*. 2023;53(3):184–195.
4 Collaborators GBDHB. Global, regional, and national burden of hepatitis B, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Gastroenterol Hepatol*. 2022;7(9):796–829.
5 Le MH, Yeo YH, Li X, et al. 2019 global NAFLD prevalence: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2022;20(12):2809–2817 e28.
6 Li J, Zou B, Yeo YH, et al. Prevalence, incidence, and outcome of non-alcoholic fatty liver disease in Asia, 1999–2019: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. 2019;4(5):389–398.
7 Zheng Q, Zou B, Wu Y, et al. Systematic review with meta-analysis: prevalence of hepatic steatosis, fibrosis and associated factors in chronic hepatitis B. *Aliment Pharmacol Ther*. 2021;54(9):1100–1109.
8 Tang Y, Fan R, Lan Z, et al. Impact of nonalcoholic fatty liver disease status change on antiviral efficacy of nucleos(t)ide analogues in HBeAg-positive chronic hepatitis B. *J Med Virol*. 2023;95(2):e28501.
9 Li J, Le AK, Chaung KT, et al. Fatty liver is not independently associated with the rates of complete response to oral antiviral therapy in chronic hepatitis B patients. *Liver Int*. 2020;40(5):1052–1061.
10 Choi HSJ, Brouwer WP, Zanjir WMR, et al. Nonalcoholic steatohepatitis is associated with liver-related outcomes and all-cause mortality in chronic hepatitis B. *Hepatology*. 2020;71(2):539–548.

11 Khalili M, Kleiner DE, King WC, et al. Hepatic steatosis and steatohepatitis in a large North American cohort of adults with chronic hepatitis B. *Am J Gastroenterol*. 2021;116(8):1686–1697.
12 Kisseleva T, Brenner D. Molecular and cellular mechanisms of liver fibrosis and its regression. *Nat Rev Gastroenterol Hepatol*. 2021;18(3):151–166.
13 Peiseler M, Schwabe R, Hampe J, Kubes P, Heikenwalder M, Tacke F. Immune mechanisms linking metabolic injury to inflammation and fibrosis in fatty liver disease - novel insights into cellular communication circuits. *J Hepatol*. 2022;77(4):1136–1160.
14 Mountford S, Effenberger M, Noll-Puchta H, et al. Modulation of liver inflammation and fibrosis by interleukin-37. *Front Immunol*. 2021;12:603649.
15 Mohammed S, Nicklas EH, Thadathil N, et al. Role of necroptosis in chronic hepatic inflammation and fibrosis in a mouse model of increased oxidative stress. *Free Radic Biol Med*. 2021;164:315–328.
16 European Association for the Study of the Liver. Electronic address eee, European association for the study of the L. EASL 2017 clinical practice guidelines on the management of hepatitis B virus infection. *J Hepatol*. 2017;67(2):370–398.
17 Terrault NA, Lok ASF, McMahon BJ, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. *Hepatology*. 2018;67(4):1560–1599.
18 EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J Hepatol*. 2017;67(2):370–398.
19 Sarin SK, Kumar M, Lau GK, et al. Asian-Pacific clinical practice guidelines on the management of hepatitis B: a 2015 update. *Hepatol Int*. 2016;10(1):1–98.
20 Zhang C, Liu Y, Li J, et al. Dose-response relationship between qAnti-HBc and liver inflammation in chronic hepatitis B with normal or mildly elevated alanine transaminase based on liver biopsy. *J Med Virol*. 2022;94(8):3911–3923.
21 Zhang L, Yang L, Gao Y, et al. Nomogram for evaluating obvious liver inflammation in treatment-naïve HBeAg positive chronic hepatitis B virus infection patients with normal ALT. *Virulence*. 2023;14(1):2158710.
22 Chen S, Huang H. Clinical non-invasive model to predict liver inflammation in chronic hepatitis B with alanine aminotransferase ≤ 2 upper limit of normal. *Front Med*. 2021;8:661725.
23 Huang H, Shiffman ML, Friedman S, et al. A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. *Hepatology*. 2007;46(2):297–306.
24 Vanderbeck S, Bockhorst J, Komorowski R, Kleiner DE, Gawrieh S. Automatic classification of white regions in liver biopsies by supervised machine learning. *Hum Pathol*. 2014;45(4):785–792.
25 Spann A, Yasodhara A, Kang J, et al. Applying machine learning in liver disease and transplantation: a comprehensive review. *Hepatology*. 2020;71(3):1093–1105.
26 Ahn JC, Connell A, Simonetto DA, Hughes C, Shah VH. Application of artificial intelligence for the diagnosis and treatment of liver diseases. *Hepatology*. 2021;73(6):2546–2563.
27 Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. *J Hepatol*. 1991;13(3):372–374.
28 Shu X, Sun H, Yang X, et al. Correlation of effective hepatic blood flow with liver pathology in patients with hepatitis B virus. *Liver Research*. 2021;5(4):243–250.
29 Bowd C, Belghith A, Proudfoot JA, et al. Gradient-boosting classifiers combining vessel density and tissue thickness measurements for classifying early to moderate glaucoma. *Am J Ophthalmol*. 2020;217:131–139.
30 Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.
31 Chang D, Truong E, Mena EA, et al. Machine learning models are superior to noninvasive tests in identifying clinically significant stages of NAFLD and NAFLD-related cirrhosis. *Hepatology*. 2023;77(2):546–557.
32 Zhang K, Wang J, Liu T, Luo Y, Loh XJ, Chen X. Machine learning-reinforced noninvasive biosensors for healthcare. *Adv Healthc Mater*. 2021;10(17):e2100734.
33 Ester M, Kriegel HP, Xu X. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22Nd ACM SIGKDD international conference on knowledge discovery and data mining (vol, pg 785, 2016)*. Geogr Anal; 2022.
34 Avidan S. Ensemble tracking. *IEEE Trans Pattern Anal Mach Intell*. 2007;29(2):261–271.
35 Freund Y, Schapire R. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci*. 1997;55(1):119–139.
36 Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. *Mach Learn*. 1997;29(2–3):103–130.
37 Jiang X, Zhang Y, Li Y, Zhang B. Forecast and analysis of aircraft passenger satisfaction based on RF-RFE-LR model. *Sci Rep*. 2022;12(1):11174.
38 Wong T-T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recogn*. 2015;48(9):2839–2846.
39 Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. *arXiv*. 2018.
40 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29–36.
41 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845.
42 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2016;26(6):565–574.
43 Kerr KF, Brown MD, Zhu K, Janes H. Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. *J Clin Oncol*. 2016;34(21):2534–2540.
44 Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models. *Epidemiology*. 2010;21(1):128–138.
45 Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. *JAMA*. 2015;314(18):1966–1967.
46 Neshat SY, Quiroz VM, Wang Y, Tamayo S, Doloff JC. Liver disease: induction, progression, immunological mechanisms, and therapeutic interventions. *Int J Mol Sci*. 2021;22(13):6777.
47 Li X, Xing Y, Zhou D, et al. A non-invasive model for predicting liver inflammation in chronic hepatitis B patients with normal serum alanine aminotransferase levels. *Front Med*. 2021;8:688091.
48 Newsome PN, Sasso M, Deeks JJ, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. *Lancet Gastroenterol Hepatol*. 2020;5(4):362–373.
49 Andersson A, Kelly M, Imajo K, et al. Clinical utility of magnetic resonance imaging biomarkers for identifying nonalcoholic steatohepatitis patients at high risk of progression: a multicenter pooled data and meta-analysis. *Clin Gastroenterol Hepatol*. 2022;20(11):2451–2461.e3.
50 Chuah KH, Chan WK. Non-invasive biomarkers for liver inflammation in non-alcoholic fatty liver disease: present and future. *Clin Mol Hepatol*. 2023;29(2):401–403.
51 Wang C, Zheng L, Li Y, et al. Noninvasive assessment of liver fibrosis and inflammation in chronic hepatitis B: a dual-task convolutional neural network (DtCNN) model based on ultrasound shear wave elastography. *J Clin Transl Hepatol*. 2022;10(6):1077–1085.
52 Forlano R, Mullish BH, Giannakeas N, et al. High-throughput, machine learning-based quantification of steatosis, inflammation, ballooning, and fibrosis in biopsies from patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2020;18(9):2081–2090.e9.
53 Destrempes F, Gesnik M, Chayer B, et al. Quantitative ultrasound, elastography, and machine learning for assessment of steatosis, inflammation, and fibrosis in chronic liver disease. *PLoS One*. 2022;17(1):e0262291.
54 Zhou JY, Song LW, Yuan R, Lu XP, Wang GQ. Prediction of hepatic inflammation in chronic hepatitis B patients with a random forest-backward feature elimination algorithm. *World J Gastroenterol*. 2021;27(21):2910–2920.
55 Sterling RK, Lissen E, Clumeck N, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*. 2006;43(6):1317–1325.
56 Zeng DW, Liu YR, Zhang JM, et al. Serum ceruloplasmin levels correlate negatively with liver fibrosis in males with chronic hepatitis B: a new noninvasive model for predicting liver fibrosis in HBV-related liver disease. *PLoS One*. 2013;8(10):e77942.
57 Yao M, Wang L, You H, et al. Serum GP73 combined AST and GGT reflects moderate to severe liver inflammation in chronic hepatitis B. *Clin Chim Acta*. 2019;493:92–97.
58 Peck-Radosavljevic M. Thrombocytopenia in chronic liver disease. *Liver Int*. 2017;37(6):778–793.
59 Olson SR, Koprowski S, Hum J, McCarty OJT, DeLoughery TG, Shatzel JJ. Chronic liver disease, thrombocytopenia and procedural

bleeding risk; are novel thrombopoietin mimetics the solution? *Platelets*. 2019;30(6):796–798.
60 Bernardi M, Angeli P, Claria J, et al. Albumin in decompensated cirrhosis: new concepts and perspectives. *Gut*. 2020;69(6):1127–1138.
61 Tripodi A, Caldwell SH, Hoffman M, Trotter JF, Sanyal AJ. Review article: the prothrombin time test as a measure of bleeding risk and prognosis in liver disease. *Aliment Pharmacol Ther*. 2007;26(2):141–148.
62 Angulo P, Hui JM, Marchesini G, et al. The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology*. 2007;45(4):846–854.
63 Yuan T, Yang T, Chen H, et al. New insights into oxidative stress and inflammation during diabetes mellitus-accelerated atherosclerosis. *Redox Biol*. 2019;20:247–260.
64 Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. *Nat Rev Gastroenterol Hepatol*. 2016;14(1):32–42.