



OPEN

Assessment of compensated advanced chronic liver disease based on serum bile acids in chronic hepatitis B patients

Fei Chen¹, Yaning Yao¹, Zhen Li², Long Deng¹ & Ruiling He²✉

Patients with chronic liver disease progressed to compensated advanced chronic liver disease (cACLD), the risk of liver-related decompensation increased significantly. This study aimed to develop prediction model based on individual bile acid (BA) profiles to identify cACLD. This study prospectively recruited 159 patients with hepatitis B virus (HBV) infection and 60 healthy volunteers undergoing liver stiffness measurement (LSM). With the value of LSM, patients were categorized as three groups: F1 [$LSM \leq 7.0$ kilopascals (kPa)], F2 ($7.1 < LSM \leq 8.0$ kPa), and cACLD group ($LSM \geq 8.1$ kPa). Random forest (RF) and support vector machine (SVM) were applied to develop two classification models to distinguish patients with different degrees of fibrosis. The content of individual BA in the serum increased significantly with the degree of fibrosis, especially glycine-conjugated BA and taurine-conjugated BA. The Marco-Precise, Marco-Recall, and Marco-F1 score of the optimized RF model were all 0.82. For the optimized SVM model, corresponding score were 0.86, 0.84, and 0.85, respectively. RF and SVM models were applied to identify individual BA features that successfully distinguish patients with cACLD caused by HBV. This study provides a new tool for identifying cACLD that can enable clinicians to better manage patients with chronic liver disease.

Abbreviations

TCA	Taurocholic acid
TCDCA	Taurochenodeoxycholic acid
TUDCA	Tauroursodeoxycholic acid
TDCA	Taurodeoxycholic acid
THDCA	Taurohyodeoxycholic acid
TLCA	Taurolithocholic acid
GUDCA	Glycoursodeoxycholic acid
GDCA	Glycodeoxycholic acid
GCDCA	Glycochenodeoxycholic acid
GLCA	Glycolithocholic acid
GCA	Glycocholic acid
CA	Cholic acid
UDCA	Ursodeoxycholic acid
DCA	Deoxycholic acid
CDCA	Chenodeoxycholic acid
LCA	Lithocholic acid

Hepatitis B virus (HBV) infection is a major public health problem worldwide. The World Health Organization estimated that 257 million people were living with chronic HBV infection in the worldwide, making these patients at a high risk of developing cirrhosis, hepatic decompensation, and hepatocellular carcinoma¹. China has the world's largest burden of HBV infection. It is estimated that there are about 70 million HBsAg carriers (5–6% prevalence) and 20–30 million people with chronic hepatitis B in China². Liver biopsy is the gold standard for the assessment of liver fibrosis and cirrhosis in patients with chronic liver disease. However, it is an invasive

¹Department of Ultrasound, The First Hospital of Lanzhou University, Lanzhou 730000, China. ²Department of Ultrasound, Donggang Branch, The First Hospital of Lanzhou University, Lanzhou 730000, China. ✉email: 1037039754@qq.com

procedure that may be complicated by pain, liver parenchyma injury and hemorrhage³. The current indication for liver biopsy is mainly to determine the cause of liver disease in selected cases, and not to stage fibrosis. Some non-invasive methods are commonly used to assess hepatic fibrosis and cirrhosis, with suggested cutoff values being applied to guide clinical decision making^{4–6}.

Since it is usually impossible to distinguish between severe fibrosis and cirrhosis in asymptomatic patients on clinical grounds, the Baveno VI consensus proposed the term “compensated advanced chronic liver disease (cACLD)” to better reflect that the spectrum of the two is a continuum⁷. A pragmatic definition of cACLD based on liver stiffness measurement (LSM) is aimed at stratifying the risk of decompensation, irrespective of histological stage or the ability of LSM to identify these stages⁸. Currently, LSM by transient elastography (TE) is the most widely used non-invasive method for assessing cACLD clinically. The renewing Baveno VII consensus suggested that patients with LSM-TE < 10 kilopascals (kPa) in the absence of other known clinical/imaging signs are sufficient to rule out cACLD⁸. Compared with TE, on the basis of conventional ultrasound images, two-dimensional shear wave elastography (2D-SWE) uses acoustic radiation force to generate shear waves, and can also form color coded images with different stiffness in the sampling frame, so as to effectively avoid non target structures and obtain more reliable tissue stiffness values^{9,10}. Many studies have shown that the performance of 2D-SWE is equivalent to or even better than that of TE in assessing liver fibrosis and cirrhosis^{11–13}.

Recently, bile acid (BA), endogenous compounds that undergo efficient enterohepatic circulation, have been confirmed as an important factor in the pathophysiology of the dynamic component of portal hypertension in both animal models and in humans^{14,15}. Therefore, the aim of our study was to develop prediction model based on individual BA profiles of serum samples to identify cACLD in patients with HBV infection.

Methods

Patient selection. This study was approved by The First Hospital of Lanzhou University Ethics Committee (approval number: LDYYLL2022-111). Data were collected from January to October 2022. Figure 1 shows the flow chat of the study population. We prospectively selected patients who met the following criteria: age between 18 and 75 years; positive serum hepatitis B surface antigen (HBsAg) for at least 6 months; accepted examination of LSM by 2D-SWE. Informed consent was obtained from all the study participants. Exclusion criteria were as follows: patients who used drugs against BA accumulation, such as ursodeoxycholic acid; hepatitis A, C, D, E viral infections; autoimmune liver disease; non-alcoholic fatty liver disease; alcoholic liver disease; drug-induced hepatitis; hepatocellular carcinoma; history of endocrine diseases and cholestasis. Cholestasis was defined as serum alkaline phosphatase (ALP) level > 1.5 ULN and gamma-glutamyl transpeptidase (GGT) level > 3 ULN. The ULNs of ALP and GGT were 125 U/L and 69 U/L, respectively.

In addition, 60 healthy volunteers were recruited as control with older than 18 years. They were all volunteers who had no substantial past medical history of chronic liver disease. They had no substantial alcohol intake

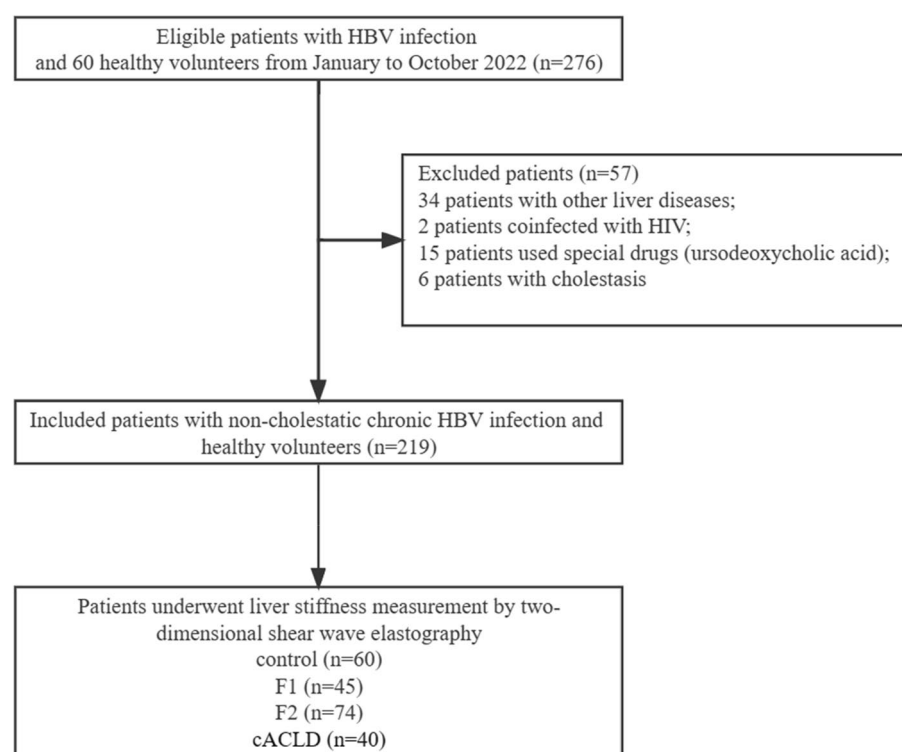


Figure 1. The flow chat of the study population. *HBV* hepatitis B virus, *HIV* human immunodeficiency virus, *cACLD* compensated advanced chronic liver disease.

(< 30 g/day for man, < 20 g/day for women) and had negative hepatitis B, hepatitis C, and human immunodeficiency virus.

All methods were performed in accordance with the relevant guidelines and regulations.

Laboratory tests. Blood samples were collected from all subjects in a fasting state in the early morning. Biochemical parameters, including serum aspartate aminotransferase (AST); alanine aminotransferase (ALT); ALP; GGT; total bilirubin (TBIL); direct bilirubin (DBIL); indirect bilirubin (IBIL) and total bile acids (TBA) were determined using a fully automated biochemical analyzer (Olympus AU400, Japan).

Two-dimensional shear wave elastography. All subjects underwent LSM based on 2D-SWE by a single professionally trained operator in fasting patients (fasted for > 6 h). An Aixplorer[®] ultrasound system (SuperSonic Imagine, SSI, France) with an abdominal 3.5-MHz curved array probe was used. LSM was measured in the right lobe using the 7th to 9th rib intercostal approach, with the right arm in maximum abduction. The region of interest, size 4 cm × 3 cm and fan-shaped, was placed in an area of parenchyma free of large vessels and bile ducts, avoiding noisy areas from rib shadowing. Start the 2D-SWE measurement, place the Q-BOX at least 1 cm and no more than 6 cm away from the liver capsule, and the diameter is not less than 1.5 cm. The value of LSM was depicted in kPa. The latest European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines were followed, the stiffness of the liver was measured five times in each case, and the median values were recorded⁹. The reliable value of LSM was defined as the stability index of image ≥ 80% and interquartile range/median ratio < 30%⁹.

The value of LSM ≤ 7.0 kPa was defined as F1, 7.1–8.0 kPa F2 and ≥ 8.1 kPa equal to or higher than F3. Patients with cACLD were considered to have the same or higher stage of liver fibrosis as F3¹⁶.

Sample preparation and high performance liquid chromatography (HPLC) analysis. We measured BA in the serum using previously described Liquid Chromatograph Mass Spectrometer method¹⁷. Simple protein precipitation using methanol was used to prepare the serum samples. Briefly, 200 µL methanol was added to 100 µL of serum spiked with 100 µL of ISs (*d*₄-chenodeoxycholic acid and Nor-desoxycholic acid). Subsequently, all the mixtures were vortexed for 1 min and centrifuged at 13,000×g for 5 min. The supernatant was aspirated for further analysis. An Agilent 1260 Infinity HPLC coupled with an Agilent 6460 triple-quadrupole mass spectrometer equipped with an electrospray ionization interface was used for the analysis of serum. Chromatographic resolution was performed on an Agilent HC-C18 column (4.6 mm × 250 mm, 5-µm particles), guarded by an Agilent Eclipse XDB-C18 4.6 mm × 12.5 mm analytical guard column (Agilent Technologies, USA). The mobile phase consisted of methanol (solvent B) and 7.5 mM ammonium acetate containing 0.1% ammonium hydroxide (solvent A, deionized water), pH 7.5, at a total flow of 1 mL/min, and post column splitting (1:4) was applied to give optimal interface flow rates (0.2 mL/min) for MS detection.

Principal component analysis (PCA). Principal component analysis (PCA) was performed using the prcomp (version 4.0.2) package to visualize the distribution of individual BA in different classes.

Random forest (RF). RF was introduced as a classifier owing to its attractive characteristics, including the need for few tunable parameters, automatic handling of missing data, and insensitivity to overfitting¹⁸. By using all the descriptors in the training set to build an RF classification model based on the cross-validation method, the importance of each descriptor with respect to prediction ability was determined. Subsequently, the order of importance for all descriptors was obtained. The resulting model was implemented in the statistical language R based on STATISTICA 10.0 with the default settings. The BA profiles of the control, F1, F2 and cACLD group were randomly divided into training and test sets at a 4:1 ratio, respectively 175 and 44 patients. There are two important parameters, ntree (the number of trees) and mtry (the number of features to split on each node), that must be optimized. In this study, to obtain the optimal model, the value of ntree was tuned from 1 to 219 with a step of 100. Meanwhile, the value of mtry was tuned from 1 to 50 with a step of 1 in each tuning step of ntree.

Support vector machine (SVM). The SVM method is a novel small-sample learning method, which can be used to deal with highly nonlinear regression and classification problems¹⁹. In brief, it is a supervised learning method that predicts the corresponding category of the new training sample by learning the category of the known sample and judging the relationship between the sample and the category. Similarly, all subjects were randomly divided into training and test sets with the ratio of 4:1. SVM was developed in a training set of 175 patients and tested in a validation set of 44 patients.

Model validation. The performance of the classification models was evaluated using the following metrics: Marco-Precise, Marco-Recall, Marco-F1 score, total accuracy, and Kappa coefficient.

Statistical analysis. Continuous variables were reported as median with interquartile range or mean with standard deviation. Categorical data, presented as number and frequencies (%). Differences among groups were analyzed by one-way analysis of variance with Dunnett's multiple comparison test or Mann–Whitney test using SPSS 25.0.02 (IBM, New York, U.S.). The difference was considered statistically significant when *p* < 0.05. RF and SVM data acquisition and quantification were performed using STATISTICA 10.0.

Results

Characteristics of the study population. A total of 159 patients with HBV infection and 60 healthy volunteers from the First Hospital of Lanzhou University between January 2022 to October 2022 were included in the final analysis. The characteristics of the study population are summarized in Table 1. The mean age was (44.7 ± 11.6) years and 62% were males. F2 group was present in 74 patients, accounting for the largest proportion (33.8%). This was followed by control (27.4%, 60/219) and F1 group (20.5%, 45/219). cACLD was found in 18.3% (40/219) of patients. There was no significant difference in biochemical indexes between F1 and control group ($p > 0.05$). Patients with F2 and cACLD group had significantly elevated levels of AST, TBIL, DBIL, IBIL, and TBA compared to those in the control group ($p < 0.05$). In addition, the subjects with cACLD exhibited notable increase in ALT, ALP, and GGT comparing with the control ($p < 0.05$).

PCA. PCA was performed to visualize the distributions of individual BA profiles in different degrees of liver fibrosis. As shown in Fig. 2, the BA of cACLD (pink) patients yielded higher PC1, PC2 and PC3 values, and only a few outliers were evident between control, F1, and F2. Therefore, the PCA method yielded a clear separation of cACLD patients from chronic liver disease caused by HBV using the individual BA.

	Control (n = 60)	F1 (n = 45)	F2 (n = 74)	Cacld (n = 40)
Age (years), mean (SD)	44 ± 11	41 ± 10	46 ± 11	48 ± 10
Male, n (%)	66.7%	73.3%	63.0%	60.0%
AST (U/L), median (IQR)	26.0 (11.8)	27.0 (9.2)	33.3 (18.6)**	43.0 (29)**
ALT (U/L), median (IQR)	25.5 (17.3)	29.5 (17.8)	32.9 (21.8)	32.3 (31.1)
ALP (U/L), mean (SD)	90.1 ± 25.3	86.5 ± 27.4	93.9 ± 27.5	105.2 ± 30.6*
GGT (U/L), mean (SD)	30.9 ± 12.6	28.8 ± 17.8	32.9 ± 18.8**	40.4 ± 27.1**
TBIL (μmol/L), median (IQR)	15.5 (8.2)	16.9 (6.7)	17.9 (9.8)	23.6 (15.6)**
DBIL (μmol/L), median (IQR)	7.1 (3.0)	7.8 (2.8)	8.4 (3.4)	11.0 (5.2)**
IBIL (μmol/L), median (IQR)	8.8 (4.6)	9.6 (4.3)	10.6 (5.3)	13.4 (5.8)**
TBA, mean (SD)	11.9 ± 8.7	15.5 ± 5.7	27.4 ± 15.4*	48.7 ± 17.2**

Table 1. Basic characteristics of the included patients. *SD* standard deviation, *IQR* interquartile range, *cACLD* compensated advanced chronic liver disease, *AST* aspartate aminotransferase, *ALT* alanine aminotransferase, *ALP* alkaline phosphatase, *GGT* γ-glutamyl transpeptidase, *TBIL* total bilirubin, *DBIL* direct bilirubin, *IBIL* indirect bilirubin, *TBA* total bile acids. * $p < 0.05$, ** $p < 0.01$ when compared to patients with control group.

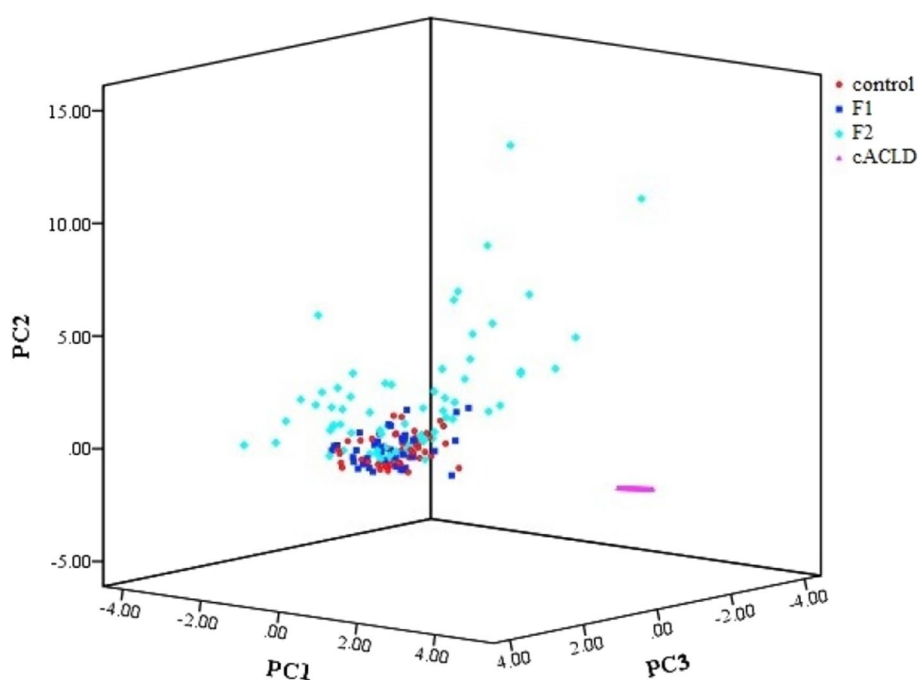


Figure 2. Principal component analysis in the control, F1, F2 and cACLD group. cACLD, compensated advanced chronic liver disease. Control (red); F1 (blue); F2 (green); cACLD (pink).

The alteration of serum bile acids in different liver fibrosis stage. Compared to control, the subjects with F1, F2, and cACLD exhibited an increase in total serum primary BA, while the proportion of total secondary BA decreased (Fig. 3a). In addition, patients with F1, F2, and cACLD exhibited significant increases in glycine-conjugated BA and taurine-conjugated BA compared with the control (Fig. 3b). The heat map displays the spectrum of the bile acid profiles across different fibrosis stage (Fig. 3c). Compared to control, there was a significant increase in the percentage of glycine-conjugated BA in F1 (2%), F2 (9%), and cACLD patients (14%). Taurine-conjugated BA exhibited a higher percentage in F1 (1.1%) and F2 patients (2.2%) and a significantly higher percentage in cACLD patients (12%). Moreover, the percent of unconjugated BA was 36% for the control, 33% for F2, 24% for F3, and 10% for cACLD group.

The change of serum individual bile acids in different liver fibrosis stage. Compared to the control, the sum of unconjugated, glycine-conjugated, and taurine-conjugated BA contents in the serum was significantly increased in F1, F2, and cACLD patients (Fig. 4a). Compared to the control, F1 patients showed a significant increase in the content of CDCA ($p < 0.001$), GCA ($p < 0.05$), and GCDCA ($p < 0.01$) (Fig. 4b, c), but there was no significant difference in other individual BA. Unconjugated BA such as CA, CDCA, DCA, LCA, and UDCA in the serum of patients with F2 were significantly increased ($p < 0.001$), and glycine-conjugated BA such as GCA, GCDCA, GDCA, GLCA, and GUDCA were also significantly increased ($p < 0.001$) in those with F2 (Fig. 4c). Furthermore, TCDCA, THDCA, TLCA, and TUDCA were significantly increased ($p < 0.001$) in patients with F2 (Fig. 4d). All individual BA were significantly increased in the serum of cACLD patients, except for TDCA.

Classification performance of RF and SVM. The number of decision trees was set to 20, and the maximum tree size was set to 15 based on the results of the parameter tuning tests. A summary of the RF response of classification is shown in Fig. 5a. For the RF method, a regression algorithm based on importance ranking was used to extract the features of the impact factors, select the optimal feature variable set, and achieve the goal of dimension reduction. The SVM model was optimized using tenfold cross-validation, and the selected samples were trained and predicted using the SVM model. The Macro-Precise, Macro-Recall, Macro-F1 score, kappa coefficient, and accuracy of the optimized RF model were 0.82, 0.82, 0.82, 0.74, and 0.81, respectively. The importance of 16 individual BA features was calculated using the RF method. The importance status is shown in Fig. 5b, showing that all the assigned individual BA features had the capacity to discriminate between different liver fibrosis stage. The CA, CDCA, DCA, GCA, and GCDCA features gained the highest importance. The Macro-Precise, Macro-Recall, Macro-F1 score, kappa coefficient, and accuracy of the optimized SVM model were 0.86, 0.84, 0.85, 0.76, and 0.82, respectively. The performances of the built RF and SVM models for identifying different liver fibrosis stage are shown in Table 2.

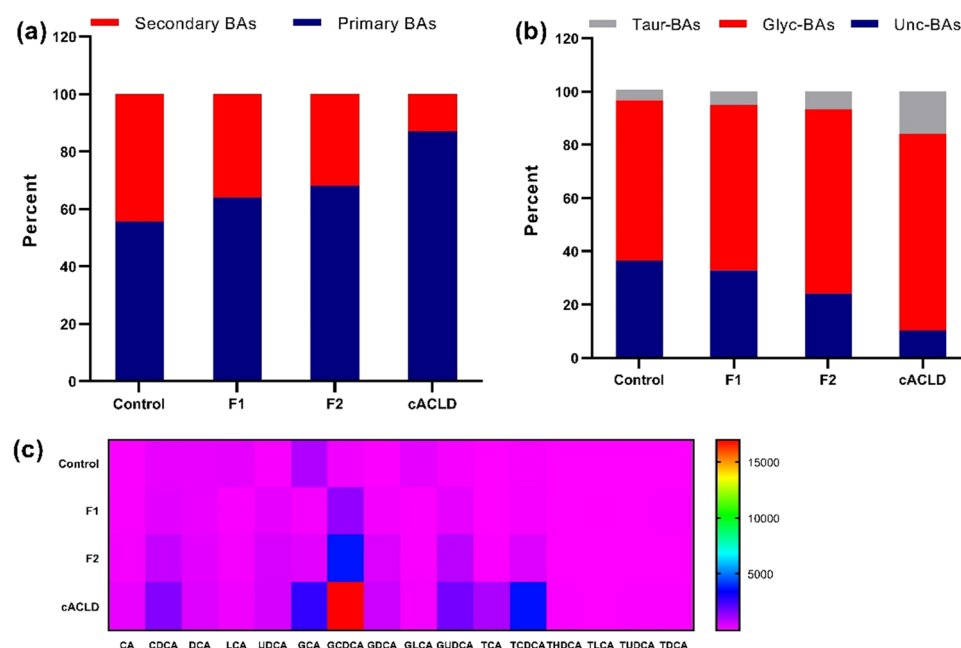


Figure 3. The alteration of serum bile acids in the control, F1, F2 and cACLD group. (a) Stack bar plot representing proportion of unconjugated (Unc-BAs), glycine conjugated (Glyc-BAs) and taurine conjugated (Taur-BAs) bile acids. (b) Stack bar plot representing proportion of primary bile acids and secondary bile acids; (c) heat map display the spectrum of bile acids profile across different fibrosis stage.

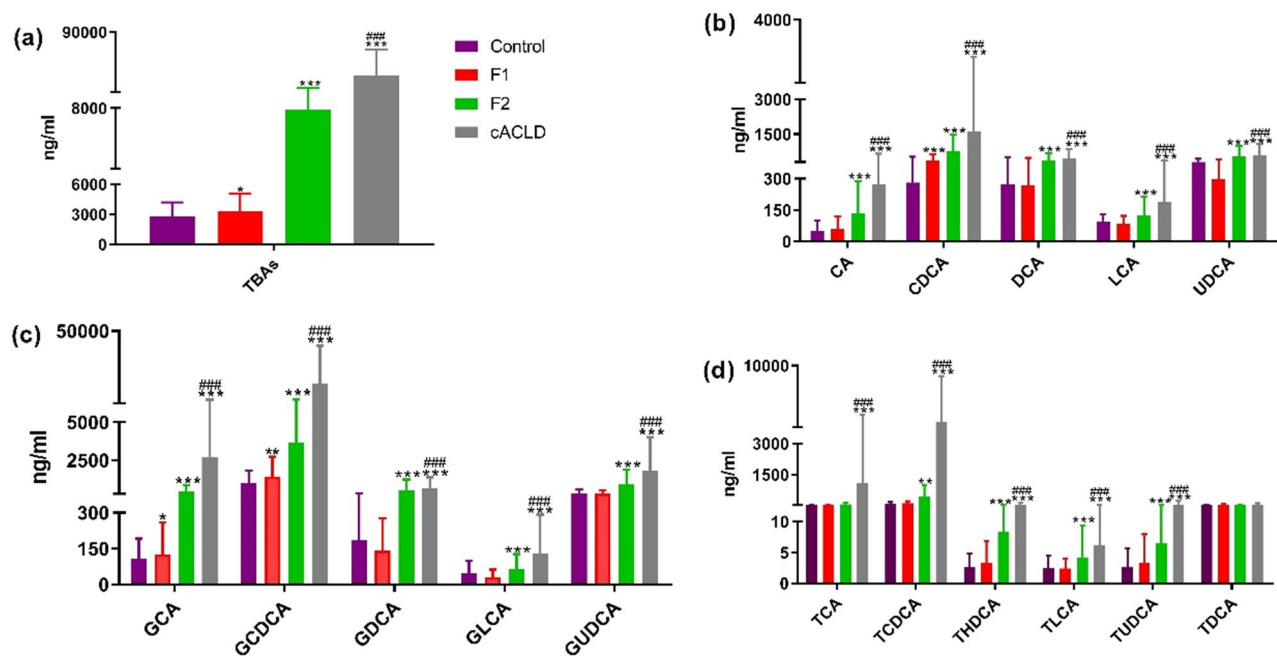


Figure 4. The change of TBAs, unconjugated, glycine-conjugated and taurine-conjugated bile acids of (mean \pm SD) in the control, F1, F2 and cACLD group. TBAs: the sum of unconjugated, glycine-conjugated and taurine-conjugated bile acids content in the serum. cACLD compensated advanced chronic liver disease. Compared with the control, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Compared with F2 group, ### $p < 0.001$.

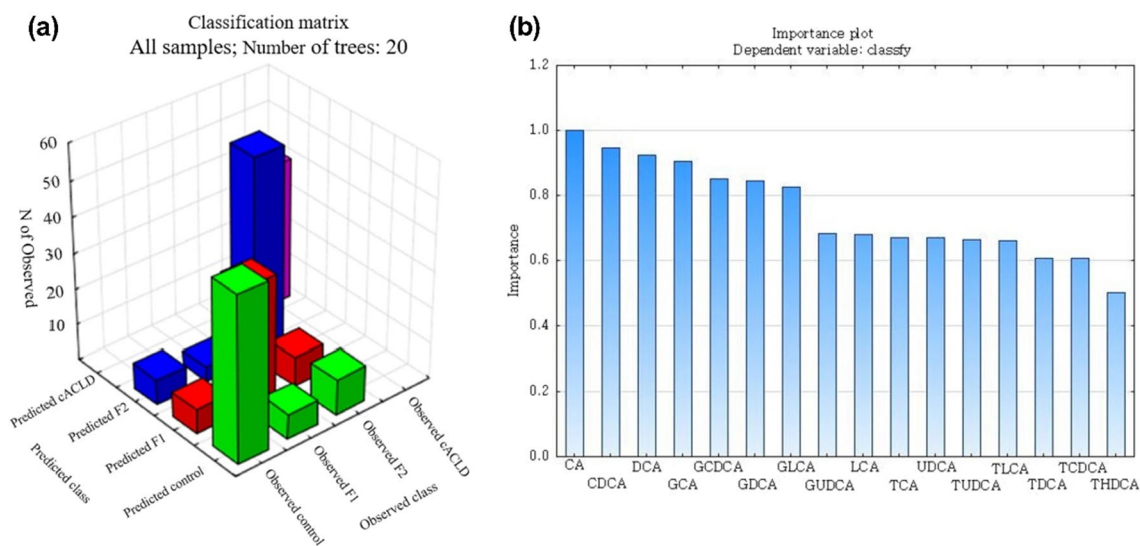


Figure 5. The performance of build Random Forest (RF) model. (a) Classification matrix of all samples, number of trees: 20; (b) importance plot of individual bile acids.

Model	Marco-Precise	Marco-Recall	Marco-F1 score	Kappa coefficient	Accuracy
RF	0.82	0.82	0.82	0.74	0.81
SVM	0.86	0.84	0.85	0.76	0.82

Table 2. Performance of the built random forest (RF) and support vector machine (SVM) models for identification different liver fibrosis stage. Marco-F1 score: $2 \times \text{Marco-Precise} \times \text{Marco-Recall} / (\text{Marco-Precise} + \text{Marco-Recall})$.

Discussion

The degree of liver fibrosis in patients with chronic liver disease predicts the likelihood of developing liver-related morbidity and death²⁰. When these patients progressed to cACLD, the risk of liver-related decompensation events and death increased significantly. Thus, assessment of cACLD is an essential part of the evaluation of chronic liver disease patients in order to prognosticate, stratify therapeutic and surveillance strategies⁸. Research showed that elevated serum BA concentrations have been shown to be a more sensitive test for the detection of liver cirrhosis than conventional liver function tests¹⁴. Therefore, this study developed two different models (RF and SVM) based on individual BA profiles of serum samples to recognize cACLD in patients with HBV infection.

In the present study we analysed the relationship between serum concentrations of individual BA and the degree of liver fibrosis in patients with chronic liver diseases. We found that PCA method using the individual BA can distinguish cACLD from chronic liver disease patients. Furthermore, the total serum primary BA increased while the proportion of total secondary BA decreased in subjects with the controlled, F1, F2, and cACLD group. Meanwhile, the glycine-conjugated BA and taurine-conjugated BA increased, while the unconjugated BA decreased. More importantly, conjugated BAs, including GCDCA and TCDCA, increased significantly in patients with cACLD. It seems to be consistent with the research of Žižalová et al., which found that GCDCA and TCDCA are significantly related to portal pressure in patients with cirrhosis¹⁴. Oehler showed that the synthesis of primary BA, CA and CDCA, was significantly increased by the strong induction of hCYP7A1 (the rate-limiting enzyme converting cholesterol to BA) in human liver chimeric mice infected with HBV²¹. Although the relative excess of CDCA over CA derivatives seem to be a common feature of liver cirrhosis as well as non-alcoholic fatty liver disease, the mechanism behind this remains somewhat enigmatic²². The Žižalová K's study aimed to identify clinically significant portal hypertension in patients with cirrhosis through BA, while our study intended to identify cACLD in patients with chronic liver disease, which represent an important point for timely intervention to prevent further progression.

Furthermore, the RF model and SVM model derived from our current BA analysis showed good separation between different fibrosis stage, highlighting the diagnostic potential of this noninvasive analytical approach. Five serum BA, CA, CDCA, DCA, GCA, and GCDCA, gained the highest importance, suggesting that unconjugated and glycine-conjugated BA may be indicators of liver dysfunction in chronic hepatitis.

The limitation of the current study was that we were unable to compare noninvasive biomarkers with the gold standard of liver biopsy. In addition, the subjects included in present study are all patients with HBV. The results need to be verified in patients with other causes, such as hepatitis C virus, alcoholic, non-alcoholic fatty liver disease, autoimmune liver disease, etc.

In conclusion, RF and SVM models were applied to identify individual BA features that successfully distinguish patients with cACLD caused by HBV. This study provides a new tool for identifying cACLD patients that can enable clinicians to better manage patients with chronic liver disease.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article.

Received: 26 March 2023; Accepted: 2 August 2023

Published online: 08 August 2023

References

1. WHO. *Global Hepatitis Report, 2017*[EB/OL], [2019-11-06] (2019).
2. Liu, J., Liang, W., Jing, W. & Liu, M. Countdown to 2030: Eliminating hepatitis B disease, China. *Bull. World Health Organ.* **97**, 230–238 (2019).
3. Loomba, R. & Adams, L. A. Advances in non-invasive assessment of hepatic fibrosis. *Gut* **69**, 1343–1352 (2020).
4. Munteanu, M. et al. Long-term prognostic value of the FibroTest in patients with non-alcoholic fatty liver disease, compared to chronic hepatitis C, B, and alcoholic liver disease. *Aliment. Pharmacol. Ther.* **48**, 1117–1127 (2018).
5. Thiele, M. et al. Accuracy of the enhanced liver fibrosis test vs FibroTest, elastography, and indirect markers in detection of advanced fibrosis in patients with alcoholic liver disease. *Gastroenterology* **154**, 1369–1379 (2018).
6. Eddowes, P. J. et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology* **156**, 1717–1730 (2019).
7. de Franchis, R. & Baveno VI, F. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. *J. Hepatol.* **63**, 743–752 (2015).
8. de Franchis, R., Bosch, J., Garcia-Tsao, G., Reiberger, T., Ripoll, C. & Baveno VII, F. Baveno VII—Renewing consensus in portal hypertension. *J. Hepatol.* **76**, 959–974 (2022).
9. Dietrich, C. F. et al. EFSUMB guidelines and recommendations on the clinical use of liver ultrasound elastography, update 2017 (long version). *Ultraschall Med.* **38**, e16–e47 (2017).
10. Ferraioli, G. et al. WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 3: Liver. *Ultrasound Med. Biol.* **41**, 1161–1179 (2015).
11. Gao, Y. et al. Liver fibrosis with two-dimensional US shear-wave elastography in participants with chronic hepatitis B: A prospective multicenter study. *Radiology* **289**, 407–415 (2018).
12. Cassinotto, C. et al. Agreement between 2-dimensional shear wave and transient elastography values for diagnosis of advanced chronic liver disease. *Clin. Gastroenterol. Hepatol.* **18**, 2971–2979.e3 (2020).
13. Thiele, M. et al. Transient and 2-dimensional shear-wave elastography provide comparable assessment of alcoholic liver fibrosis and cirrhosis. *Gastroenterology* **150**, 123–133 (2016).
14. Žižalová, K. et al. Serum concentration of taurochenodeoxycholic acid predicts clinically significant portal hypertension. *Liver Int.* **43**, 888–895 (2023).
15. Königshofer, P. et al. Distinct structural and dynamic components of portal hypertension in different animal models and human liver disease etiologies. *Hepatology* **75**, 610–622 (2022).
16. Herrmann, E. et al. Assessment of biopsy-proven liver fibrosis by two-dimensional shear wave elastography: An individual patient data-based meta-analysis. *Hepatology* **67**, 260–272 (2018).

17. Jin, Y. W. *et al.* Development of a LC-MS/MS method for simultaneous determination of bile acids and their conjugates in hepatocytes, tissue and fluids in rat. *Curr. Pharm. Anal.* **14**, 331–341 (2018).
18. Zhou, J. Y., Song, L. W., Yuan, R., Lu, X. P. & Wang, G. Q. Prediction of hepatic inflammation in chronic hepatitis B patients with a random forest-backward feature elimination algorithm. *World J. Gastroenterol.* **27**, 2910–2920 (2021).
19. Huang, S. *et al.* Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics* **15**, 41–51 (2018).
20. Huang, Y. *et al.* Image analysis of liver biopsy samples measures fibrosis and predicts clinical outcome. *J. Hepatol.* **61**, 22–27 (2014).
21. Oehler, N. *et al.* Binding of hepatitis B virus to its cellular receptor alters the expression profile of genes of bile acid metabolism. *Hepatology* **60**, 1483–1493 (2014).
22. Caussy, C. *et al.* Serum bile acid patterns are associated with the presence of NAFLD in twins, and dose-dependent changes with increase in fibrosis stage in patients with biopsy-proven NAFLD. *Aliment. Pharmacol. Ther.* **49**, 183–193 (2019).

Acknowledgements

We appreciate all the staff of the Department of Infectious Diseases, The first Hospital of Lanzhou University, for their joint effort and sincere cooperation.

Author contributions

F.C. and R.H. conceived and designed the project; F.C., Y.Y., Z.L. and L.D. collected the data; F.C., Y.Y., Z.L., L.D. and R.H. drafted and revised the paper, supervised the analyses, and suggested revisions of the paper. All the authors have read and approved the final manuscript.

Funding

This work was supported by the Foundation of the First Hospital of Lanzhou University (grant no. ldyyn2021-62).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023