Retrospective Study

Serum N-glycan markers for diagnosing liver fibrosis induced by hepatitis B virus

Xi Cao, Qing-Hua Shang, Xiao-Ling Chi, Wei Zhang, Huan-Ming Xiao, Mi-Mi Sun, Gang Chen, Yong An, Chun-Lei Lv, Lin Wang, Yue-Min Nan, Cui-Ying Chen, Zong-Nan Tan, Xue-En Liu, Hui Zhuang

ORCID number: Xi Cao (0000-0002-9297-0439); Qing-Hua Shang (0000-0001-5439-7921); Xiao-Ling Chi (0000-0003-3193-1941); Wei Zhang (0000-0003-1520-7568); Huan-Ming Xiao (0000-0002-8739-0720); Mi-Mi Sun (0000-0002-0359-0240); Gang Chen (0000-0001-5909-541X); Yong An (0000-0002-2590-0916); Chun-Lei Lv (0000-0003-0664-6414); Lin Wang (0000-0003-0835-7410); Yue-Min Nan (0000-0003-4192-099X); Cui-Ying Chen (0000-0002-2136-5723); Zong-Nan Tan (0000-0003-4056-9886); Xue-En Liu (0000-0002-6574-4989); Hui Zhuang (0000-0001-9119-6325).

Author contributions: All authors performed the research, Liu XE and Zhuang H designed the study. Cao X, Wang L and Tan ZN did the experiments. Cao X analyzed data and drafted the manuscript. Shang QH, Chi XL, Zhang W, Xiao HM, Sun MM, Chen G, An Y, Lv CL and Nan YM were responsible for serum sample collection and clinical data acquisition. Liu XE, Zhuang H and Chen CY critically reviewed the manuscript for important intellectual content. All authors agreed to be accountable for all aspects of the work and approved the final version of the manuscript.

Supported by: Major Science and Technology Special Project of China Thirteenth Five-Year Plan, No. 2018ZX10220-001-019; Guangxi Key Laboratory for the Prevention and Control of Viral Hepatitis, No. GXXDCKL201901.

Abstract

BACKGROUND

Hepatitis B virus (HBV) infection is the primary cause of hepatitis with chronic HBV infection, which may develop into liver fibrosis, cirrhosis and hepatocellular carcinoma. Detection of early-stage fibrosis related to HBV infection is of great clinical significance to block the progression of liver lesion. Direct liver biopsy is regarded as the gold standard to detect and assess fibrosis; however, this method is invasive and prone to clinical sampling error. In order to address these issues, we attempted to find more convenient and effective serum markers for detecting HBV-induced early-stage liver fibrosis.

AIM

To investigate serum N-glycan profiling related to HBV-induced liver fibrosis and verify multiparameter diagnostic models related to serum N-glycan changes.

METHODS

N-glycan profiles from the sera of 432 HBV-infected patients with liver fibrosis were analyzed. Significant changed N-glycan levels (peaks) ($P < 0.05$) in different
fibrosis stages were selected in the modeling group, and multiparameter diagnostic models were established based on changed N-glycan levels by logistic regression analysis. The receiver operating characteristic (ROC) curve analysis was performed to evaluate diagnostic efficacy of N-glycans models. These models were then compared with the aspartate aminotransferase to platelet ratio index (APRI), fibrosis index based on the four factors (FIB-4), glutamyltranspeptidase platelet albumin index (5 index), GlycoCirrho-test, and GlycoFibro-test. Furthermore, we combined multiparameter diagnostic models with alanine aminotransferase (ALT) and platelet (PLT) tests and compared their diagnostic power. In addition, the diagnostic accuracy of N-glycan models was also verified in the validation group of patients.

RESULTS

Multiparameter diagnostic models constructed based on N-glycan peak 1, 3, 4 and 8 could distinguish between different stages of liver fibrosis. The area under ROC curves (AUROCs) of Model A and Model B were 0.890 and 0.752, respectively differentiating fibrosis F0-F1 from F2-F4, and F0-F2 from F3-F4, and surpassing other serum panels. However, AUROC (0.747) in Model C used for the diagnosis of F4 from F0-F3 was lower than AUROC (0.795) in FIB-4. In combination with ALT and PLT, the multiparameter models showed better diagnostic power (AUROC = 0.912, 0.829, 0.885, respectively) when compared with other models. In the validation group, the AUROCs of the three combined models (0.929, 0.858, and 0.867, respectively) were still satisfactory. We also applied the combined models to distinguish adjacent fibrosis stages of 432 patients (F0-F1/F2/F3/F4), and the AUROCs were 0.917, 0.720 and 0.785.

CONCLUSION

Multiparameter models based on serum N-glycans are effective supplementary markers to distinguish between adjacent fibrosis stages of patients caused by HBV, especially in combination with ALT and PLT.

Key words: Chronic hepatitis B; Liver fibrosis; N-glycan; Multiparameter diagnostic models; Receiver operating characteristic curve analysis; Diagnostic power

INTRODUCTION

Hepatitis B virus (HBV) infection is the primary cause of hepatitis in humans, with 257 million individuals worldwide having chronic hepatitis B (CHB) infection in the year of 2015[1]. Chronic HBV infection correlates with an increased risk for the...
development of cirrhosis and hepatocellular carcinoma (HCC). Most HCC patients experience a three-stage process of "hepatitis - liver fibrosis/cirrhosis - HCC", in which the end-stage liver disease eventually leads to death[4]. Therefore, accurate detection of the early-stage fibrosis is of great clinical significance to implement intervention and block the progression of liver disorders.

Efforts have been made to search for diagnostic biomarkers of liver fibrosis for many years. Various imaging techniques, such as FibroScan and magnetic resonance imaging have been used with higher accuracy in the diagnosis of advanced fibrosis and cirrhosis when compared to mild fibrosis[3], which linked to obesity[4]. Although liver biopsy was the gold standard to assess fibrosis, the invasive nature of the procedure and its potential sampling error among different patients limits its application. Therefore, it is meaningful to find more definitive serum markers for detecting early-stage liver fibrosis. In recent years, diagnostic models based on serum markers have been established as valuable supplementary tools to diagnose and monitor liver fibrosis. Currently, non-invasive serological models including aspartate aminotransferase to platelet ratio index (APRI), fibrosis index based on the four factors (FIB-4), and glutamyltranspeptidase platelet albumin index (S index)[5-7] have been widely used in clinical practice. However, the sensitivity and specificity of these models for diagnosing mild fibrosis cannot completely meet clinical requirements.

Protein glycosylation, the most common post-translational modification, has become a new focus in the diagnosis of liver diseases. Most N-linked glycoproteins in serum are synthesized by the liver and B-lymphocytes, therefore any changes in total serum N-glycan could reflect an alteration of liver or B-lymphocyte physiology. In addition, sugar chains of glycoproteins are important for maintaining the ordered behavior of differentiated cells in multicellular organisms. Moreover, changes in sugar chains resulted in the molecular basis of abnormalities, such as alterations in the N-linked sugar chains, which were indeed found in various tumors[8,9]. Callewaert et al[10] developed the "DNA sequencer-assisted fluorophore-assisted carbohydrate electrophoresis" (DSA-FACE) for the systemic and rapid analysis of N-glycan profiles in serum and other body fluids. By using of DSA-FACE, a cirrhosis biomarker (GlycoCirrhoTest) and a fibrosis biomarker (GlycoFibroTest) were utilized in patients with chronic HCV infection[11,12]. However, we found that serum altered N-glycan in HBV patients were quite different from HCV patients in the initial study. In our current study, we enlarged clinical samples with HBV-related hepatic fibrosis and attempted to establish new diagnostic models which can be applied in distinguishing different stages of liver fibrosis. We also evaluated the clinical value of the diagnostic models in the validation group.

**MATERIALS AND METHODS**

**Patients**

Patients were recruited from No. 88 Hospital of Chinese People’s Liberation Army in Tai'an, China, and from Guangdong Provincial Hospital of Chinese Medicine in Guangdong, China between 2011 and 2017. A total of 432 HBV-infected patients with chronic liver diseases were recruited in the retrospective study. All the enrolled patients were positive for hepatitis B surface antigen for at least six months without any anti-viral treatment and were histopathologically assessed by liver biopsy. Histological specimens were > 15 mm in length and contained at least six portal tracks. Patients with sufficient serum samples collected at the day of liver biopsy were qualified for serum N-glycan profiling analysis. Patients with hepatitis A, hepatitis C, hepatitis D, hepatitis E, autoimmune liver disease, drug-induced liver injury, alcoholic liver disease, obstructive jaundice, or liver cancer were excluded from the study. The study was approved by the Institutional Review Board of Peking University Health Science Centre.

**Clinical laboratory tests**

Clinical data of the patients are summarized in Tables 1 and 2. Serum alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ-glutamyl-transpeptidase (GGT), alkaline phosphatase, total serum protein, albumin, total bilirubin, indirect bilirubin, blood platelet (PLT), alpha-fetoprotein, serum HBV markers including hepatitis B surface antigen, anti-HBs, HBeAg, anti-HBe, anti-HBc and HBV DNA levels were tested at hospital clinics.

**Liver fibrosis stage**

All patients underwent a liver biopsy procedure and results of the histopathological assessment were obtained. Liver fibrosis was determined histologically by...
Table 1  Characteristics of 432 hepatitis B virus-infected patients in different fibrosis stages diagnosed by liver biopsy (mean ± SD)

| Variables           | Fibrosis stage | F0-F1 | F2      | F3      | F4      | P value |
|---------------------|----------------|-------|---------|---------|---------|---------|
| Cases, n            |                | 95    | 126     | 103     | 108     | /       |
| Male, n (%)         |                | 62 (65.26) | 88 (69.84) | 82 (79.61) | 89 (82.41) | /       |
| Age (yr)            |                | 37.91 ± 10.71 | 36.52 ± 11.11 | 41.71 ± 10.01 | 46.63 ± 9.95 | < 0.001 |
| HBV DNA (IU/mL)     |                | 1.26 × 10^8 ± 3.64 × 10^6 | 1.33 × 10^8 ± 4.01 × 10^6 | 6.39 × 10^7 ± 1.62 × 10^8 | 5.63 × 10^7 ± 1.56 × 10^8 | 0.229   |
| AFP (ng/mL)         |                | 3.20 ± 3.01 | 19.37 ± 100.27 | 12.56 ± 23.25 | 40.98 ± 160.29 | 0.118   |
| ALT (IU/L)          |                | 42.58 ± 61.49 | 125.21 ± 211.29 | 263.82 ± 421.57 | 115.84 ± 165.90 | < 0.001 |
| AST (IU/L)          |                | 30.43 ± 26.74 | 85.65 ± 50.37 | 185.61 ± 29.25 | 92.18 ± 127.81 | < 0.001 |
| GGT (IU/L)          |                | 75.63 ± 21.27 | 198.36 ± 55.72 | 185.36 ± 58.29 | 134.95 ± 46.41 | < 0.001 |
| ALP (IU/L)          |                | 106.91 ± 31.14 | 198.36 ± 55.72 | 185.36 ± 58.29 | 134.95 ± 46.41 | < 0.001 |
| PLT (10^9/L)        |                | 206.91 ± 51.14 | 198.36 ± 55.72 | 185.36 ± 58.29 | 134.95 ± 46.41 | < 0.001 |
| Total bilirubin (μmol/L) |          | 13.09 ± 5.30 | 18.34 ± 16.99 | 29.93 ± 31.44 | 22.67 ± 12.60 | < 0.001 |
| Indirect bilirubin (μmol/L) |        | 7.78 ± 4.31 | 8.40 ± 9.40 | 13.46 ± 17.05 | 10.01 ± 7.94 | 0.003 |
| Total protein (g/L) |                | 73.46 ± 4.31 | 73.61 ± 10.53 | 72.48 ± 31.44 | 74.44 ± 5.39 | 0.320   |
| Albumin (g/L)       |                | 45.80 ± 3.30 | 45.44 ± 4.40 | 45.33 ± 6.51 | 44.39 ± 5.39 | 0.320   |

*p < 0.01.

Serum protein N-Glycome analysis

Callewaert et al. used DSA-FACE to profile sialidase-treated N-glycan from whole serum. We recently adapted DSA-FACE by capillary electrophoresis-based ABI 3500 Genetic Analyzer. N-glycan profiling data were analyzed with GeneMapper software version 4.1. We measured the heights of the 9 specific peaks (N-glycan) that were detected in all the samples and described them by normalizing their height to the sum of all peak heights. Researchers were blind to the histopathological results of patients when serum N-glycan analysis was performed.

Statistical analysis

Statistical analysis was performed with SPSS 18.0 for Windows software (SPSS, Chicago, IL). All quantitative variables were presented as means ± SD and compared with ANOVA analysis. Diagnostic models were constructed by stepwise logistic
Table 2 Inflammation grade and fibrosis stage of 432 hepatitis B virus-infected patients

| Fibrosis stage | Inflammation grade |
|---------------|--------------------|
|               | G0 | G1 | G2 | G3 | G4 | Total |
| F0-F1         | 8  | 71 | 16 | 0  | 0  | 95    |
| F2            | 0  | 31 | 83 | 12 | 0  | 126   |
| F3            | 0  | 12 | 40 | 47 | 4  | 103   |
| F4            | 0  | 11 | 38 | 55 | 4  | 108   |
| Total         | 8  | 125| 177|114 |8  |432    |

regression analysis. The diagnostic performance of a single marker or the diagnostic model was evaluated by the receiver operating characteristic (ROC) curves analysis. The area under ROC curve (AUROC) was used as an index of accuracy, with an AUROC value of 1.0 indicating high diagnostic accuracy. Using the maximum of Youden Index to best assess sensitivity and specificity, we calculated sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV). \( P < 0.05 \) was considered statistically significant.

RESULTS

N-glycome profiles in different fibrosis stages

N-glycome profiles in sera of 432 patients with HBV-related fibrosis were examined by DSA-FACE. Representative N-glycan profiling patterns of liver fibrosis F0-F4 are shown in Figure 1. The structures of 9 peaks had been characterized previously by Liu et al.\(^1\). We quantified each peak by normalizing its height to the sum of the heights of all peaks. The average relative abundances of N-glycan in different fibrosis stages are presented in Table 3. In addition, the changes in the abundance of each peak were compared between different fibrosis stages of patients in the modeling group, and results are shown in Table 3.

Construction and assessment of diagnostic models based on N-glycan markers

There were significant differences in abundances of peaks 1, 3, 4 and 8 between group F0-F1 and group F2-F4, group F0-F2 and group F3-F4, group F0-F3 and group F4. Based on the four peaks, three multivariate logistic regression models were constructed by logistic regression analysis. Model A = 0.237 × peak 1 + 0.495 × peak 3 - 1.155 × peak 4 - 0.054 × peak 8 + 4.632 for distinguishing F0-F1 from F2-F4; Model B = 0.129 × peak 1 + 0.175 × peak 3 - 0.983 × peak 4 - 0.119 × peak 8 + 4.576 for distinguishing F0-F2 from F3-F4; Model C = 0.231 × peak 1 - 0.060 × peak 3 - 0.838 × peak 4 - 0.160 × peak 8 + 3.340 for distinguishing F0-F3 from F4. The areas under ROC (AUROCs) of each model, and sensitivity, specificity, PPV, NPV at the defined cutoff value in the patients of modeling group are summarized in Table 4 and Supplementary Table 1.

In addition, based on the previous research, we adopted the log ratio of some N-glycome markers designated as GlycoFibroTest \([\log(\text{peak 2/peak 8})]\) and GlycoCirrhoTest \([\log(\text{peak 7/peak 8})]\) for diagnosing liver fibrosis and cirrhosis\(^1\). We analyzed the diagnostic power of these two tests in the modeling group by ROC curve analysis and compared them with Model A, Model B and Model C, respectively (Figure 2). The results showed that the AUROCs of Model A-Model C were significantly superior to the GlycoFibroTest and GlycoCirrhoTest.

Comparing N-glycome models with serological marker panels

Serum marker panels including S index, APRI and FIB-4 had been selected to diagnose liver fibrosis. We evaluated the diagnostic value of these panels in the modeling group. The fibrosis scores were calculated using the formulas previously reported \([\text{APRI} = \frac{\text{AST (IU/L)}}{100} \times \frac{\text{PLT (10^9/L)}}{\text{AST (IU/L)}}] ; \text{FIB-4} = [\text{age (years)} \times \text{AST (IU/L)}]/[\text{PLT} (10^9/L)] \times \text{ALT}^{1/2} (\text{IU/L})] ; \text{S index} = [1000 \times \text{GGT} (\text{IU/L})]/[\text{PLT} (10^9/L) \times \text{ALB}^2 (\text{g/L})]\). The sensitivity, specificity, PPV, NPV and AUROCs of APRI, FIB-4 and S index were analyzed and the results are listed in Table 4.

In the differentiation of fibrosis F0-F1 from F2-F4, the diagnostic power of Model A (AUROC = 0.890) was much higher than APRI (AUROC = 0.790), FIB-4 (AUROC = 0.599) and S index (AUROC = 0.701), respectively. In the differentiation of fibrosis F0-F2 from F3-F4, the diagnostic power of Model B (AUROC = 0.752) was also superior.
Cao X et al. Serum N-glycan markers for diagnosing LF induced by HBV
Figure 1 A typical desialylated N-glycan profile from the total serum protein is shown in panels. The structures of the N-glycan peaks are shown below the panels. Peak 1 indicates an agalacto core-α-1,6-fucosylated biantennary glycan (NGA2F), peak 2 indicates an agalacto core-α-1,6-fucosylated bisecting biantennary glycan (NGA2FB), peaks 3 and 4 indicate a single agalacto core-α-1,6-fucosylated biantennary glycan (NG1A2F), peak 5 indicates a bigalacto biantennary glycan (NA2), peak 6 indicates a bigalacto core-α-1,6-fucosylated biantennary glycan (NA2F), peak 7 indicates a bigalacto core-α-1,6-fucosylated bisecting biantennary glycan (NA2FB), peak 8 indicates a triantennary glycan (NA3), peak 9 indicates a branching α-1,3-fucosylated triantennary glycan (NA3Fb).

The diagnostic power of the N-glycome model demonstrated a slight decline with respect to diagnosing cirrhosis (F4) from fibrosis (F0-F3). FIB-4 (AUROC = 0.795) showed better diagnostic efficacy than Model C (AUROC = 0.747). The AUROC of Model A (0.890) was close to 0.9, and the AUROCs of Model B (0.752) and Model C (0.747) were below 0.8. We found that N-glycome diagnostic models combined with ALT and PLT showed the best diagnostic potency in the differentiation of different fibrosis stages. Three combined models were also constructed using logistic regression analysis: Model A + ALT + PLT = 5.816 × Model A + 0.004 × ALT - 0.006 × PLT - 2.130; Model B + ALT + PLT = 4.538 × Model B + 0.001 × ALT - 0.016 × PLT + 0.236; Model C + ALT + PLT = 5.943 × Model C - 0.003 × ALT -
Table 3. N-glycome profiling results in hepatitis B virus-infected patients with liver fibrosis F0-F4 and differences between different fibrosis stages in the modeling group (mean ± SD)

| Peak | F0-F1 Mean ± SD | F2 Mean ± SD | F3 Mean ± SD | F4 Mean ± SD | F0-F1 vs F2 Mean ± SD | F0-F2 vs F3 Mean ± SD | F0-F3 vs F4 Mean ± SD |
|------|-----------------|--------------|--------------|--------------|----------------------|----------------------|----------------------|
| 1    | 5.92 ± 1.46     | 7.62 ± 2.57  | 7.51 ± 2.09  | 8.68 ± 3.60  | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   |
| 2    | 1.95 ± 0.67     | 1.78 ± 0.76  | 1.60 ± 0.57  | 1.85 ± 0.76  | 0.161                | 0.235                | 0.312                |
| 3    | 4.57 ± 1.16     | 5.96 ± 1.43  | 6.04 ± 1.17  | 6.21 ± 1.66  | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   | 0.003<sup>b</sup>     |
| 4    | 7.34 ± 1.37     | 5.96 ± 1.15  | 5.54 ± 0.81  | 5.43 ± 0.81  | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   |
| 5    | 47.90 ± 3.77    | 46.93 ± 4.64 | 46.75 ± 4.59 | 46.27 ± 6.23 | 0.097                | 0.163                | 0.197                |
| 6    | 16.89 ± 2.92    | 17.66 ± 3.01 | 18.13 ± 2.89 | 17.30 ± 2.58 | 0.086                | 0.323                | 0.413                |
| 7    | 4.48 ± 1.08     | 4.53 ± 0.95  | 4.86 ± 1.22  | 4.90 ± 1.16  | 0.089                | 0.004<sup>b</sup>     | 0.058                |
| 8    | 8.07 ± 2.07     | 6.91 ± 1.90  | 6.74 ± 1.70  | 6.22 ± 2.16  | < 0.001<sup>c</sup>   | < 0.001<sup>c</sup>   | 0.001<sup>b</sup>     |
| 9    | 2.87 ± 1.27     | 2.65 ± 1.01  | 2.82 ± 1.12  | 3.12 ± 1.33  | 0.684                | 0.136                | 0.035<sup>c</sup>     |

<sup>a</sup><sup>P</sup> < 0.05.  
<sup>b</sup><sup>P</sup> < 0.01.  
<sup>c</sup><sup>P</sup> < 0.001.

0.030 × PLT + 2.400. The combination of different markers such as Model A + ALT + PLT, Model B + ALT + PLT and Model C + ALT + PLT improved diagnostic power significantly, the AUROCs were 0.912, 0.829 and 0.885, respectively (Table 4).

**Verifying the diagnostic power of N-glycome models in the validation group**

We applied N-glycome models (Model A-Model C) in patients in the validation group. The AUROCs of each model are demonstrated in Supplementary Figure 1. The cutoff value was determined in the modeling group, and sensitivity, specificity, PPV, NPV of Model A, Model B and Model C were calculated, respectively, and the data are shown in Supplementary Table 2. The diagnostic powers of the three models were all significantly improved in the validation group. For example, the AUROC of Model A for detecting fibrosis F0-F1 from F2-F4 was 0.915, and the cutoff value was set at 0.8032, yielding a sensitivity of 94.6% and a specificity of 78.8%. The AUROCs of N-glycome models (Model A-Model C) were all higher than those of other diagnostic panels in the validation cohort (Supplementary Table 3). The combined markers (Model A + ALT + PLT, Model B + ALT + PLT and Model C + ALT + PLT) also performed better in diagnostic efficacy than N-glycome model alone, and the AUROCs were 0.929, 0.858 and 0.867, respectively.

**Distinguishing adjacent fibrosis stages in 432 patients with CHB using combined N-glycome markers**

We tried to distinguish adjacent fibrosis stages of 432 patients using three combined N-glycome markers (Model A + ALT + PLT, Model B + ALT + PLT and Model C + ALT + PLT) and analyzed the diagnostic power. It was found that the diagnostic power of the combined marker (Model A + ALT + PLT) was effective in the differentiation of fibrosis F0-F1 from F2-F4 (AUROC = 0.917) (Supplementary Table 4). We then applied the combined markers (Model B + ALT + PLT) to differentiate F2 from F3-F4 among fibrosis patients with stages F2-F4 who were diagnosed by the combined markers (Model A + ALT + PLT), and AUROC was 0.720 (Supplementary Table 4). Lastly, we used the combined markers (Model C + ALT + PLT) to differentiate F3 from F4 among fibrosis patients with stages F3-F4 who were diagnosed by the combined markers (Model B + ALT + PLT), the AUROC was 0.785 (Supplementary Table 4). These results demonstrated that application of the three markers together yielded favorable effects in practicability for distinguishing adjacent fibrosis stages.

**DISCUSSION**

Emerging genomics, proteomics and glycomics are increasingly applied to the diagnosis and differentiation of disease progression. Glycan is a component of bioactive molecules that participate in regulating inflammation, pathogen infection, immune response and expression of markers on cancer cells<sup>[14,15]</sup>. Physiologically, the liver is the major organ responsible for glycoprotein biosynthesis and modification besides immunoglobulin-producing B lymphocytes. Liver abnormalities associated with aberrant glycosylation can influence the structure, quantity and function of sugar...
Table 4 Predictive value of fibrosis diagnostic models in the modeling group

| Fibrosis stage | Model                  | Cutoff value | Sen   | Spec  | PPV   | NPV   | AUROC |
|----------------|------------------------|--------------|-------|-------|-------|-------|-------|
| F0-F1 vs F2-F4 | Model A                | 0.8032       | 0.871 | 0.839 | 0.951 | 0.642 | 0.890 |
|                | GlycoFibroTest         | -0.4404      | 0.280 | 0.903 | 0.913 | 0.257 | 0.538 |
|                | GlycoCirrhoTest        | -0.2946      | 0.778 | 0.452 | 0.837 | 0.359 | 0.644 |
|                | APRI                   | 0.4871       | 0.626 | 0.839 | 0.933 | 0.385 | 0.790 |
|                | FIB-4                  | 0.4720       | 0.401 | 0.855 | 0.908 | 0.285 | 0.599 |
|                | S index                | 0.0543       | 0.742 | 0.596 | 0.871 | 0.386 | 0.701 |
|                | Model A + ALT + PLT    | 0.7779       | 0.923 | 0.806 | 0.945 | 0.746 | 0.912 |
| F0-F2 vs F3-F4 | Model B                | 0.3418       | 0.986 | 0.514 | 0.662 | 0.974 | 0.752 |
|                | GlycoFibroTest         | -0.2295      | 0.128 | 1.000 | 1.000 | 0.543 | 0.524 |
|                | GlycoCirrhoTest        | 0.0033       | 0.270 | 0.911 | 0.745 | 0.564 | 0.618 |
|                | APRI                   | 0.3964       | 0.819 | 0.562 | 0.638 | 0.766 | 0.740 |
|                | FIB-4                  | 0.3930       | 0.609 | 0.726 | 0.677 | 0.663 | 0.699 |
|                | S index                | 0.0890       | 0.664 | 0.711 | 0.690 | 0.686 | 0.738 |
|                | Model B + ALT + PLT    | 0.4404       | 0.855 | 0.685 | 0.720 | 0.833 | 0.829 |
| F0-F3 vs F4    | Model C                | 0.2258       | 0.808 | 0.598 | 0.407 | 0.901 | 0.747 |
|                | GlycoFibroTest         | -0.3200      | 0.274 | 0.939 | 0.606 | 0.791 | 0.571 |
|                | GlycoCirrhoTest        | -0.0287      | 0.356 | 0.827 | 0.413 | 0.790 | 0.607 |
|                | APRI                   | 0.4206       | 0.859 | 0.488 | 0.359 | 0.912 | 0.666 |
|                | FIB-4                  | 0.3762       | 0.831 | 0.657 | 0.447 | 0.921 | 0.795 |
|                | S index                | 0.0817       | 0.765 | 0.571 | 0.380 | 0.876 | 0.717 |
|                | Model C + ALT + PLT    | 0.3010       | 0.789 | 0.845 | 0.629 | 0.923 | 0.885 |

F1 APRI = [AST (IU/L) / 40] × 100/PLT (10^9/L). F2 FIB-4 = [age (yr) × AST (IU/L)] / [PLT (10^9/L) × ALT^1/2 (IU/L)]. F3 S index = 1000 × GGT (IU/L) / [PLT (10^9/L) × ALB (g/L)]. Sen: Sensitivity; Spec: Specificity; PPV: Positive predictive value; NPV: Negative predictive value; ALT: Alanine aminotransferase; GGT: γ-glutamyl-transpeptidase; PLT: Platelet; AUROC: Area under receiver operating characteristic curves.

Glycosylation is highly sensitive to the disease environment and has been implicated in many diseases including liver fibrosis/cirrhosis and cancer[11,12,16]. Screening for diagnostic and prognostic glycan markers for liver diseases will be of great clinical significance.

Serum markers are non-invasive, easy to perform, and convenient to monitor the progression of the liver diseases. Previous studies supported our assumption and revealed that serum N-glycan profiling was helpful in identifying fibrosis/cirrhosis induced by HCV[11,12] or HBV[17]. We enlarged the number of clinical cases to 432 HBV-infected patients and examined the diagnostic power of GlycoFibroTest and GlycoCirrhoTest made up of two N-glycan structures. Our results indicated dissatisfactory diagnostic power of GlycoFibroTest and GlycoCirrhoTest in HBV-related fibrosis patients (GlycoFibroTest: AUROC = 0.538, 0.524 and 0.571; GlycoCirrhoTest: AUROC = 0.644, 0.618 and 0.607 for F0-F1 from F2-F4, F0-F2 from F3-F4, and F0-F3 from F4; respectively, Table 4). Therefore, we attempted to construct more effective and robust models for diagnosing fibrosis induced by HBV and finally found that the diagnostic power of the models containing four N-glycan markers were far more effective than GlycoCirrhoTest and GlycoFibroTest.

In the present study, we used DSA-FACE technology to analyze serum N-glycan changes of 432 HBV infected patients with hepatic fibrosis and found that peak 1 (an agalacto core-α-1,6-fucosylated biantennary glycan, NGA2F), peak 3/4 (a single agalacto core-α-1,6-fucosylated biantennary glycan, NG1A2F), and peak 8 (a tri-antennary glycan, NA3) were correlated with liver fibrosis at varying stages. Using logistic regression analysis, we constructed three mathematic diagnostic models that can be used for monitoring continuous progression of fibrosis, which were defined as Model A (for fibrosis F0-F1 specified from F2-F4), Model B (for fibrosis F0-F2 specified from F3-F4) and Model C (for fibrosis F0-F3 specified from F4). Furthermore, the three models were compared with previously reported N-glycomic markers (GlycoFibroTest and GlycoCirrhoTest) and serological marker panels (S index, APRI and FIB-4). Results showed that Model A (AUROC = 0.890) and Model B (AUROC =...
Figure 2  Receiver operating characteristic curve. A: Receiver operating characteristic (ROC) curve for the prediction of fibrosis F0-F1 from fibrosis F2-F4 in the modeling group. [area under ROC curves (AUROCs): Model A (0.890) > GlycoCirrhoTest (0.644) > GlycoFibroTest (0.538)]; B: ROC curve for the prediction of fibrosis F0-F2 from fibrosis F3-F4. [AUROCs: Model B (0.752) > GlycoCirrhoTest (0.618) > GlycoFibroTest (0.524)]; C: ROC curve for the prediction of fibrosis F0-F3 from fibrosis F4. [AUROCs: Model C (0.747) > GlycoCirrhoTest (0.607) > GlycoFibroTest (0.571)].

It was indicated that biantennary peak 1(NGA2F) and peak 3 (NG1A2F) increased gradually, and biantennary peak 4 (NG1A2F) and triantennary peak 8 (NA3) decreased gradually with the fibrosis progression in patients with CHB. Generally, most serum N-linked glycoproteins are synthesized and secreted by the liver and B-lymphocytes. The N-glycans on serum IgG are mostly core fucosylated biantennary glycans[18], including peak 1 (NGA2F) and peak 3 (NG1A2F), which allow the follow-up of liver fibrosis accompanied inflammation and immune response in HBV infection. Peak 3 (NG1A2F) and peak 4 (NG1A2F) are isomers, which could lead to increased concentration of peak 3 and the decreased level of peak 4 which could be explained by the increased activity of peak 3 synthesis possibly competing for the substrate of peak 4. Triantennary glycan (NA3, peak 8), the abundance of which decreased with advancing fibrosis, was hepatic specific and presented on non-IgG proteins. Peak 8 synthesized by the liver in HBV-induced fibrotic patients could reflect, at least in part, alteration of the physiological condition of the liver. N-glycan precursors are synthesized in the endoplasmic reticulum/Golgi apparatus of hepatocyte and changes of N-glycan profile in the serum could be related to the expression levels of glycosyltransferases in liver cells, leading to modifications in both the core structure of glycans and the terminal structure[19]. N-acetylglucosaminyltransferase IV (GnT-IV) and N-acetylglucosaminyltransferase V (GnT-V) are responsible for multiantennary glycans[20,21], which may cause changes of peak 8. While peak 1, peak 3 and peak 4 were altered in parallel with the fibrosis progression, correlative glycosyltransferases remain to be investigated.

As the three mathematic models (renamed as Model A, Model B and Model C) constructed in the modeling group provided considerable diagnostic power for assessment of HBV-related liver fibrosis. We observed their application in the validation group and found the diagnostic accuracy of the three N-glycan models performed directly increased in the validation cohort (AUROC = 0.915, 0.815 and 0.772, respectively). Comparing the diagnostic power with other marker panels, we found that the diagnostic efficacies of the combined models (Model A + ALT + PLT , Model B + ALT + PLT and Model C + ALT + PLT) were all superior to those of other panels (or tests), AUROC was 0.929, 0.858 and 0.867, respectively in the validation group (Supplementary Table 3).

Imaging techniques are widely applied to diagnose liver fibrosis, such as transient elastography (TE), which are useful in diagnosing significant fibrosis (≥ F3) in chronic hepatitis patients. In our study, Model A exhibited a desirable diagnostic power in distinguishing fibrosis (F0-F1) from fibrosis (F2-F4) (AUROC = 0.890 and 0.915 for the modeling and validation groups, respectively), and showed advantages over TE (AUROC = 0.800, Sen = 64%, Spe = 85%)[22]. N-glycan markers (Model A) in
combination with ALT and PLT presented more effective diagnostic potency (AUROC = 0.912 and 0.929 for the modeling and validation groups, respectively). Unfortunately, it was impossible to compare the diagnostic accuracy between N-glycan markers and TE as TE data was unavailable in the patients at the time of enrollment. Importantly, patients with severe fibrosis (≥ F2) need clinical treatment. N-glycan analysis facilitated the diagnosis of severe liver fibrosis (≥ F2) from mild liver fibrosis (F0-F1) more powerfully. As liver fibrosis commonly undergoes pathological changes after HBV infection, the F4 stage of fibrosis (liver cirrhosis) was regarded as a high risk factor for the development of HCC. Patients with cirrhosis need a lifelong antiviral treatment; the combined model (Model C + ALT + PLT) could distinguish cirrhosis from fibrosis successfully. (AUROC = 0.885, Sen = 78.9%, Spe = 84.5%) (Table 4). We also found that the three combined markers (Model A + ALT + PLT, Model B + ALT + PLT and Model C + ALT + PLT) are promising and innovative tools in stratifying adjacent degrees of fibrosis (F0-F1/F2/F3/F4) (AUROC = 0.917, 0.720, 0.785, respectively) (Supplementary Table 4).

In addition, we analyzed serum N-glycan data of 159 chronic hepatitis C patients and used the three N-glycan diagnostic models to determine the liver fibrosis of chronic hepatitis C patients (Supplementary Table 5 and 6). The diagnostic power of mathematical scoring formulas was decreased as compared with the result obtained from HBV patients (Supplementary Table 7, Supplementary Figure 2). Results support that liver fibrosis induced by HCV or HBV were distinguishable and identifiable. The reason may be that N-glycan changes in HBV-infected patients during the development of fibrosis did not completely resemble the changes occurring in HCV patients. We propose that the changed abundances of N-glycan peaks were possibly associated with infection of different viruses (i.e., HBV is DNA virus, and HCV is RNA virus). Although the molecular mechanism involved is unknown, the altered serum N-glycan in CHB patients or chronic hepatitis C patients could attribute to alterations of the glycosylation machinery in the hepatocyte during the progression of viral infection. Therefore, it is essential and meaningful to establish an N-glycan diagnostic model for distinguishing liver fibrosis induced by HBV alone.

Our study indicates that serum N-glycan profiling is a promising non-invasive method for diagnosing liver fibrosis induced by HBV. However, individual variation in the serum N-glycan profile of patients still existed and might influence model effectiveness. We observed that mean ALT level in some patients with the obvious liver fibrosis (≥ F2) exceeded the double upper limit of normal (ULN). We attempted to select these patients to analyze their N-glycan profiles and found that the N-glycomic changes in patients with higher serum ALT levels (> 2 ULN) were different from the patients who had lower ALT levels (< 2 ULN). It was supposed that higher ALT levels which commonly suggest a serious liver necroinflammation state could influence N-glycan profiles in serum. Additionally, PLT levels were all significantly different (P < 0.05) between patients with liver fibrosis F2-F4 and F0-F1, F3-F4 and F0-F2, F4 and F0-F3. Therefore, we combined N-glycan markers with ALT and PLT to evaluate liver fibrosis induced by HBV, which eliminated the effects of ALT and PLT and achieved a higher diagnostic accuracy.

In summary, we found the alterations of serum N-glycan associated with HBV related liver fibrosis and established three mathematic diagnostic models based on serum N-glycan markers for assessing hepatic fibrosis. The results showed that multiparameter N-glycan models are powerful in diagnosing fibrosis, especially early-stage fibrosis (≤ F2). When combining with ALT and PLT, the diagnostic accuracy was improved considerably. Serum N-glycan profiling experiment requires a standard DNA sequencer equipped in a clinical laboratory, and three μL serum is sufficient to diagnose liver fibrosis accurately, which potentially reduces the need for liver biopsy. Further research focusing on the development of N-glycome marker-based diagnostic kits will be of great clinical significance. Prospective studies with a larger number of clinical samples to investigate relevant molecular regulatory mechanisms related to the changes of N-glycan markers are required.

**ARTICLE HIGHLIGHTS**

**Research background**

Chronic hepatitis B virus (HBV) infection is a risk factor for the development of cirrhosis and hepatocellular carcinoma, in which the end-stage liver disease eventually leads to death. Accurate detection of the early-stage fibrosis is of great clinical significance to implement the intervention and block the progression of liver disorders. Liver biopsy as the gold standard is invasive and prone to clinical sampling error. So, we attempted to find more convenient and effective serum markers for detecting HBV-induced early-stage liver fibrosis.
Cao X et al. Serum N-glycan markers for diagnosing LF induced by HBV

Research motivation
GlycoFibroTest and GlycoCirrhoTest were unsuitable for assessing stages of fibrosis induced by HBV. Therefore, we investigated characteristics of serum N-glycan profiling of HBV-related liver fibrosis and tried to find effective diagnostic markers.

Research objectives
To investigate serum N-glycan profiling of patients with HBV-related liver fibrosis and verify multiparameter diagnostic models constructed based on serum N-glycan changes.

Research methods
N-glycan profiles were analyzed from sera of 432 HBV-infected patients with liver fibrosis by DNA sequencer-assisted fluorophore-assisted carbohydrate electrophoresis. Multiparameter diagnostic models were established based on statistically significant changed N-glycans using logistic regression analysis. The receiver operating characteristic (ROC) curve analysis was performed to evaluate diagnostic efficacy of N-glycan models, and compared with APRI, FIB-4, S-index and GlycoCirrho-test, GlycoFibro-test.

Research results
Multiparameter models containing four N-glycans (peak 1, peak 3/4 and peak 8) were powerful in diagnosing liver fibrosis, especially at the early stage of fibrosis (≤ F2) with area under ROC curve of 0.890. When combining with alanine aminotransferase and platelet tests, the diagnostic accuracy was improved effectively with area under ROC curve of 0.912, 0.829, and 0.885, respectively, for discriminating fibrosis F0-F1 from F2-F4, F0-F2 from F3-F4, and F0-F2 from F4, and these models were also promising in distinguishing adjacent fibrosis (F0-F1/F2/F3/F4). We confirmed that diagnostic power of both multiparameter and combined models was high in the validation group.

Research conclusions
Multiparameter models based on serum N-glycans are effective supplementary markers to distinguish between adjacent fibrosis stages caused by HBV, especially in combination with alanine aminotransferase and platelet tests.

Research perspectives
This study provides insight into the role of serum N-glycan profiles in diagnosing HBV-related liver fibrosis and a new diagnostic approach for other liver diseases.

ACKNOWLEDGEMENTS
Authors are grateful to Sunlon pharmaceutical Co., Ltd for helping in the sample collection. We are grateful to Dr. Joseph Carpenito from Division of Pulmonary and Critical Care Medicine, New York University School of Medicine, United States for proof reading and editing the manuscript.

REFERENCES
1. Lledó JL, Fernández C, Gutiérrez ML, Ocaña S. Management of occult hepatitis B virus infection: an update for the clinician. World J Gastroenterol 2011; 17: 1563-1568 [PMID: 21472122 DOI: 10.3748/wjg.v17.i12.1563]
2. World Health Organization. Global hepatitis report, 2017. Available from: https://www.who.int/hepatitis/publications/global-hepatitis-report2017/en/
3. Kennedy P, Wagner M, Castilíos L, Hong CW, Johnson CL, Sirlin CB, Taouli B. Quantitative Elastography Methods in Liver Disease: Current Evidence and Future Directions. Radiology 2018; 286: 738-763 [PMID: 29461940 DOI: 10.1148/radiol.2018170601]
4. Sharma P, Dhawan S, Bansal R, Tyagi P, Bansal N, Singla V, Kumar A, Matin A, Arora A. Usefulness of transient elastography by FibroScan for the evaluation of liver fibrosis. Indian J Gastroenterol 2014; 33: 445-451 [PMID: 25138767 DOI: 10.1007/s12664-014-0491-x]
5. Wai CT, Greenenson JK, Fontana RJ, Kalbfeisch JD, Marrero JA, Conjeevaran HS, Lok AS. A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 2003; 38: 518-526 [PMID: 12883497 DOI: 10.1053/jhep.2003.50346]
6. Sterling RK, Lissen E, Chumcek N, Sola R, Correa MC, Montaner J, S Sulkowski M, Torriani FJ, Dieterich DT, Thomas DL, Messinger D, Nelson M, APRIOCT Clinical Investigators. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 2006; 43: 1317-1325 [PMID: 16729309 DOI: 10.1002/hep.21178]
7. Zhou K, Gao CF, Zhao YP, Liu HL, Zheng RD, Xian JC, Xu HT, Mao YM, Zeng MD, Lu LG. Simpler score of routine laboratory tests predicts liver fibrosis in patients with chronic hepatitis B. J Gastroenterol Hepatol 2010; 25: 1569-1577 [PMID: 20796157 DOI: 10.1111/j.1440-1746.2010.06383.x]
8. Lan Y, Hao C, Zeng X, He Y, Zeng P, Guo Z, Zhang L. Serum glycoprotein-derived N- and O-linked glycans as cancer biomarkers. Am J Cancer Res 2016; 6: 2390-2415 [PMID: 27904760]
9. Abbott KL, Nairn AV, Hall EM, Horton MB, McDonald JF, Moremen KW, Dinulescu DM, Pierce M. Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. Proteomics 2008; 8: 3218-3229 [PMID: 18696643 DOI: 10.1002/pmic.200800917]
10. Callewaert N, Geysens S, Molemans F, Contreras R. Ultrasensitive profiling and sequencing of N-linked
oligosaccharides using standard DNA-sequencing equipment. *Glycobiology* 2001; **11**: 275-281 [PMID: 11358876 DOI: 10.1093/glycob/11.4.275]

11 Callewaert N, Van Vlierberghe H, Van Hecke A, Laroy W, Delanghe J, Contreras R. Noninvasive diagnosis of liver cirrhosis using DNA sequencer-based total serum protein glycomics. *Nat Med* 2004; **10**: 429-434 [PMID: 15152612 DOI: 10.1038/nn1006]

12 Vanderschaeghe D, Laroy W, Sahlon E, Hafon P, Van Hecke A, Delanghe J, Callewaert N. GlycoFibroTest is a highly performant liver fibrosis biomarker derived from DNA sequencer-based serum protein glycomics. *Mol Cell Proteomics* 2009; **8**: 986-994 [PMID: 19181623 DOI: 10.1074/mcp.M800470-MCP200]

13 Liu XE, Desmyter L, Gao CF, Laroy W, Dewaele S, Vanhooren V, Wang L, Zhuang H, Callewaert N, Libert C, Contreras R, Chen C. N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus. *Hepatology* 2007; **46**: 1426-1435 [PMID: 17683101 DOI: 10.1002/hep.21855]

14 Drake RR. Glycosylation and cancer: moving glycomics to the forefront. *Adv Cancer Res* 2015; **126**: 1-10 [PMID: 25727144 DOI: 10.1016/bs.acr.2014.12.002]

15 An HJ, Kronewitter SR, de Leoz ML, Lebrilla CB. Glycomics and disease markers. *Curr Opin Chem Biol* 2009; **13**: 601-607 [PMID: 19775929 DOI: 10.1016/j.cob.2009.08.015]

16 de Oliveira RM, Ornelas-Ricart CA, Araujo Martins AM. Use of Mass Spectrometry to Screen Glycan Early Markers in Hepatocellular Carcinoma. *Front Oncol* 2017; **7**: 328 [PMID: 29379771 DOI: 10.3389/fonc.2017.00328]

17 Qu Y, Gao CF, Zhou K, Zhao YP, Xu MY, Lu LG. Serum N-glycomic markers in combination with panels improves the diagnosis of chronic hepatitis B. *Ann Hepatol* 2012; **11**: 202-212 [PMID: 22345337]

18 Arnold JN, Royle L, Dwek RA, Rudd PM, Sim RB. Human immunoglobulin glycosylation and the lectin pathway of complement activation. *Adv Exp Med Biol* 2005; **564**: 27-43 [PMID: 16400805 DOI: 10.1007/0-387-25515-X_9]

19 Trombetta ES, Parodi AJ. Glycoprotein reglucosylation. *Methods* 2005; **35**: 328-337 [PMID: 15804604 DOI: 10.1016/j.ymeth.2004.10.004]

20 Yao M, Zhou DP, Jiang SM, Wang QH, Zhou XD, Tang ZY, Gu JX. Elevated activity of N-acetylglucosaminyltransferase V in human hepatocellular carcinoma. *J Cancer Res Clin Oncol* 1998; **124**: 27-30 [PMID: 9498831 DOI: 10.1007/s004320050129]

21 Ito Y, Miyoshi E, Sakom M, Takeda T, Noda K, Tsujimoto M, Ito S, Honda H, Takemura F, Wakasa K, Monden M, Matsuura N, Taniguchi N. Elevated expression of UDP-N-acetylglucosamine: alpha-mannosidase beta1,6-N-acetylglucosaminyltransferase is an early event in hepatocarcinogenesis. *Int J Cancer* 2001; **91**: 631-637 [PMID: 11267972]

22 Fouche J, Chanteloup E, Vergniol J, Castéra L, Le Bail B, Adhoute X, Bertet J, Couzigou P, de Ledinghen V. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut* 2006; **55**: 403-408 [PMID: 16020491 DOI: 10.1136/gut.2005.069153]
