Title Page

Title

A combined and sequential non-invasive approach to diagnosing nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease and persistently normal alanine aminotransferase levels

Short Title: Non-invasive diagnosis of NASH

Authors' name

Kenneth I. Zheng[1], Wen-Yue Liu[2], Xiao-Yan Pan[2], Hong-Lei Ma[1], Pei-Wu Zhu[3], Xi-Xi Wu[4], Giovanni Targher[5], Christopher D. Byrne[6], Xiao-Dong Wang[1,7], Yong-Ping Chen[1,7], Fengmin Lu[8*] and Ming-Hua Zheng[1,7*]

Institution

[1]NAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[2]Department of Endocrinology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[3]Department of Clinical Laboratory, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[4]Department of Gastroenterology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China;

[5]Section of Endocrinology, Diabetes and Metabolism, Department of Medicine, University and Azienda Ospedaliera Universitaria Integrata of Verona, Verona, Italy;

[6]Southampton National Institute for Health Research Biomedical Research Centre, University Hospital Southampton, Southampton General Hospital, Southampton, UK;

[7]Institute of Hepatology, Wenzhou Medical University, Wenzhou, China;

[8]Department of Microbiology & Infectious Disease Center, School of Basic Medicine, Peking University Health Science Center, Beijing, China.

***Co-corresponding author:**

Ming-Hua Zheng, MD, PhD

NAFLD Research Center, Department of Hepatology, the First Affiliated Hospital of Wenzhou Medical University; No. 2 Fuxue Lane, Wenzhou 325000, China.

E-mail: zhengmh@wmu.edu.cn; fax: (86) 577-55578522; tel: (86) 577-55579622.

Fengmin Lu, PhD

Department of Microbiology & Infectious Disease Center, School of Basic Medicine, Peking University Health Science Center; 38 Xueyuan Road, Haidian District, Beijing 100191, China.

E-mail: lu.fengmin@bjmu.edu.cn; tel: 86-10-8280-5136; fax: 86-10-8280-5136

Word count: 3923

Number of figures and tables: 3 tables, 2 figures, 4 supplementary tables, 3 supplementary figures

Author contributions

Study concept and design: Yong-Ping Chen, Fengmin Lu and Ming-Hua Zheng

Acquisition of data: Wen-Yue Liu, Xiao-Yan Pan, Hong-Lei Ma, Pei-Wu Zhu and Xi-Xi Wu

Pathology analysis: Xiao-Dong Wang

Analysis and interpretation of data: Kenneth I. Zheng

Drafting of the manuscript: Kenneth I. Zheng

Critical revision of the manuscript for important intellectual content: Giovanni Targher and Christopher D. Byrne

Study supervision: Fengmin Lu and Ming-Hua Zheng

All authors contributed to the manuscript for important intellectual content and approved the submission.

Compliance with Ethical Requirements

Acknowledgements

Declaration of personal interests: We thank Prof. Ji-Min Liu, a pathologist from McMaster University, Canada, who conducted quality control of liver pathology data. The authors thank Hotgen Biotech Inc. (Beijing, China) and Herui Biomed Company Limited (Suzhou, China) for providing ELISA kits.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. GT is supported in part by grants from the University School of Medicine of Verona, Verona, Italy. CDB is supported in part by the Southampton NIHR Biomedical Research Centre (IS-BRC-20004), UK.

Conflict of interest

The authors have no conflict of interests to declare.

The authors thank Hotgen Biotech Inc. (Beijing, China) and Herui Biomed Company Limited (Suzhou, China) for providing ELISA kits.

Data sharing

Data generated in this study is not open to public access.

Ethical approval

The research protocol of this study was approved by the ethics committee of the First Affiliated Hospital of Wenzhou Medical University.

Informed consent

All procedures followed were in accordance with the local ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. Informed consent was obtained from all patients for being included in the study.

Abstract

Background & Aim: Imaging-confirmed nonalcoholic fatty liver disease (NAFLD) with normal alanine aminotransferase (nALT) levels is infrequently the subject for further evaluation. Early diagnosis of nonalcoholic steatohepatitis (NASH) is needed for preventing disease progression. Thus, we tested the clinical utility of serum Golgi protein 73 (GP73) levels and developed a new non-invasive score to diagnose NASH in patients with biopsy-confirmed NAFLD and persistently nALT levels.

Methods: Serum GP73 and cytokeratin-18 M30 fragments (CK18-M30) levels were measured in 345 patients with biopsy-proven NAFLD. We developed a new score, named G-NASH model (by incorporating serum GP73), and combined it with serum CK18-M30 measurement in a sequential non-invasive approach to accurately identify NASH among patients with NAFLD and persistently nALT levels.

Results: 105 (30.4%) patients had persistently nALT, 53 of which had histologically-confirmed NASH. Both serum GP73 and CK18-M30 levels alone had poor diagnostic accuracy for identifying NASH (55.2% and 51.6% respectively) in these patients. Conversely, G-NASH model performed better than other established non-invasive scoring systems, and by using our proposed sequential non-invasive approach, 82.9% of patients with NASH were correctly identified.

Conclusions: NASH is highly prevalent in NAFLD patients with persistently nALT levels. The G-NASH model accurately identifies NASH in this patient group.

Keywords

Golgi protein 73, cytokeratin-18, non-invasive methods, nonalcoholic fatty liver disease, alanine aminotransferase, nonalcoholic steatohepatitis, liver biopsy, combined and sequential approach, misdiagnosis, decision curve analysis.

Abbreviations

NAFLD, nonalcoholic fatty liver disease; ALT, alanine aminotransferase; NASH, nonalcoholic steatohepatitis; ULN, upper limit of normal; abnALT abnormal alanine aminotransferase; nALT, normal alanine aminotransferase; CK18-M30, cytokeratin 18 M30 fragments; GP73, Golgi membrane glycoprotein 73; BMI, body mass index; HOMA-IR, homeostasis model assessment of insulin resistance; ROC, receiver operating characteristic curve; AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; DCA, decision curve analysis; NB, net benefit; SDPV, standard deviation of platelet volume; CBR, cost-benefit ratio; Mis, misdiagnosis; IFN-γ, interferon gamma

Highlight

What is already known about this subject?

- Liver disease severity is often not investigated further in patients with imaging-confirmed NAFLD, who have normal ALT levels.
- Diagnosing NASH non-invasively remains a big challenge.

What does this study add?

- We have tested whether serum Golgi protein 73 (GP73) levels can predict NASH in patients with biopsy-confirmed NAFLD and normal ALT levels.
- Our newly developed non-invasive score (G-NASH) performed best for accurately identifying NASH when compared to other widely used non-invasive scoring systems.

How might these results change the focus of research or clinical practice?

- By using our proposed combined and sequential non-invasive approach, 82.9% of patients with NASH were correctly identified.
- This is a step toward identifying NASH patients with normal alanine aminotransferase as well as reduce numbers of patients who may be subject to unnecessary liver biopsy.

Introduction

With an estimated global prevalence of nearly 25%, nonalcoholic fatty liver disease (NAFLD) is predicted to cause considerable health, social and economic burdens in the next few years.[1] In patients with NAFLD, abnormal serum alanine aminotransferase (ALT) levels are commonly used (both in China and in other countries) as a surrogate marker of a more severe liver condition, such as nonalcoholic steatohepatitis (NASH), and when NASH occurs with advanced fibrosis or cirrhosis, there is also a marked increase in risk of hepatocellular carcinoma or end-stage liver disease.[2] As a consequence, NAFLD patients with persistently normal ALT levels are often neglected from further assessment of the severity of their liver disease.[3] This is clinically important because there is little evidence that the risk of developing liver-related complications or extra-hepatic diseases in this group of patients is any different from that of their counterparts with abnormal ALT levels.[4] Moreover, the prevalence of patients with NAFLD who have normal serum ALT levels is high. For example, the prevalence of patients with serum ALT levels within the upper limit of normal (ULN) was reported to be over 90% of United States adult individuals with NAFLD by National Health and Nutrition Examination Survey III.[5]

To date, an understanding of the liver disease severity and characteristics of patients with NAFLD who have normal ALT levels is limited. While most studies have included patients with NAFLD and serum levels of ALT above the ULN (abnALT), patients with persistently normal ALT (nALT) have been included only in a few studies of NAFLD.[6, 7] Although serum ALT levels are a poor predictor of the histological severity of NAFLD [2, 7], serum ALT levels

remain to date the most widely used laboratory parameter, both in China and in other countries, for referral to liver specialist(s) in those patients with suspected more severe liver disease.[3]

Increased levels of serum cytokeratin 18 (CK18) have shown promising results for diagnosing NASH in some studies, but is uncertain whether serum CK18 levels may be also useful for diagnosing NASH in patients with persistently nALT.[8]

Golgi membrane glycoprotein 73 (GP73), also known as Golgi phosphoprotein 2, is a type II Golgi-localized integral membrane protein found in the blood when it is cleaved by the proprotein convertase furin. In a previous study, we found that elevated serum GP73 levels were a reliable marker of liver fibrosis and cirrhosis.[9] Elevated serum GP73 levels were also found to be useful for assessing liver inflammation in chronic viral hepatitis B.[10] Also, Lu et al. confirmed the usefulness of GP73 levels in NAFLD-related cirrhosis.[11] Because GP73 is released into the serum as a result of hepatic injury, and it is associated with the severity of liver inflammation in chronic liver diseases, we believe that its diagnostic utility in patients with NASH with persistently nALT should be also examined.

Therefore, in patients with biopsy-proven NAFLD, the aim of our study was to test the diagnostic performance of a combined and sequential strategy, using serum CK18-M30 and GP73 levels, to predict NASH in patients with persistently nALT levels.

Methods

Patients

This is an observational study of the well-characterized Prospective Epidemic Research Specifically Of NASH (PERSONS) cohort.[12] Patients enrolled in this study fulfilled the following criteria: (1) age between 18-75 years; (2) presence of hepatic steatosis on liver ultrasound or other imaging techniques, and then confirmed by liver biopsy, in the absence of other competing causes of chronic liver diseases (as specified in the Results section); and (3) provided informed written consent prior to their inclusion in the study. Persistent nALT status was defined as a serum ALT level below the ULN (i.e., <40 U/L for both sexes, according to our local laboratory reference) for at least 3 months prior to study entry.

Laboratory, anthropometric and clinical data

As previously reported, anthropometric measurements were recorded and all blood samples were obtained on the same day of liver biopsy, after at least an 8-hour fast, and routine biochemical tests performed in our central laboratory.[13] Serum insulin and total thyroxine concentrations were measured by using commercially available radioimmunoassay kits. Body mass index (BMI) was calculated as weight (kg) divided by squared height (m).

The presence of diabetes was defined as a self-reported physician established diagnosis of diabetes, current use of any hypoglycaemic drugs, a fasting glucose level ≥7.0 mmol/L (≥126 mg/dL) or a hemoglobin A1c level ≥6.5% (≥48 mmol/mol). Diagnosis of metabolic syndrome was based on the presence of at least three of the following metabolic abnormalities: (1) waist circumference >90 cm in men and >80 cm in women; (2) triglycerides ≥1.70 mmol/L; (3) high-density lipoprotein <1.0 mmol/L in men and <1.3 mmol/L in women; (4) blood pressure ≥130/85 mmHg or drug treatment; (5) fasting glucose ≥5.6 mmol/L. Hypertension was defined as blood

pressure ≥140/90 mmHg or use of any anti-hypertensive agents. Hyperuricemia was defined as a serum uric acid level >428 μmol/L or use of allopurinol. Homeostasis model assessment (HOMA-IR) was calculated by published formula.[14] Obesity for Asian individuals was defined as BMI ≥25 kg/m^2.[15]

Clinical scoring systems for predicting NASH

The currently proposed soring systems for predicting NASH include the NAFIC score, the Palekar et al.'s risk factors panel, the Cao et al.'s score system, and the Nice model.[16-19]

Measurements of serum GP73 and CK18

Serum GP73 and CK18-M30 concentrations were measured by two commercially enzyme-linked immunosorbent assay kits that were provided by Hotgen Biotech Inc. (Beijing, China) and Herui Biomed Company Limited (Suzhou, China), respectively. Serum samples were stored at -80 °C prior to testing. Both tests had coefficients of variation <15% and were performed by a technician, who was blinded to participants' clinical details. The upper normal limits of serum GP73 and CK18-M30 levels in healthy population, provided by the manufacturer of commercial kits, were 53 ng/mL and 200 U/L, respectively.

Liver biopsy

In brief, liver biopsy was performed and histology was assessed, separately, by an expert liver pathologist (Wang X.D.),[20] who was blinded to clinical and biochemical data of participants, according to the NASH-Steatohepatitis Clinical Research Network scoring system.[21] In this study NASH was defined as presence of NAFLD Activity Score (NAS) ≥4 with at least one

point for each of the three individual features of NAFLD (i.e. steatosis, lobular inflammation, and ballooning), in the absence of other known causes of chronic liver disease. Liver fibrosis was scored from zero to 4 according to the Brunt's criteria.[22]

Statistical analysis

Continuous variables were expressed as mean ± SD and compared by using either the Student's *t*-test for normally distributed variables or the Mann-Whitney test for non-normally distributed variables. Difference between categorical variables was examined with the chi-squared test or the Fisher's exact test as appropriate. A *P*-value <0.05 was considered statistically significant. Data management and analysis were performed using R software (v3.5.2, R Foundation for Statistical Computing, Vienna, Austria).

Clinically relevant variables were pre-selected and subjected to univariable regression analysis. Since normal ALT levels may differ between populations,[23] exploratory analysis was performed with a lower cut-off value of serum ALT levels (≤35 U/L for men and ≤23 U/L for women), referring to our previous study in the Chinese Han population.[24] A multivariable regression model was also determined by (1) backward stepwise logistic regression (SLR), and (2) best-subset regression (BSR). Variables in univariable regression analysis with a P-value <0.10 entered into SLR, and the model was determined with the lowest value of Akaike's Information Criterion. Variables in univariable analysis with a P-value <0.30 were selected to form a whole-set for exhaustive search by BSR.

Subsequently, a receiver operating characteristic curve (ROC) was performed to evaluate the aforementioned models. The area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated to assess the diagnostic accuracy of models or predictors. Sensitivity and specificity were given under the maximal Youden's Index (sensitivity + specificity -1).

To further determine whether incorporation of the models or predictors could improve outcome(s) of patients in clinical practice, a decision curve analysis (DCA) was undertaken using the *rmda* R package. DCA uses net benefit (NB) to assess the performance of predictive risk models. The standardized net benefit (sNB) was a convenient form of NB to be interpreted, which was defined as NB divided by prevalence of disease. The cost-benefit ratio (CBR) was calculated as the odds of P_t ($CBR = P_t / (1-P_t)$).[25]

Results

Baseline characteristics

Between December 2016 and September 2018, a total of 588 individuals with imaging-defined fatty liver were screened. Among these, 244 patients were excluded for the following reasons: 79 had alcoholic fatty liver, 75 had steatosis <5% on liver histology, 71 had viral hepatitis, 7 had autoimmune hepatitis, 3 had drug-induced hepatitis and 15 had missing data for serum GP73 levels. As a result, 345 patients with biopsy-confirmed NAFLD were included in statistical analyses and were then divided into two subgroups according to their serum ALT levels. The flow chart for patient selection process was summarized in **Figure 1**.

Table 1 shows the main clinical and biochemical characteristics of patients with biopsy-proven proven NAFLD stratified by abnALT and persistently nALT (i.e. using a cut-off of serum ALT <40 U/L for both sexes). There were 105 (30.4%) patients with persistently nALT (67 men, 38 women). Mean ALT levels for the abnALT and persistently nALT groups were 106.4±81.4 U/L and 27.7±7.8 U/L, respectively. Serum levels of aspartate aminotransferase (AST), gamma-glutamyl transferase, uric acid, ferritin, GP73, type III pro-collagen, type IV collagen and CK18-M30 were significantly lower in the persistently nALT group, while this latter group of patients were more likely to be older, leaner and of female sex, compared to the abnALT group. Presence of NASH was greater in the abnALT group than in the persistently nALT group (77.9% *vs*. 50.5%, P<0.001). Histologically, the abnALT group had more severe hepatic steatosis and lobular inflammation compared to the persistently nALT group, but comparable stages of liver fibrosis.

Table 1. Baseline characteristics of patients with biopsy-proven NAFLD

		Serum ALT level (U/L)		
	Overall (n = 345)	abnALT group (n = 240)	nALT group (n = 105)	P
Clinical parameters				
Female sex, n(%)	83 (24.1)	45 (18.8)	38 (36.2)	**0.001**
Age, years	40.2 ± 12.5	37.8 ± 12.1	45.5 ± 11.6	**<0.001**
WBC, $\times 10^9$/L	6.22 ± 1.50	6.41 ± 1.60	5.80 ± 1.12	**<0.001**
RBC, $\times 10^{12}$/L	5.01 ± 0.50	5.11 ± 0.50	4.78 ± 0.44	**<0.001**
PLT, $\times 10^9$/L	245.48 ± 57.72	243.62 ± 58.69	249.72 ± 55.50	0.367
MPV, fL	10.92 ± 0.99	10.95 ± 1.02	10.87 ± 0.92	0.506
SDPV, fL	13.80 ± 2.77	13.94 ± 3.02	13.48 ± 2.09	0.160
Albumin, g/L	46.76 ± 3.74	47.34 ± 3.19	45.43 ± 4.50	**<0.001**
ALT, U/L	82.47 ± 77.04	106.43 ± 81.38	27.70 ± 7.77	**<0.001**
AST, U/L	49.08 ± 33.58	59.28 ± 35.50	25.77 ± 6.75	**<0.001**
ALP, U/L	88.43 ± 39.17	90.38 ± 35.94	83.96 ± 45.57	0.161
GGT, U/L	75.79 ± 99.39	89.71 ± 114.30	43.98 ± 34.37	**<0.001**
Fasting glucose, mmol/L	5.60 ± 1.42	5.47 ± 1.24	5.91 ± 1.74	**0.007**
BUN, mmol/L	4.83 ± 1.28	4.80 ± 1.25	4.90 ± 1.34	0.505
Creatinine, μmol/L	67.94 ± 14.49	68.72 ± 14.12	66.17 ± 15.22	0.133

Uric acid, μmol/L	403.35 ± 105.20	420.57 ± 106.75	364.00 ± 90.37	**<0.001**
Total cholesterol, mmol/L	5.05 ± 1.11	5.14 ± 1.09	4.84 ± 1.15	**0.023**
Triglycerides, mmol/L	2.30 ± 1.41	2.31 ± 1.28	2.27 ± 1.69	0.774
HDL cholesterol, mmol/L	1.01 ± 0.23	1.00 ± 0.23	1.04 ± 0.23	0.232
LDL cholesterol, mmol/L	3.08 ± 0.92	3.18 ± 0.89	2.86 ± 0.95	**0.002**
GP73, ng/mL	80.96 ± 40.05	85.55 ± 43.02	70.45 ± 29.88	**0.001**
CK18-M30, U/L	354.21 ± 633.71	457.84 ± 726.10	102.21 ± 88.79	**<0.001**
Zinc, μmol/L	18.75 ± 4.00	18.98 ± 4.00	18.24 ± 3.98	0.169
Thyroxine, nmol/L	105.69 ± 18.49	105.59 ± 18.20	105.93 ± 19.24	0.874
HbA1c, %	6.03 ± 1.27	5.92 ± 1.12	6.28 ± 1.53	**0.014**
Fasting insulin, pmol/L	151.54 ± 216.56	159.67 ± 238.69	132.76 ± 153.19	0.290
HOMA-IR	5.97 ± 10.15	5.96 ± 10.35	5.99 ± 9.70	0.980
AFP, ng/mL	2.99 ± 1.42	3.01 ± 1.34	2.94 ± 1.60	0.674
Ferritin, μg/L	347.71 ± 242.07	393.36 ± 252.50	243.25 ± 178.13	**<0.001**
Type III pro-collagen, ng/mL	21.38 ± 19.74	22.80 ± 23.42	18.21 ± 4.77	**0.049**
Type IV collagen, ng/mL	20.57 ± 9.57	21.73 ± 10.87	18.00 ± 4.80	**0.001**
Hyaluronic acid, ng/mL	48.26 ± 39.89	48.49 ± 46.30	47.74 ± 19.22	0.874
Comorbid diseases				
Diabetes, n (%)	104 (30.1)	66 (27.5)	38 (36.2)	0.136
Metabolic syndrome, n (%)	199 (60.3)	144 (62.9)	55 (54.5)	0.187
Hypertension, n (%)	111 (32.2)	74 (30.8)	37 (35.2)	0.496
Dyslipidemia, n (%)	313 (91.5)	221 (92.9)	92 (88.5)	0.258
Hyperuricemia, n (%)	131 (38.0)	107 (44.6)	24 (22.9)	**<0.001**
Obesity BMI 25-29.9 kg/m^2, n (%)	190 (55.1)	143 (59.6)	47 (44.8)	**0.015**
Obesity BMI ≥30 kg/m^2, n (%)	48 (14.0)	37 (15.5)	11 (10.5)	0.287
Anthropometry				
Waist circumference, cm	92.25 ± 8.70	93.30 ± 8.44	89.87 ± 8.85	**0.001**
Hip circumference, cm	99.69 ± 7.46	100.63 ± 7.27	97.57 ± 7.50	**0.001**
BMI, kg/m^2	26.93 ± 4.20	27.31 ± 4.49	26.05 ± 3.33	**0.010**
Liver histology				
NASH, n (%)	240 (69.6)	187 (77.9)	53 (50.5)	**<0.001**
Steatosis, n (%)				**<0.001**
1	142 (41.2)	67 (27.9)	75 (71.4)	
2	91 (26.4)	73 (30.4)	18 (17.1)	
3	112 (32.5)	100 (41.7)	12 (11.4)	
Hepatocyte ballooning, n (%)				0.272
0	13 (3.8)	8 (3.3)	5 (4.8)	
1	182 (52.8)	121 (50.4)	61 (58.1)	
2	150 (43.5)	111 (46.2)	39 (37.1)	
Lobular inflammation, n (%)				**0.004**
0	21 (6.1)	11 (4.6)	10 (9.5)	
1	249 (72.2)	165 (68.8)	84 (80.0)	

2	70 (20.3)	60 (25.0)	10 (9.5)	
3	5 (1.4)	4 (1.7)	1 (1.0)	
Fibrosis, n (%)				0.059
0	111 (32.3)	69 (28.9)	42 (40.0)	
1	161 (46.8)	124 (51.9)	37 (35.2)	
2	55 (16.0)	34 (14.2)	21 (20.0)	
3	16 (4.7)	11 (4.6)	5 (4.8)	
4	1 (0.3)	1 (0.4)	0 (0.0)	

Notes: Continuous variables are presented as mean ± SD; categorical variables are presented as No. (%); P-values <0.05 are in bold font.
Abbreviations: AFP, alpha-fetoprotein; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; BUN, blood urea nitrogen; CK18-M30, cytokeratin 18 M30 fragments; GGT, gamma-glutamyltransferase; GP73, Golgi protein 73; Hb, hemoglobin; HbA1c, hemoglobin A1c; HCT, hematocrit; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein; MPV, mean platelet volume; NASH, nonalcoholic steatohepatitis; PLT, platelet count; RBC, red blood cell; SDPV, standard deviation of platelet volume; WBC, white blood cell.

To further examine the clinical and biochemical profile of patients with persistently nALT, we subdivided these patients according to the presence/absence of NASH (**Table 2**). In this subgroup, 53 out of 105 (50.5%) patients had NASH. Interestingly, almost all clinical and biochemical variables (including also serum CK18-M30) were comparable between the two groups, with the only exception for serum AST and GP73 levels that were significantly higher in the NASH group than in the non-NASH group.

Table 2. Baseline characteristics of patients with persistently nALT stratified by NASH status

	Overall (n = 105)	non-NASH (n = 52)	NASH (n = 53)	P
Clinical parameters				
Female sex, n(%)	38 (36.2)	16 (30.8)	22 (41.5)	0.346
Age, years	45.5 ± 11.6	46.2 ± 9.5	44.9 ± 13.4	0.571
WBC, $\times 10^9$/L	5.80 ± 1.12	5.76 ± 1.17	5.84 ± 1.09	0.735
RBC, $\times 10^{12}$/L	4.78 ± 0.44	4.80 ± 0.46	4.76 ± 0.41	0.626
PLT, $\times 10^9$/L	249.72 ± 55.50	246.44 ± 48.22	252.94 ± 62.13	0.551
MPV, fL	10.87 ± 0.92	10.99 ± 0.79	10.76 ± 1.02	0.197
SDPV, fL	13.48 ± 2.09	13.88 ± 1.95	13.09 ± 2.18	0.055

Albumin, g/L	45.43 ± 4.50	45.27 ± 3.10	45.59 ± 5.58	0.715
ALT, U/L	27.70 ± 7.77	26.88 ± 8.07	28.49 ± 7.45	0.292
AST, U/L	25.77 ± 6.75	24.02 ± 3.88	27.49 ± 8.39	**0.008**
ALP, U/L	83.96 ± 45.57	81.06 ± 24.40	86.81 ± 59.60	0.520
GGT, U/L	43.98 ± 34.37	41.08 ± 27.35	46.83 ± 40.15	0.394
Fasting glucose, mmol/L	5.91 ± 1.74	5.61 ± 1.36	6.21 ± 2.01	0.080
BUN, mmol/L	4.90 ± 1.34	4.82 ± 1.09	4.98 ± 1.55	0.539
Creatinine, μmol/L	66.17 ± 15.22	68.25 ± 13.77	64.13 ± 16.39	0.167
Uric acid, μmol/L	364.00 ± 90.37	368.27 ± 89.92	359.81 ± 91.48	0.634
Total cholesterol, mmol/L	4.84 ± 1.15	4.90 ± 1.03	4.79 ± 1.26	0.622
Triglycerides, mmol/L	2.27 ± 1.69	2.30 ± 1.92	2.24 ± 1.43	0.859
HDL cholesterol, mmol/L	1.04 ± 0.23	1.01 ± 0.23	1.06 ± 0.22	0.243
LDL cholesterol, mmol/L	2.86 ± 0.95	2.91 ± 0.90	2.80 ± 0.99	0.579
GP73, ng/mL	70.45 ± 29.88	64.59 ± 22.07	76.20 ± 35.21	**0.046**
CK18-M30, U/L	102.21 ± 88.79	95.76 ± 55.88	108.02 ± 110.67	0.504
Zinc, μmol/L	18.24 ± 3.98	17.46 ± 3.61	19.05 ± 4.24	0.075
Thyroxine, nmol/L	105.93 ± 19.24	103.50 ± 14.29	108.32 ± 22.99	0.201
HbA1c, %	6.28 ± 1.53	6.12 ± 1.60	6.45 ± 1.46	0.281
Fasting insulin, pmol/L	132.76 ± 153.19	125.66 ± 169.29	139.86 ± 136.49	0.639
HOMA-IR	5.99 ± 9.70	5.48 ± 10.05	6.51 ± 9.41	0.592
AFP, ng/mL	2.94 ± 1.60	2.96 ± 1.90	2.91 ± 1.24	0.877
Ferritin, μg/L	243.25 ± 178.13	214.96 ± 172.49	272.52 ± 182.12	0.218
Type III pro-collagen, ng/mL	18.21 ± 4.77	17.30 ± 4.54	19.13 ± 4.86	0.050
Type IV collagen, ng/mL	18.00 ± 4.80	17.11 ± 4.31	18.90 ± 5.14	0.057
Hyaluronic acid, ng/mL	47.74 ± 19.22	44.20 ± 13.93	51.28 ± 22.94	0.060
Comorbid diseases				
Diabetes, n (%)	38 (36.2)	16 (30.8)	22 (41.5)	0.346
Metabolic syndrome, n (%)	55 (54.5)	25 (50.0)	30 (58.8)	0.49
Hypertension, n (%)	37 (35.2)	19 (36.5)	18 (34.0)	0.943
Dyslipidemia, n (%)	92 (88.5)	48 (92.3)	44 (84.6)	0.357
Hyperuricemia, n (%)	24 (22.9)	12 (23.1)	12 (22.6)	0.999
Obesity BMI 25-29.9 kg/m^2, n (%)	47 (44.8)	22 (42.3)	35 (47.2)	0.761
Obesity BMI ≥30 kg/m^2, n (%)	11 (10.5)	3 (5.8)	8 (15.1)	0.214
Anthropometry				
Waist circumference, cm	89.87 ± 8.85	88.57 ± 9.41	91.12 ± 8.18	0.148
Hip circumference, cm	97.57 ± 7.50	96.38 ± 7.20	98.72 ± 7.67	0.115
BMI, kg/m^2	26.05 ± 3.33	25.63 ± 3.26	26.47 ± 3.38	0.195
Liver histology				
Steatosis, n (%)				**<0.001**
1	75 (71.4)	51 (98.1)	24 (45.3)	
2	18 (17.1)	1 (1.9)	17 (32.1)	
3	12 (11.4)	0 (0.0)	12 (22.6)	
Hepatocyte ballooning, n (%)				**<0.001**

0	5 (4.8)	5 (9.6)	0 (0.0)	
1	61 (58.1)	46 (88.5)	15 (28.3)	
2	39 (37.1)	1 (1.9)	38 (71.7)	
Lobular inflammation, n (%)				**<0.001**
0	10 (9.5)	10 (19.2)	0 (0.0)	
1	84 (80.0)	42 (80.8)	42 (79.2)	
2	10 (9.5)	0 (0.0)	10 (18.9)	
3	1 (1.0)	0 (0.0)	1 (1.9)	
Fibrosis, n (%)				0.156
0	42 (40.0)	24 (46.2)	18 (34.0)	
1	37 (35.2)	20 (38.5)	17 (32.1)	
2	21 (20.0)	7 (13.5)	14 (26.4)	
3	5 (4.8)	1 (1.9)	4 (7.5)	
4	0 (0.0)	0 (0.0)	0 (0.0)	

Notes: Continuous variables are presented as mean ± SD; categorical variables are presented as No. (%); P-values <0.05 are in bold font.
Abbreviations: AFP, alpha-fetoprotein; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; BUN, blood urea nitrogen; CK18-M30, cytokeratin 18 M30 fragments; GGT, gamma-glutamyltransferase; GP73, Golgi protein 73; Hb, hemoglobin; HbA1c, hemoglobin A1c; HCT, hematocrit; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein; MPV, mean platelet volume; PLT, platelet count; RBC, red blood cell; SDPV, standard deviation of platelet volume; WBC, white blood cell;

Diagnostic performance of serum GP73 and CK18-M30 in identifying NASH in patients with persistently nALT

Both concentrations of GP73 and CK18-M30 were increased in patients with NASH, regardless of ALT levels (as summarized in **Supplementary Table 2**). **Table 3** shows the ability to discriminate NASH from simple steatosis with either serum GP73 or CK18-M30 levels alone in the subgroup of NAFLD patients with persistently nALT (defined as a serum ALT level ≤40 U/L). Fifty-eight patients (55.2%) were accurately identified with NASH by using serum GP73 alone, whereas 49 patients (51.6%) were accurately identified by using serum CK18-M30 alone. Therefore, the diagnostic accuracies of GP73 or CK18-M30 alone for identifying NASH were low and essentially comparable (P=0.707). Moreover, as also shown in the table, serum GP73

had higher sensitivity and NPV, but lower specificity and PPV than serum CK18-M30. Similar results were found when we used a lower cut-off value of ALT (≤35 U/L for men and ≤23 U/L for women, respectively) for defining the nALT status (**Supplementary Table 3**).

Table 3. Performance of serum GP73 and CK18-M30 levels in diagnosing NASH in patients with persistently nALT

	Tests result(s)	Liver pathology result (s)		Accuracy (%)	Se (%)	Sp (%)	PPV	NPV	P-value*
		NASH	non-NASH						
GP73	Positive predictive	41	35	0.552	0.773	0.327	0.539	0.586	0.707
	Negative predictive	12	17						
CK18-M30	Positive predictive	5	1	0.516	0.100	0.978	0.821	0.516	
	Negative predictive	45	44						

Note: 10 missing values in serum CK18-M30. Diagnostic accuracy was calculated as the sum of true positive and true negative values divided by the total number of observations. The cut-off value for normal serum ALT was <40U/L for both sexes. Test results above the cut-off value (53ng/mL for GP73 and 200U/L for CK18-M30) were considered as predictive of NASH.

* chi-square test for comparing diagnostic accuracy of serum GP73 and CK18-M30 levels.

Abbreviations: Se, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value.

Development of a prediction model for diagnosing NASH in patients with persistently nALT

To further differentiate NASH from non-NASH in patients with persistently nALT, we performed both univariable and multivariable logistic regression analyses. Values of GP73, AST, zinc and standard deviation of platelet volume (SDPV) were independently associated with the presence of NASH, and a SLR model was also created (**Supplementary Table 4**): 0.0186 × GP73 (ng/mL) + 0.1405 × AST (U/L) + 0.1352 × serum zinc (μmol/L) - 0.3969 × SDPV (fL). The resulting BSR-model contained values of serum GP73, zinc, AST, total thyroxine, SDPV and presence of BMI ≥30 kg/m^2. Thus, a final BSR model (named *G-NASH* as it primarily uses GP73 to predict NASH) was calculated as follows: 0.02 × GP73 (ng/mL) +0.123 × AST (U/L) +

0.1576 × serum zinc (μmol/L) + 0.0227 × total thyroxine (nmol/L) - 0.4525 × SDPV (fL) + 2.0789 × (BMI ≥30 kg/m^2, yes = 1, no = 0). A nomogram is also provided in **Supplementary Figure 3** to show the possible application of the G-NASH model in clinical practice.

Performance of clinical scoring systems for predicting NASH

To compare the performance of the G-NASH model with other available clinical scoring systems for identifying NASH, we performed ROC curve comparisons (**Figure 2**). The G-NASH model showed the best AUROC [0.846 (95% CI, 0.757-0.934)] for diagnosing NASH compared to the SLR model (AUROC = 0.806, 95% CI, 0.708-0.903) (**Figure 2A, Supplementary Table 1**). The good diagnostic performance of the G-NASH model was further confirmed even when we used a lower cut-off value of serum ALT levels (≤35 U/L for men and ≤23 U/L for women, respectively) (**Figure 2B**).[23, 24] The possible reasons for the poorer performances of other available clinical scoring systems for identifying NASH we observed in our study are most likely due to differences in the study populations.

Moreover, the clinical value of the G-NASH model was also evaluated by DCA. As shown in **Supplementary Figure 1A**, decision curves were plotted with sNB *versus* P_t and the cost-benefit ratio (CBR) in all predictive models and predictors, as well as with the "treat all" and "treat none" strategy. By using the G-NASH model and a P_t higher than 0.22, almost all of these patients had the highest sNB compared with all other predictive decision models or predictor(s). We also found an sNB = 0.62 with P_t = 0.40; these data indicate that in patients with persistently nALT, the G-NASH model had a 62% accuracy for identifying NASH and 100% accuracy for identifying non-NASH.

Supplementary Figure 1B shows the predictive ability of the G-NASH model with different P_t thresholds for identifying NASH. The solid line denotes the true positive rate and the dotted line denotes the false positive rate. A P_t of between 0.4 and 0.5 yielded the best test performance and provided the optimal sensitivity and specificity for accurately identifying NASH in patients with persistently nALT.

Supplementary Figure 1C shows the potential clinical impact curve of the G-NASH model. The solid line denotes the number of patients identified as NASH by the G-NASH model, given a specific P_t (high risk), whereas the dotted line denotes the number of patients who truly had NASH at the P_t (high risk with NASH). When the P_t was 40%, 58 out of 100 patients were identified as NASH, compared to 40 patients who were confirmed with NASH by liver biopsy.

Combined use of non-invasive markers to predict NASH in patients with persistently nALT

Serum CK18-M30 has better specificity than serum GP73, whilst GP73 has better sensitivity for diagnosing NASH. Therefore, we devised a combined sequential strategy for identifying patients with persistently nALT who require liver biopsy because of a high probability of NASH. Utilising the better specificity of CK18-M30, patients would firstly be assessed by serum CK18-M30 measurement alone. Those patients with serum CK18-M30 higher than the upper normal limit, i.e., equal or greater than 200 U/L, were deemed as having NASH, whereas those patients below this CK-18-M30 threshold should undergo the G-NASH assessment. Patients with G-NASH score of less than two were deemed as having NAFL (simple steatosis), while those greater than four were diagnosed as NASH. Patients with G-NASH score between two and four

("grey zone") would undergo liver biopsy for staging of NAFLD (**Supplementary Figure 2**). Among the patients assessed with our proposed combined and sequential strategy, five out of seventy subjects were directly deemed as having NASH according to serum CK18-M30 measurement alone, with one individual actually having NAFL, so that the rate of misdiagnosis (Mis) was 20% [Mis. % (n/N): 20% (1/5)]. Sixty-five patients were further assessed by the G-NASH model and six of them were deemed as having NAFL and thirty of them as having NASH, with one patient [Mis. % (n/N): 16.7% (1/6)] and five patients [Mis. % (n/N): 16.7% (5/30)] misdiagnosed, respectively. In summary, seven out of forty-one patients were misdiagnosed by our proposed combined and sequential approach without a biopsy, yielding a misdiagnosis rate of 17.1% [Mis. % (n/N): 17.1% (7/41)]. By this two-step pre-selection, 41.4% (29 out of 70) of our patients with persistently nALT needed biopsy, compared with 92.9% (65 out of 70) of patients when we used only serum CK18-M30 levels alone (by the standard that patients with serum CK18-M30 levels <200 U/L would undergo liver biopsy and those with serum CK18-M30 levels ≥200U/L were diagnosed as having NASH, without further application of the G-NASH scoring system). Of note, 29 out of 70 patients with persistently nALT still need liver biopsy by our sequential approach, which is a result less than ideal. However, it is also important to note that 41 out of 70 patients may be spared from liver biopsy examination, which is a sizable improvement especially if we consider the poor specificity and sensibility of currently available non-invasive tests for diagnosing NASH among patients with NAFLD and persistently nALT.

Discussion

To the best of our knowledge, this is the first study that compared the performance of clinical scoring systems and serum biomarkers for predicting NASH in patients with biopsy-confirmed NAFLD who have persistently nALT levels. In contrast to the few available retrospective studies that have included NAFLD patients with nALT levels,[6, 7] our study involves a relatively large cohort of well-characterized patients with biopsy-proven NAFLD including also a substantial number of patients with persistently nALT levels.

We found that measurement of either serum CK18-M30 or GP73 levels alone were not able to accurately identify NASH in the subgroup of NAFLD patients with persistently nALT levels. However, we found that serum CK18-M30 levels can be usefully used as an initial screening test, and then sequentially combined with a newly developed G-NASH model that includes measurement of serum GP73 and other clinical/biochemical parameters in its equation. Indeed, our results show that the G-NASH model has good diagnostic accuracy for identifying NASH, and by using this model it should be possible to accurately identify those subjects who require a liver biopsy.

Serum ALT levels are the most widely used laboratory test for assessing and monitoring liver injury. Both in China and in many other countries, serum ALT levels are often used by primary care clinicians for referring patients to gastroenterologists/hepatologists for further investigation. However, the fact that increased serum ALT levels can be present only in a proportion of patients with NAFLD (be it simple steatosis or NASH) necessitates finding alternative non-invasive tests to refine specialist referral pathways.[7] Previous efforts have been made to discover novel biomarkers or test panels to better distinguish NASH from simple steatosis. Some

studies have suggested that measurement of serum CK18-M30 levels may be useful;[26, 27] although one study has concluded that CK18-M30 measurement alone lacked sufficient sensitivity for diagnosing NASH.[28] Similarly, in the present study, we found that a substantial number of patients with persistently nALT (53 out of 105) had NASH on histology, and more importantly, serum CK18-M30 levels alone did not have sufficiently high diagnostic accuracy for identifying NASH in this group of patients. Thus, based on these findings, we could make the following inferences: (1) screening for NASH is required in patients with NAFLD who have persistently nALT, because of the progressive potential of their liver disease; and (2) serum CK18-M30 measurement alone does not have sufficient sensitivity for detecting the early stages of NASH in this group of patients.

We have also tested the performance of serum GP73 levels for identifying NASH in NAFLD patients with persistently nALT. A recent study pertaining to hepatocellular carcinoma development showed increased GP73 secretion by hepatocyte endoplasmic reticulum (ER) stress.[29] In view of the recognised relevance of ER stress in NAFLD pathogenesis, it is also possible to speculate that up-regulation of gene expression and increased serum levels of GP73 may also occur in NASH.[30] Additionally, it has been reported that interferon gamma (IFN-γ) was responsible for increased expression of GP73 in hepatocytes[31], and elevated IFN-γ-secreting T helper cell 17 infiltration was also found in the hepatic tissue relative to peripheral blood mononuclear cells.[32] Collectively, these findings might, at least in part, explain the significant association we observed between serum GP73 levels and the presence of NASH.

Increased platelet volume may occur with inflammatory states[33] and alterations in serum zinc[34] and thyroxine levels[35] have been noted to be associated with NAFLD. Zinc is mostly found in skeletal muscle rather than liver and fluctuation of serum zinc levels may occur with sarcopenia[36] that has been noted with NAFLD.[37] Therefore, since NASH is a pro-inflammatory state, there is emerging evidence supporting the inclusion of SDPV, serum zinc and thyroxine levels in the prediction model.

In clinical practice, the use of liver biopsy is limited mostly due to its poor acceptability to patients, high costs, and high risk to benefit ratio. Among NAFLD patients with nALT levels, this issue is even more pertinent, as clinicians are more inclined to investigate only patients with abnALT levels. Since it is known that most of the available non-invasive tests/methods perform poorly to diagnose NASH, better non-invasive tests are increasingly needed, especially for screening the subgroup of NAFLD patients with nALT levels, who may be otherwise misdiagnosed as having less harmful disease by non-specialists. In this regard, serum CK18-M30 levels and other non-invasive clinical scoring systems have been previously investigated for their ability to accurately identify NASH in patients with NAFLD and abnALT.[27, 38] However, these non-invasive clinical scoring systems are complex combinations that include “omics” methodology, which is usually not available in clinical practice.

To resolve this conundrum, we have proposed a sequential non-invasive approach by using serum CK18-M30 measurement first and then applying the G-NASH model. In our cohort of patients with biopsy-confirmed NAFLD, serum GP73 had better sensitivity than serum CK18-M30 alone for the diagnosis of NASH among those patients with persistently nALT, thus

suggesting that serum GP73 might be superior to CK18-M30 alone for diagnosing the early stages of NASH. However, it is noteworthy that increased serum CK18-M30 might reflect prominent hepatic necro-inflammation, which typically occurs in NASH, but not NAFL.[39] Therefore, we propose that serum CK18-M30 measurement alone could be used first to screen out apparent cases of NASH, with subsequent use of the G-NASH model (that includes serum GP73 concentration along with BMI and other laboratory parameters) for accurately differentiating NASH from simple steatosis in the remaining patients. Notably, this proposed combined and sequential approach for non-invasively identifying NASH was effective in reducing unnecessary liver biopsies with an acceptable misdiagnosis rate.

Among the major limitations of our single-center, observational study we need to mention the lack of an external validation cohort to test our proposed strategy of using a sequential combination of serum CK18-M30 and the G-NASH model for non-invasively identifying NASH patients among NAFLD with persistently nALT. In addition, serum CK18-M30 and GP73 levels are not measured in routine clinical practice. Our study also included a relatively small number of patients (n=53) with NASH who had persistently nALT. With this small number of subjects we were unable to stratify the analyses by sex in order to test the prediction model in men and women, separately. Patients included in our study were of Chinese Han ethnicity precluding extrapolation of these results to other ethnic groups. Thus, our results need to be verified in larger studies that also include other ethnic groups.

In conclusion, our results show that ~50% of NAFLD patients with persistently nALT levels have NASH, and the sequential combination of serum CK18-M30 measurement and the G-

NASH model can be used to identify NASH. Not only does this approach accurately identify patients with NASH, but importantly, it also reduces the numbers of patients who may be subjected to an unnecessary liver biopsy.

References

1. Younossi ZM, Blissett D, Blissett R, Henry L, Stepanova M, Younossi Y et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology (Baltimore, Md) 2016;64:1577-1586
2. Chalasani N, Younossi Z, Lavine JE, Charlton M, Cusi K, Rinella M et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. Hepatology (Baltimore, Md) 2018;67:328-357
3. [Expert recommendations on standardized diagnosis and treatment for fatty liver disease in China (2019 revised edition)]. Zhonghua gan zang bing za zhi = Zhonghua ganzangbing zazhi = Chinese journal of hepatology 2019;27:748-753
4. Byrne CD, Patel J, Scorletti E, Targher G. Tests for diagnosing and monitoring non-alcoholic fatty liver disease in adults. Bmj 2018;362:k2734
5. Health USDo, Control HSCfD, Statistics PNCfH. National Health and Nutrition Examination Survey III, 1988-1994. In: Inter-university Consortium for Political and Social Research [distributor]; 1998.
6. Gawrieh S, Wilson LA, Cummings OW, Clark JM, Loomba R, Hameed B et al. Histologic Findings of Advanced Fibrosis and Cirrhosis in Patients With Nonalcoholic Fatty Liver Disease Who Have Normal Aminotransferase Levels. The American journal of gastroenterology 2019;114:1626-1635
7. Wong VW, Wong GL, Tsang SW, Hui AY, Chan AW, Choi PC et al. Metabolic and histological features of non-alcoholic fatty liver disease patients with different serum alanine aminotransferase levels. Alimentary pharmacology & therapeutics 2009;29:387-396
8. Yilmaz Y, Kurt R, Kalayci C. Apoptosis in nonalcoholic steatohepatitis with normal aminotransferase values: zooming in on cytokeratin 18 fragments. Biomarkers in medicine 2010;4:743-745
9. Yao M, Wang L, Leung PSC, Li Y, Liu S, Wang L et al. The Clinical Significance of GP73 in Immunologically Mediated Chronic Liver Diseases: Experimental Data and Literature Review. Clinical reviews in allergy & immunology 2018;54:282-294
10. Yao M, Wang L, You H, Wang J, Liao H, Yang D et al. Serum GP73 combined AST and GGT reflects moderate to severe liver inflammation in chronic hepatitis B. Clinica chimica acta; international journal of clinical chemistry 2019;493:92-97
11. Lu FM, Zhang Y. [Diagnostic application of serum GP73 and the relevant mechanism in the diagnosis of liver cirrhosis]. Zhonghua gan zang bing za zhi = Zhonghua ganzangbing zazhi = Chinese journal of hepatology 2018;26:321-324

12. Zhou Y-J, Ye F-Z, Li Y-Y, Pan X-Y, Chen Y-X, Wu X-X et al. Individualized risk prediction of significant fibrosis in non-alcoholic fatty liver disease using a novel nomogram. United European Gastroenterology Journal 2019;7:1124-1134
13. Liu WY, Zheng KI, Pan XY, Ma HL, Zhu PW, Wu XX et al. Effect of PNPLA3 polymorphism on diagnostic performance of various noninvasive markers for diagnosing and staging nonalcoholic fatty liver disease. J Gastroenterol Hepatol 2019
14. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412-419
15. Leung JC, Loong TC, Wei JL, Wong GL, Chan AW, Choi PC et al. Histological severity and clinical outcomes of nonalcoholic fatty liver disease in nonobese patients. Hepatology (Baltimore, Md) 2017;65:54-64
16. Anty R, Iannelli A, Patouraux S, Bonnafous S, Lavallard VJ, Senni-Buratti M et al. A new composite model including metabolic syndrome, alanine aminotransferase and cytokeratin-18 for the diagnosis of non-alcoholic steatohepatitis in morbidly obese patients. Alimentary pharmacology & therapeutics 2010;32:1315-1322
17. Cao W, Zhao C, Shen C, Wang Y. Cytokeratin 18, alanine aminotransferase, platelets and triglycerides predict the presence of nonalcoholic steatohepatitis. PloS one 2013;8:e82092
18. Palekar NA, Naus R, Larson SP, Ward J, Harrison SA. Clinical model for distinguishing nonalcoholic steatohepatitis from simple steatosis in patients with nonalcoholic fatty liver disease. Liver Int 2006;26:151-156
19. Sumida Y, Yoneda M, Hyogo H, Yamaguchi K, Ono M, Fujii H et al. A simple clinical scoring system using ferritin, fasting insulin, and type IV collagen 7S for predicting steatohepatitis in nonalcoholic fatty liver disease. J Gastroenterol 2011;46:257-268
20. Ye FZ, Liu WY, Zheng KI, Pan XY, Ma HL, Wang XD et al. Homeostatic model assessment of insulin resistance closely related to lobular inflammation in nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol 2020;32:80-86
21. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313-1321
22. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. The American journal of gastroenterology 1999;94:2467-2474
23. Kwo PY, Cohen SM, Lim JK. ACG Clinical Guideline: Evaluation of Abnormal Liver Chemistries. The American journal of gastroenterology 2017;112:18-35
24. Zheng MH, Shi KQ, Fan YC, Liu WY, Lin XF, Li LF et al. Upper limits of normal for serum alanine aminotransferase levels in Chinese Han population. PloS one 2012;7:e43736
25. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol 2018;74:796-804
26. Tada T, Kumada T, Toyoda H, Saibara T, Ono M, Kage M. New scoring system combining the FIB-4 index and cytokeratin-18 fragments for predicting steatohepatitis and liver fibrosis in patients with nonalcoholic fatty liver disease. Biomarkers 2018;23:328-334

27. Miyasato M, Murase-Mishiba Y, Bessho M, Miyawaki M, Imbe H, Tsutsumi C et al. The cytokeratin-18 fragment level as a biomarker of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus. Clinica chimica acta; international journal of clinical chemistry 2014;433:184-189
28. Cusi K, Chang Z, Harrison S, Lomonaco R, Bril F, Orsak B et al. Limited value of plasma cytokeratin-18 as a biomarker for NASH and fibrosis in patients with non-alcoholic fatty liver disease. Journal of hepatology 2014;60:167-174
29. Wei C, Yang X, Liu N, Geng J, Tai Y, Sun Z et al. Tumor Microenvironment Regulation by the Endoplasmic Reticulum Stress Transmission Mediator Golgi Protein 73 in Mice. Hepatology (Baltimore, Md) 2019;70:851-870
30. Lebeaupin C, Vallee D, Hazari Y, Hetz C, Chevet E, Bailly-Maitre B. Endoplasmic reticulum stress signalling and the pathogenesis of non-alcoholic fatty liver disease. Journal of hepatology 2018;69:927-947
31. Kladney RD, Cui X, Bulla GA, Brunt EM, Fimmel CJ. Expression of GP73, a resident Golgi membrane protein, in viral and nonviral liver disease. Hepatology 2002;35:1431-1440
32. Rau M, Schilling AK, Meertens J, Hering I, Weiss J, Jurowich C et al. Progression from Nonalcoholic Fatty Liver to Nonalcoholic Steatohepatitis Is Marked by a Higher Frequency of Th17 Cells in the Liver and an Increased Th17/Resting Regulatory T Cell Ratio in Peripheral Blood and in the Liver. J Immunol 2016;196:97-105
33. Korniluk A, Koper-Lenkiewicz OM, Kaminska J, Kemona H, Dymicka-Piekarska V. Mean Platelet Volume (MPV): New Perspectives for an Old Marker in the Course and Prognosis of Inflammatory Conditions. Mediators Inflamm 2019;2019:9213074
34. Ito T, Ishigami M, Ishizu Y, Kuzuya T, Honda T, Ishikawa T et al. Correlation of serum zinc levels with pathological and laboratory findings in patients with nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol 2019
35. Lonardo A, Mantovani A, Lugari S, Targher G. NAFLD in Some Common Endocrine Diseases: Prevalence, Pathophysiology, and Principles of Diagnosis and Management. Int J Mol Sci 2019;20
36. Hanai T, Shiraki M, Nishimura K, Ohnishi S, Imai K, Suetsugu A et al. Sarcopenia impairs prognosis of patients with liver cirrhosis. Nutrition 2015;31:193-199
37. De Fre CH, De Fre MA, Kwanten WJ, Op de Beeck BJ, Van Gaal LF, Francque SM. Sarcopenia in patients with non-alcoholic fatty liver disease: is it a clinically significant entity? Obes Rev 2019;20:353-363
38. Feldstein AE, Wieckowska A, Lopez AR, Liu YC, Zein NN, McCullough AJ. Cytokeratin-18 fragment levels as noninvasive biomarkers for nonalcoholic steatohepatitis: a multicenter validation study. Hepatology (Baltimore, Md) 2009;50:1072-1078
39. Tsutsui M, Tanaka N, Kawakubo M, Sheena Y, Horiuchi A, Komatsu M et al. Serum fragmented cytokeratin 18 levels reflect the histologic activity score of nonalcoholic fatty liver disease more accurately than serum alanine aminotransferase levels. Journal of clinical gastroenterology 2010;44:440-447

Tables Legend

Table 1. Baseline characteristics of patients with biopsy-proven NAFLD

Table 2. Baseline characteristics of patients with persistently nALT stratified by NASH status

Table 3. Performance of serum GP73 and CK18-M30 levels for diagnosing NASH in patients with persistently nALT levels

Supplementary Table 1. Performance of the G-NASH model and other predictive modalities for diagnosing NASH.

Supplementary Table 2. Baseline characteristics of all patients stratified by NASH status

Supplementary Table 3. Performance of GP73 and CK18-M30 for diagnosing NASH in patients with persistently nALT levels using a lower threshold.

Supplementary Table 4. Results of univariable and multivariable analyses: parameters associated with NASH in patients with persistently normal ALT levels

Figure legends

Figure 1. Flow diagram of the study.

Abbreviations: AFL, alcoholic fatty liver; AIH, autoimmune hepatitis; ALT, alanine aminotransferase; DIH, drug-induced hepatitis; GP73, Golgi protein 73; NAFL, nonalcoholic fatty liver; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD Activity Score; NASH, nonalcoholic steatohepatitis.

Figure 2. Receiver operating characteristic curves of various non-invasive clinical scoring systems for the prediction of NASH among NAFLD patients with persistently nALT.

(A) using a cut-off of serum ALT levels <40 U/L for both sexes. (B) using a lower cut-off of serum ALT ≤35 U/L in men (n = 48) and ≤23 U/L in women (n = 16). The y-axis represents the sensitivity in ascending order. The x-axis represents the specificity in descending order.

Abbreviations: AUROC, area under the receiver operating characteristic curve. G-NASH, model constructed by best-subset regression. Stepwise, logistic regression model constructed by backward stepwise selection. NAFIC, NAFIC score. NASH, nonalcoholic steatohepatitis. CK18-M30, cytokeratin 18 M30 fragments.

Supplementary Figure 1. Decision curves, ROC component curves and clinical impact curves. (A) Decision curves for all the predictive models and the predictors. The y-axis represents the standardized net benefit (sNB). The x-axis is composed of two parts: the threshold probability (P_t) and the cost-benefit ratio (CBR). The sNB can be interpreted as the percentage of treated cases under the condition that no controls were falsely treated under a specific P_t. CBR in decision theory equates to the odds of P_t [$CBR = P_t / (1-P_t)$]. The grey line and the black line

represent the sNB of carrying out a liver biopsy for all and for none of the patients under the different value of P_t, respectively. The G-NASH model achieved the best performance when P_t is larger than 0.22, as labelled by the dotted vertical line. (B) ROC component curve of the G-NASH model. The solid line represents the true positive result and the dotted line represents the false positive result under the different value of P_t. (C) Clinical impact curve of the G-NASH model. The curve represents a simulation of a population of 100 patients who are assessed for their risk of NASH. The solid line denotes those who were identified as potentially having NASH (high risk), and the dotted line represents those patients who had NASH confirmed by histological examination (high risk with NASH).

Supplementary Figure 2. Combined and sequential non-invasive approach to predict NASH among NAFLD patients with persistently nALT levels by using serum CK18-M30 levels and the G-NASH model. NB: Missing data for 35 patients to fully undergo the approach. Mis.: misdiagnosis.

Supplementary Figure 3. A nomogram of the G-NASH model. Six factors from the G-NASH model are shown in blue bold font and the probability of NASH calculated by the model is shown in red bold font. Every single factor in the model can be matched to the corresponding number of points using the uppermost axis and the original value/concentration for the factor in question. The sum of points for the six factors is then used to derive the probability of NASH by the Total Points axis (probability of NASH [in red font]).

Abbreviations: AST, aspartate aminotransferase; GP73, Golgi protein 73; NASH, nonalcoholic steatohepatitis; PLT, platelet count; SD, standard deviation.