Medicine®

OPEN

A nomogram for estimating the probability of nonalcoholic fatty liver disease in a Chinese population
A retrospective cohort study

Weining Xie, MD* , Shengxin Chen, MD

Abstract

Studies have showed that dyslipidemia is closely related to nonalcoholic fatty liver disease (NAFLD). However, less attention has been paid to the relationship between early dyslipidemia and long-term risk of NAFLD. Therefore, we aimed to develop a simple-to-use nomogram to predict early dyslipidemia and long-term risk of NAFLD onset.

A retrospective cohort study including 3621 employees (including retirees) from 7 companies was conducted between 2012 and 2019. Anthropometric, potential laboratory parameters and abdominal ultrasound were performed at baseline and after a 5-year follow-up. Cox proportional hazards model was used to determine predictors for NAFLD onset. The effects of lipids, age, body mass index (BMI), and serum uric acid (UA) on NAFLD were evaluated with the use of Kaplan–Meier curves (log-rank test). A nomogram was developed based on the Cox proportional hazard model and a 2-piecewise linear regression model. The accuracy of model was evaluated according to the area under the receiver operating characteristic curves.

A total of 1545 subjects were included in the final analysis. The mean follow-up time was 52 ± 6.6 months. Of the total subjects, 77.61% were male and 22.39% were female. The mean age at the time of initial visit was 45.21 ± 11.20 years. Five hundred fifty-five subjects (35.92% of all subjects) were finally diagnosed with NAFLD. Variables in the nomogram included age, BMI, triglycerides, high-density lipoprotein, low-density lipoprotein, and UA. The accuracy of the nomogram for predicting 5-year cumulative occurrence of NAFLD was 0.8135 (95% confidence interval: 0.7921–0.8349), and the sensitivity and specificity were 0.8108 and 0.6960, respectively.

The combination of age, BMI, triglycerides, high-density lipoprotein, low-density lipoprotein, and UA translated into a nomogram can reliably estimate the incidence of NAFLD within 5 years. It may serve as a decision support tool to determine whether to intervene at an early stage.

Abbreviations: γ-GT = γ-glutamyl transpeptidase, ALP = alkaline phosphatase, ALT = alanine aminotransferase, AST = aspartate aminotransferase, BMI = body mass index, CI = confidence interval, DBP = diastolic blood pressure, FBG = fasting blood glucose, GAM = generalized additive model, HDL = high-density lipoprotein, LDL = low-density lipoprotein, NAFLD = nonalcoholic fatty liver disease, NASH = non-alcoholic steatohepatitis, OR = odds ratio, ROC = receiver operating characteristic, SBP =systolic blood pressure, TBA = total bile acid, TC = total cholesterol, TG = triglycerides, UA = uric acid.

Keywords: dyslipidemia, nomogram, nonalcoholic fatty liver disease

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is an important issue of global public health in the 21st century.[1,2] Multiple studies have shown that the global prevalence of NAFLD is estimated to be 6% to 35%, with a median of 20% for the general public.[3] The disease affects 15% to 20% of adults in China and continues to increase due to the prevalence of overweight and obesity in the Chinese population.[4] NAFLD covers a range of liver diseases, ranging from simple steatosis to non-alcoholic steatohepatitis (NASH), with varying degrees of fibrosis, which may eventually develop into cirrhosis.[5–7] Simple steatosis is considered benign and progresses slowly over the years, while NASH may develop into cirrhosis and hepatocellular carcinoma.[8–11] Although the effects of NAFLD on cardiovascular diseases are controversial,[12,13] the incidence of NAFLD may significantly increase the risk of T2DM and chronic kidney disease.[14,15]

Therefore, it is of utmost importance to identify those who may further develop NAFLD. Early intervention in these "at risk"

patients may prevent the development of NAFLD through education on healthy diet, physical activity, weight loss, and prevention of weight gain.[16–18]

Dyslipidemia is a well-documented influencing factor of NAFLD, which is usually characterized by elevated levels of low-density lipoprotein (LDL) cholesterol and triglycerides (TG) or decreased levels of high-density lipoprotein (HDL).[19,20] Excess fat accumulates in hepatocytes in the form of lipid droplets, which are coated by several structural proteins that may be involved in the pathophysiology of liver diseases.[21–23] This intrahepatic lipid accumulation is caused by abnormal lipid metabolism, such as increased systemic lipolysis, increased hepatic free fatty acid uptake and very LDL synthesis, and decreased free fatty acid oxidation and TG export.[24,25] Serum uric acid (UA), an end product of purine metabolism in the liver. Recently, some evidence suggests that elevated UA is also associated with the development or progression of NAFLD.[26–29]

Previous studies have shown that age, body mass index (BMI), TG, total cholesterol (TC), HDL, LDL, and other factors are closely related to NAFLD.[14,30–32] However, most traditional factor analyses based on cross-sectional studies revealed risk factors for NAFLD.[33] Previous cross-sectional studies have not been able to determine whether dyslipidemia is a cause or a consequence of NAFLD. In addition, some studies have shown that lipid abnormalities were not detected before NAFLD.[34] In particular, there are conflicting data on the role of individual lipids in promoting hepatic fat accumulation.[35]

In addition, less attention has been paid to the relationship between early dyslipidemia and the long-term risk of NAFLD in general population. Therefore, in order to prevent NAFLD, more attention should be paid to patients with dyslipidemia at an early stage. Our goal was to develop a simple-to-use nomogram to predict early dyslipidemia and long-term risk of NAFLD onset. These tools will help provide early evidence for NAFLD and thus effectively prevent NAFLD.

2. Materials and methods

2.1. Study population

This was a retrospective observational cohort study involving 3621 employees (including retirees) from 7 companies who regularly underwent annual health examination at the Health Examination Center of Guangdong Integrated Chinese and Western Medicine Hospital. Data were collected between 2012 and 2019. In the cohort study, 2274 subjects without NAFLD were initially recruited. Exclusion criteria at baseline included:

1) alcoholism (male>140 g/w, female >70 g/w);
2) subjects with incomplete clinical data or lost to follow-up;
3) impaired fasting glucose (>110 mg/dL), impaired glucose tolerance or known type 2 diabetes;
4) medical history, including viral hepatitis, autoimmune hepatitis, and other known chronic liver diseases;
5) hypertension, defined as blood pressure (≥140/90 mm Hg) or hypertension under treatment. Eventually, 1545 subjects were included in the analysis (Fig. 1). The study protocol was conducted in accordance with the Declaration of Helsinki and all data used in this study were approved by the institutional
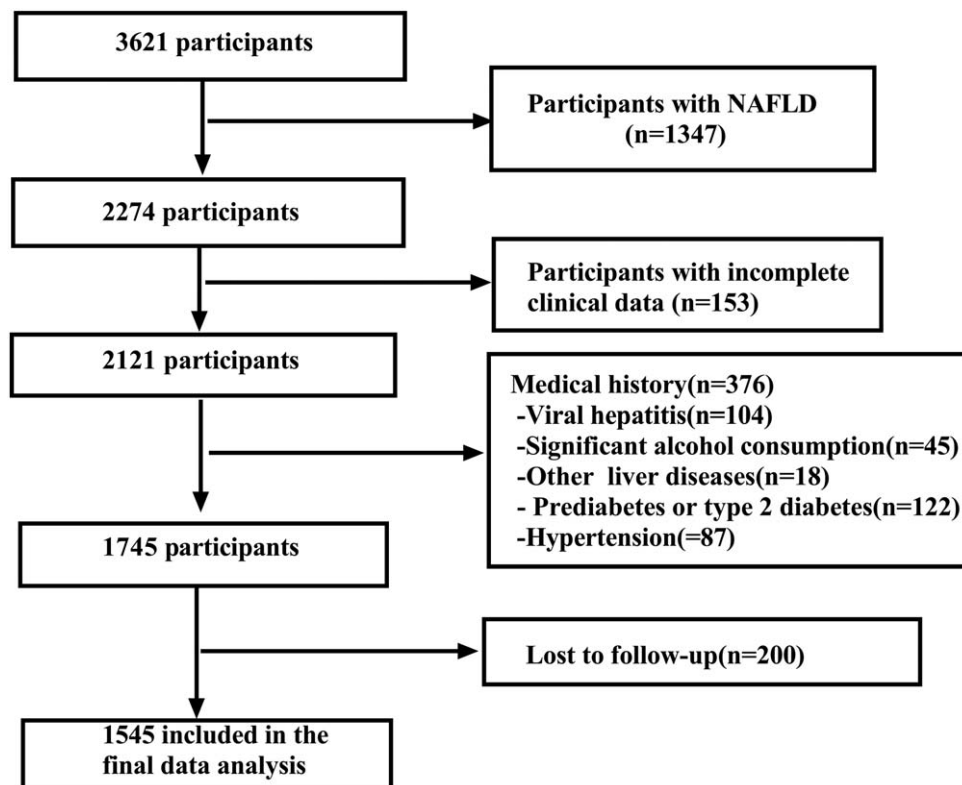
Figure 1. The flow chart of population selection. In total, 3621 employees (including retirees) were recruited from the Foshan city of Guangdong Province. Among them, 1347 participants with nonalcoholic fatty liver disease were excluded at baseline. After excluding 729 participants who met at least 1 of the exclusion criteria. the final analysis sample of 1545 participants was obtained.

review board of Guangdong Provincial Hospital of Integrated Traditional Chinese and Western Medicine.

2.2. Data collection

The trained staff would deliver the standardized spreadsheet to collect the general information, such as age, gender, height, weight, systolic blood pressure (SBP), diastolic blood pressure (DBP), medical history, etc BMI was calculated as weight/height^2. Data collection in subjects included laboratory evaluation of TG, TC, HDL, LDL, fasting blood-glucose (FBG), Alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), UA, and γ-glutamyl transpeptidase (γ-GT), total bile acid (TBA). The assays above were conducted on an Olympus AU2700 biochemical auto-analyzer (Olympus, Japan). NAFLD was diagnosed by abdominal ultrasonography (ACU-SON X700 ultrasound system, Siemens, Japan).

2.3. Definition

The diagnostic criteria for NAFLD were in accordance with the Chinese Association of Liver Diseases.[36] In general, NAFLD was diagnosed as at least 2 of 3 abnormalities, namely diffuse hyperechoicity of the liver relative to the spleen and kidney, attenuated ultrasound beams and poor visibility of intrahepatic structural details.

2.4. Follow-Up

All participants were followed up every 9 to 15 months at the Department of Health Examination Center. Within each follow-up, all participants received laboratory tests (such as TG, TC, HDL, LDL, FBG, ALT, AST, ALP, UA, γ-GT, TBA) and abdominal ultrasonography. The cut-off time was December 2019.

2.5. Statistical analysis

Continuous variables were expressed as mean standard deviation or median inter-quartile range, while categorical values were expressed using relative frequencies and proportions. Comparisons of parameters between 2 different groups were conducted with the Student t test and the Mann–Whitney U test for continuous variables with or without normal distribution and with the Chi-squared test for categorical variables. Univariate and multivariate analyses were then performed to define independent factors associated with NAFLD. Besides, we used generalized additive model (GAM) to identify the non-linear relationship between continuous variable and NAFLD according to smooth curve fitting. If the non-linear correlation was observed, a 2-piecewise linear regression model was performed to calculate the threshold effect of continuous variable on NAFLD in terms of the smoothing plot. We used a likelihood ratio test comparing the 1-line linear regression model with a 2-piecewise linear model, when the likelihood ratio test was statistically significant ($P < .05$), a 2-piecewise linear model was used to analyze the relationship between the continuous variables and NAFLD. The effects of lipids, age, BMI, and UA categories on NAFLD were evaluated with the use of Kaplan–Meier curves (log-rank test). Cox proportional hazards model and a 2-piecewise linear regression were performed to estimate the odds ratio (OR) and 95% confidence interval (CI). The nomogram was

developed based on Cox proportional hazards model and a 2-piecewise linear regression model, which allowed us to obtain NAFLD probability estimations. An automated stepwise variable selection method performed on 500 bootstrap samples was used to choose the final model. An area under the receiver operator characteristic (ROC) curve was used as a measure of the diagnostic accuracy. All of the analyses were performed with the statistical software packages R (http://www.R-project.org, The R Foundation) and EmpowerStats (http://www.empowerstats.com, X&Y Solutions, Inc., Boston, MA). A 2-tailed P value $< .05$ was considered statistically significant.

3. Results

3.1. Demographic and clinical characteristics

A total of 1545 subjects were included in the final analysis. The mean follow-up time was 52 ± 6.6 months. Of the total subjects, 77.61% were male and 22.39% were female. The mean age at the time of initial visit was 45.21 ± 11.20 years. Five hundred fifty-five subjects (35.92% of all subjects) were finally diagnosed with NAFLD. Table 1 presented a summary of patient anthropometric, clinical, and laboratory characteristics. Participants with NAFLD had higher BMI, SBP, and DBP. Serum levels of TC, TG, LDL, FBG, ALP, ALT, AST, UA, and γ-GT were significantly higher in participants with NAFLD than in those without NAFLD (all $P < .05$). However, Serum levels of HDL were significantly lower in participants with NAFLD compared to those without NAFLD (all $P < .001$), and levels of serum TBA were found to be non-associated with the presence of NAFLD ($P = .160$).

3.2. Results of univariate and multivariate analysis for risk factors of NAFLD

A univariate analysis was conducted to evaluate the associations between exposure factors and NAFLD. The results of univariate analysis showed that the levels of TG, TC, LDL, HDL, FBG, γ-GT, ALT, AST, UA, TC, LDL, TG, and ALP were significantly associated with NAFLD. However, TBA was not associated with NAFLD. After adjusting for gender, age, ALT, AST, ALP, UA, TC, HDL, LDL, TG, BMI, SBP, DBP, FBG, and γ-GT, Cox proportional hazards model showed that the risk factors for NAFLD included: ALT (HR=1.043, 95CI%: 1.030–1.055, $P < .001$), AST (HR=0.952, 95CI%: 0.931–0.975, $P < .001$), UA (HR=1.908, 95CI%: 1.469–2.479, $P < .001$), TC (HR=1.426, 95CI%: 1.162–1.751, $P = .001$), γ-GT (HR=1.007, 95CI%: 1.003–1.011, $P = .002$), FBG (HR=1.206, 95CI%: 1.076–1.350, $P < .001$), BMI (HR=1.46, 95CI%: 1.357–1.583, $P < .001$). However, HDL was an independent protective factor for NAFLD (HR=0.405, 95CI%: 0.311–0.527, $P < .001$) (Table 2).

3.3. Generalized additive model analysis of risk factors for NAFLD

It was essential to analyze the nonlinear relationship between continuous variables and NAFLD, We then entered all of continuous variable (AGE, ALT, AST, ALP, UA, TC, HDL, LDL, ALB, TG, γ-GT, FGB, DBP, SBP, and BMI) into a generalized additive modelGAM. After adjusting for the aforementioned covariates, there was a non-linear relationship between age, UA, TG, HDL, LDL, BMI, and NAFLD (all the likelihood ratio

Table 1

Demographic and clinical characteristics of subjects.

Variables	All	Non-NAFLD	NAFLD	P-value
N	1545	990	555	
Gender, male, %	1199 (77.61%)	722 (72.93%)	477 (85.95%)	<.001
Age, yr	45.21±11.20	44.75±11.80	46.05±10.01	.029
ALT, U/L	21.00 (16.00–32.00)	19.00 (14.00–26.00)	29.00 (21.00–42.75)	<.001
AST, U/L	21.00 (18.00–25.00)	20.00 (17.00–24.00)	23.00 (19.00–28.00)	<.001
ALP, mmol/L	91.39±18.00	89.68±17.90	94.43±17.79	<.001
UA, μmol/L	413.18±99.17	390.53±93.82	453.57±95.64	<.001
TC, mmol/L	5.58±1.15	5.44±1.03	5.83±1.31	<.001
HDL, mmol/L	1.44±0.33	1.51±0.33	1.31±0.27	<.001
LDL, mmol/L	3.39±0.87	3.32±0.84	3.52±0.92	<.001
TBA, mmol/L	3.64 (2.50–4.98)	3.75 (2.70–4.93)	3.47 (2.12–5.00)	.160
TG, mmol/L	1.55 (1.08–2.45)	1.32 (0.95–1.85)	2.25 (1.57–3.45)	<.001
γ-GT, mmol/L	30.00 (21.00–46.00)	25.00 (19.00–36.00)	41.00 (29.00–63.00)	<.001
FBG, mmol/L	5.38±1.42	5.16±1.03	5.77±1.86	<.001
DBP, mm Hg	121.28±13.42	118.46±12.89	126.31±12.87	<.001
SBP, mm Hg	75.03±9.05	72.84±8.27	78.93±9.07	<.001
BMI, kg/m^2	24.32±2.77	23.28±2.41	26.18±2.39	<.001

γ-GT = γ-glutamyl transpeptidase, ALP = alkaline phosphatase, ALT = alanine aminotransferase, AST = aspartate aminotransferase, BMI = body mass index, DBP = diastolic blood pressure, FBG = fasting blood glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, NAFLD = nonalcoholic fatty liver disease, SBP = systolic blood pressure, TBA = total bile acid, TC = total cholesterol, TG = Triglyceride, UA = uric acid.

tests<0.05), while the relationship between γ-GT, ALT, AST, FBG, TC, and NAFLD was a linear relationship (all $P < .05$). A 2-piecewise linear regression analysis revealed that age<47 (HR: 1.04, 95%CI: 1.01–1.07, $P = .008$); UA<500mmol/L (HR = 1.0043, 95%CI: 1.002–1.007, $P = .0002$); TG<2.7mmol/L (HR = 2.129, 95%CI: 1.705–2.659, $P<.0001$); LDL≥2.67 (HR = 1.50, 95% CI: 1.31–1.77, $P < .0001$); BMI<24 (HR = 2.70, 95% CI: 2.24–3.25, $P < .001$) and BMI≥24 (HR = 1.80, 95%CI: 1.42–2.28, $P < .001$) were the independent risk factors of NAFLD, while HDL≥1.14 (HR = 0.07, 95%CI: 0.04–0.12, $P<.0001$). Age≥47 (HR = 0.973, 95%CI: 0.997–0.0239, $P = .0239$) and

LDL<2.67 (HR = 0.61, 95% CI: 0.38–0.97, $P = .0384$) were independent protective factors of NAFLD (Fig. 2, Table 3).

3.4. Kaplan–Meier curves analysis for 5-year cumulative incidence of NAFLD

On the basis of a 2-piecewise linear regression model analysis, we found that the optimal inflection points of age, UA, HDL, LDL, TG, BMI was 47 years, 500 μmol/L, 1.14 mmol/L, 2.67 mmol/L, 2.7mmol/L and 24 kg/m^2 respectively. Therefore, according to the optimal Inflection point, patients were classified into the

Table 2

The results of univariate and multivariate analysis.

Exposures	Univariate analysis		Multivariate analysis	
	HR (95%CI)	P-value	HR (95%CI)	P-value
Gender				
male	1.0			
female	0.44 (0.33, 0.58)	<.0001		
Age, yr	1.01 (1.00, 1.02)	.0292		
ALT, U/L	1.04 (1.03, 1.05)	<.0001	1.043 (1.030, 1.055)	<.001
AST, U/L	1.03 (1.02, 1.04)	<.0001	0.952 (0.931, 0.975)	<.001
ALP, mmol/L	1.01 (1.01, 1.02)	<.0001		
UA, μmol/L	1.01 (1.01, 1.01)	<.0001	1.908 (1.469, 2.479)	<.001
TC, mmol/L	1.36 (1.23, 1.50)	<.0001	1.426 (1.162, 1.751)	.001
HDL, mmol/L	0.10 (0.07, 0.16)	<.0001	0.405 (0.311, 0.527)	<.001
LDL, mmol/L	1.31 (1.16, 1.48)	<.0001		
TBA, mmol/L	1.00 (0.98, 1.02)	.8316		
TG, mmol/L	1.63 (1.49, 1.79)	<.0001		
γ-GT, mmol/L	1.02 (1.01, 1.02)	<.0001	1.007 (1.003, 1.011)	.002
FBG, mmol/L	1.50 (1.33, 1.68)	<.0001	1.206 (1.076, 1.350)	.001
DBP, mm Hg	1.05 (1.04, 1.06)	<.0001		
SBP, mm Hg	1.09 (1.07, 1.10)	<.0001		
BMI, kg/m^2	1.74 (1.63, 1.85)	<.0001	1.46 (1.357, 1.583)	<.001

γ-GT = γ-glutamyl transpeptidase, ALP = alkaline phosphatase, ALT = alanine aminotransferase, AST = aspartate aminotransferase, BMI = body mass index, CI = confidence interval, DBP = diastolic blood pressure, FBG = fasting blood glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, NAFLD = Nonalcoholic fatty liver disease, SBP = systolic blood pressure, TBA = total bile acid, TC = total cholesterol, TG = Triglyceride, UA = uric acid.
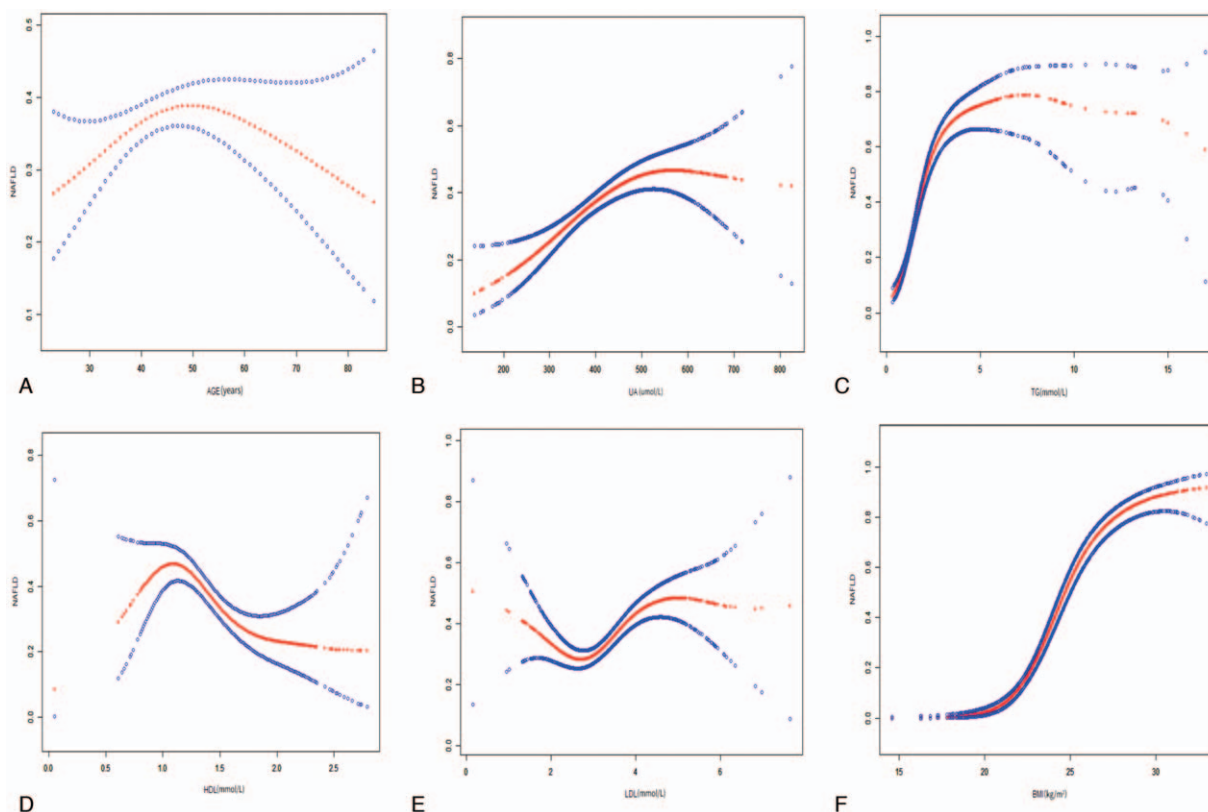
Figure 2. The relationship between age, uric acid (UA), high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides (TG), body mass index (BMI), and nonalcoholic fatty liver disease (NAFLD). A non-linear relationship between them was detected after adjusting for gender, age, alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, UA, total cholesterol, HDL, LDL, TG, BMI, systolic blood pressure, diastolic blood pressure, fasting blood-glucose, and γ-glutamyl transpeptidase. (A) age and NAFLD; (B)UA and NAFLD; (C)TG and NAFLD; (D)HDL and NAFLD; (E)LDL and NAFLD; (F)BMI and NAFLD.

different categories. The cumulative hazards of NAFLD were significantly higher in the age≧47 years, UA≧500 μmol/L, TG≧2.7mmol/L, HDL<1.14mmol/L, LDL≧2.67mmol/L and BMI ≧24 kg/m^2 groups ($P < 0.001$ for all; Fig. 3).

3.5. Predictive accuracy of nomogram for NAFLD

Risk factors based on the Cox proportional hazard model and a 2 piecewise linear regression analysis, a nomogram for evaluating the risk of NAFLD was developed for participants. The final

---

Table 3

The results of 2-piecewise linear regression model.

Variables	Inflection point	Effect size (HR)	95%CI		P-value
			Lower	Upper	
Age, yr	<47	1.04	1.01	1.07	.0080
	≧47	0.973	0.950	0.997	.0239
UA, μmol/L	<500	0.997	0.9929	1.0014	.1895
	≧500	1.004	1.0021	1.0066	.0002
TG, mmol/L	<2.70	2.129	1.705	2.659	<.0001
	≧2.70	0.959	0.861	1.070	.4549
HDL, mmol/L	<1.14	0.72	0.15	3.40	.682
	≧1.14	0.07	0.04	0.12	<.0001
LDL, mmol/L	<2.67	0.61	0.38	0.97	.0384
	≧2.67	1.50	1.31	1.77	<.0001
BMI, kg/m^2	<24	2.70	2.24	3.25	<.0001
	≧24	1.80	1.42	2.28	<.0001

Effect: NAFLD; Cause: age, UA, TG, HDL, LDL, and BMI.
Adjusted: gender, age, ALT, AST, ALP, UA, TC, HDL, LDL, TG, BMI, SBP, DBP, FBG, and γ-GT.
γ-GT = γ-glutamyl transpeptidase, ALP = alkaline phosphatase, ALT = alanine aminotransferase, AST = aspartate aminotransferase, BMI = body mass index, CI = confidence interval, DBP = diastolic blood pressure, FBG = fasting blood glucose, HDL = high-density lipoprotein, LDL = low-density lipoprotein, NAFLD = Nonalcoholic fatty liver disease, SBP = systolic blood pressure, TBA = total bile acid, TC = total cholesterol, TG = Triglyceride, UA = uric acid.
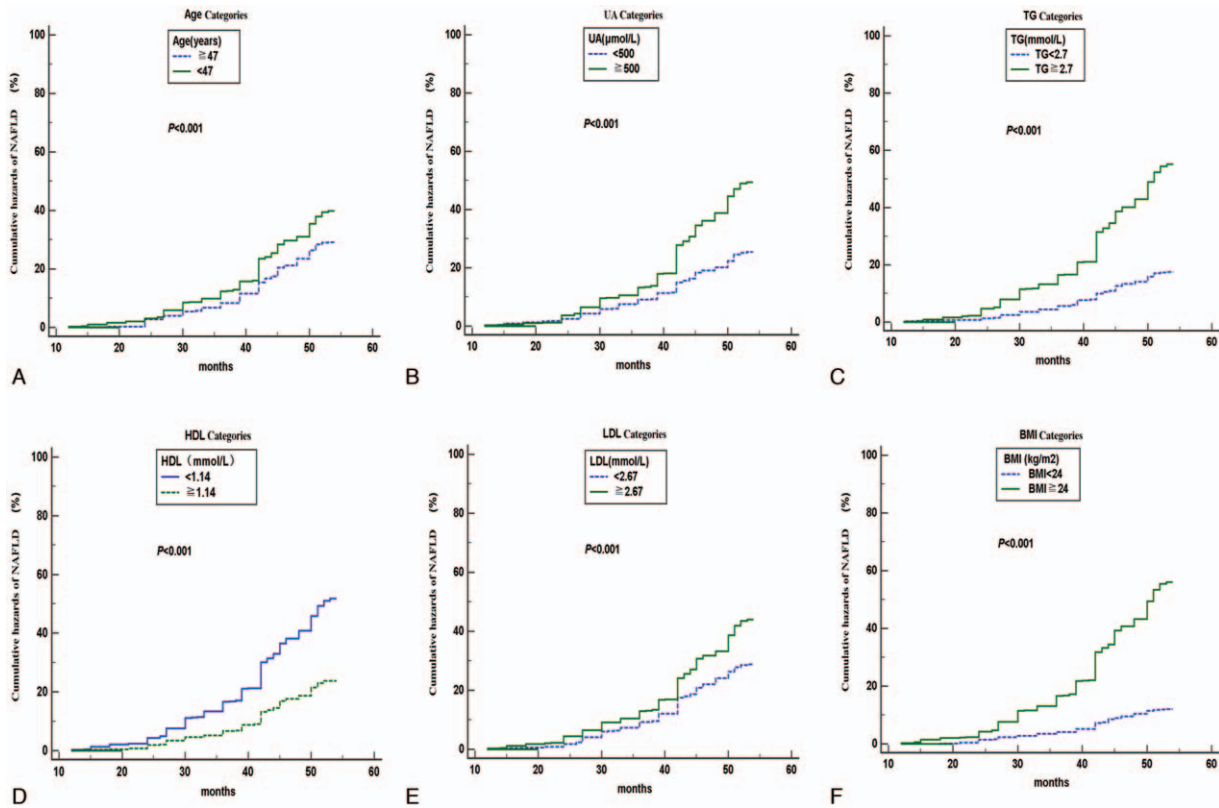
Figure 3. Kaplan–Meier curves of the cumulative incidence rate of nonalcoholic fatty liver disease stratified by the following categories: age (years)<47 or ≥47 (Fig. 2A); UA (μmol/L)<500 or≥500 (Fig. 2B); triglycerides (mmol/L)<2.7 or ≥2.7 (Fig. 2C); high-density lipoprotein (mmol/L)<1.14 or ≥1.14 (Fig. 2D); low-density lipoprotein (mmol/L)<2.27 or ≥2.27 (Fig. 2E); body mass index (kg/m²)<24 or ≥24 (Fig. 2F).

proposed model included age, UA, ALT, LDL, TG, and BMI. Six independent predictors with the best performance were incorporated into our model and presented in a nomogram (Fig. 4). The nomogram is used by locating the position on each predictor variable scale according to the value of each participant. Each scale location has a corresponding prognostic point (top axis). The point value of each predictor variable is determined first, and then the total point value (located on the bottom axis) is calculated. Where a straight line is drawn to determine the probability of having NAFLD. Calibration curves were drawn to determine whether the predicted and observed probabilities of NAFLD were consistent. Bootstrap resampling (500 resamples) was used for this plot. The accuracy of the nomogram for predicting 5-year cumulative occurrence of NAFLD was 0.8135 (95% CI: 0.7921–0.8349), and the sensitivity and specificity were 0.8108 and 0.6960, respectively (Figs. 5 and 6).

4. Discussion

Studies have found that NAFLD is an important risk factor for advanced liver disease[1,37,38] and non-liver-related diseases such as diabetes[39,40] and cardiovascular diseases.[41] Therefore, various attempts have been made to diagnose NAFLD.[42–45] However, identification of high-risk individuals may be more valuable than diagnosis. If we can predict these high-risk subjects, we can prevent disease onset by improving lifestyle, such as improving nutrition and increasing exercise.[46,47] To help meet the ongoing need to identify more useful, simple and practical

predictors of NAFLD incidence in general practice, we conducted the current retrospective cohort study.

Our investigation demonstrated the prediction effects of metabolic markers for detection of NAFLD among a Chinese people. The research findings showed that the subjects with increased TG, HDL, UA, BMI, FBG, γ-GT, BMI, and age may have higher risks and cumulative incidence of NAFLD. A nomogram was constructed to evaluate 5-years cumulative incidence of NAFLD, consisting of variables of age, UA, LDL, HDL, TG, and BMI. The accuracy of the nomogram for predicting 5-year cumulative occurrence of NAFLD was 0.8135 (95% CI: 0.7921–0.8349), and the sensitivity and specificity were 0.8108 and 0.6960, respectively.

Variables of the constructed nomogram evaluation system were easily available in clinical practice. Namely, we provided a nomogram to identify which patient was more likely to develop NAFLD and possibly for appropriate prevention at an early stage. Our study may help identify high-risk subjects for specific preventive measures. The calibration results confirmed that the model can be used to detect NAFLD.

Several studies have already reported the combination of biochemical and clinical markers to predict NAFLD. Hamaguchiet al[39] studied 3147 participants attending hospital health checkups and found that after an average of 414 days of follow-up, men and women who met the metabolic syndrome criteria at baseline were more likely to have NAFLD. Yu et al[48] showed in a prospective study of 7000 employees in China that UA and serum hemoglobin had significantly high predictive value for NAFLD. In addition, Chang et al[49] conducted a population-
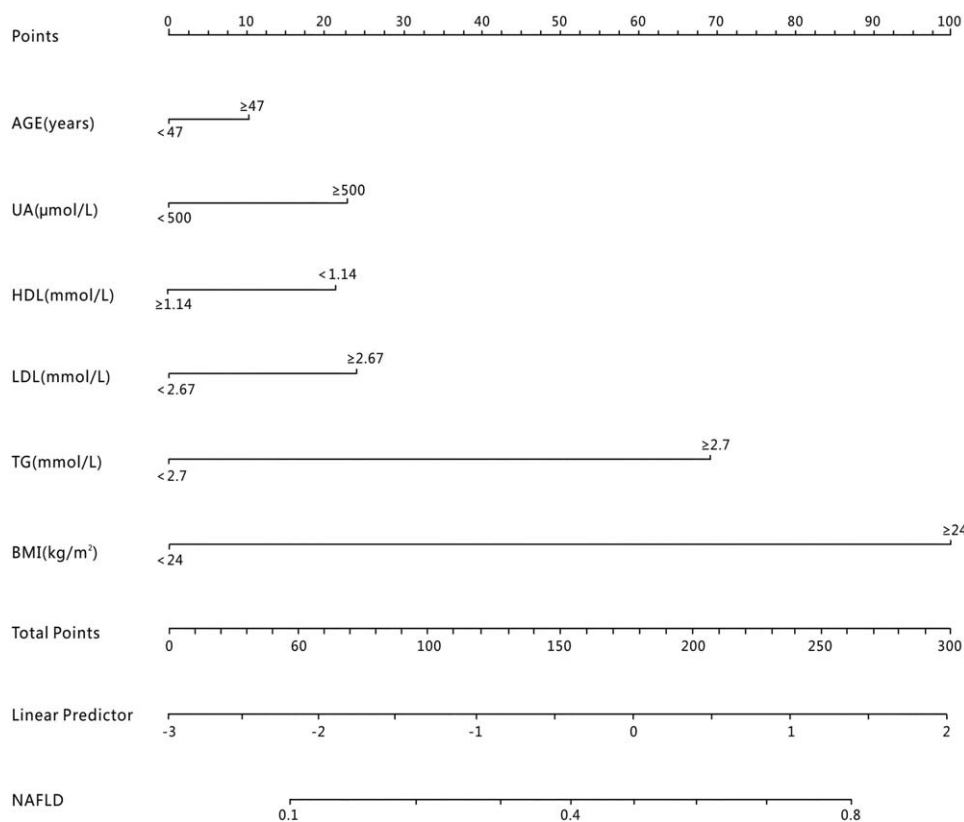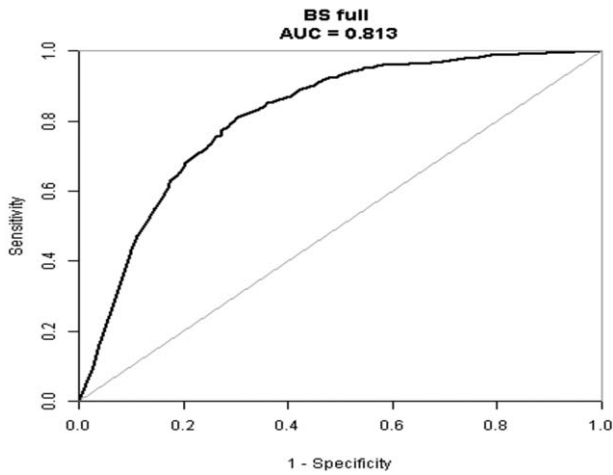
Figure 4. Nomogram for predicting nonalcoholic fatty liver disease. Values for each variable are individually plotted and correspond to point values assigned from the point scale (top). These point values are then totaled and plotted on the total point scale (bottom), which is used to assign a corresponding value for risk of nonalcoholic fatty liver disease.

based study of 5237 healthy men and found that higher ALT levels within the reference range at baseline were independent risk factors for NAFLD onset. Zheng et al[30] found that elevated serum UA levels were positively correlated with lean-NAFLD independent of other metabolic factors. The ROC curve for detection of mild fatty liver by ultrasonography was 0.735. Miyake T et al[42] showed that BMI, ALT, TG, UA, and HbA1c as well as low HDL-C levels were risk factors for NAFLD, and the ROC predicted for NAFLD was 0.807 (0.790–0.824) in man and 0.748 (0.735–0.761) in woman, respectively. Another study[50] showed that triglyceride to HDL cholesterol ratio had better predictive effects of onset of NAFLD than other lipid index. The ROC of TG/HDL-C ratio in the male was 0.70 (0.68–0.72) and 0.72 (0.70–0.75) in the female. The above research showed that dislipidemia, UA, BMI and so on were effective indexes for predicting NAFLD, which were similar to our results. However, those above were cross-sectional studies or cohort studies with short follow-up time, which cannot explain the long-term results. In addition, logistic or Cox proportional hazards model analysis was used to analyze the effect of dyslipidemia on NAFLD in most of studies above. However, in our study, we not only used logistic and Cox proportional hazards models to evaluate the relationship between dyslipidemia and NAFLD, but also used the generalized additive modelGAM to clarify the nonlinear relationship, and it was found that age, UA, TG, HDL, LDL, and BMI were non-linear correlated with NAFLD. On the basis of a 2-piecewise linear regression model analysis, we found that the optimal inflection point of age, UA, HDL, LDL, TG, BMI was

47 years, 500 μmol/L, 1.14 mmol/L, 2.67 mmol/L, 2.7 mmol/L, and 24 kg/m² respectively. And the cutoff values of predictive factors above were useful for clinicians in guiding subjects with dyslipidemia. Therefore, GAM has obvious advantages in dealing with non-linear relations and it can handle the non-parametric smoothing and will fit a regression spline to the data. The use of GAM will help us to better discover the real relationships between exposures and outcomes.

Recently, few studies have been conducted recently using a nomogram to establish a risk prediction model for the development of NAFLD.[51,52] However, they also focus on cross-sectional study, which was hard to infer causality. Nomogram analysis in our study showed that age, UA, TG, LDL, HDL and BMI had better sensitivity and specificity in early prediction of 5-year cumulative incidence of NAFLD. The nomogram helps to meet our needs for visual tools and personalized prevention. Compared with traditional prediction models, nomograms have a visual digital interface, higher accuracy and easier understanding of risk prediction, and can be more directly applied to clinical decision-making. Nomograms are simple, rapid and inexpensive and can be used to effectively identify the risk of developing NAFLD.[53] After diagnosis, medical interventions, lifestyle changes, diagnostic management and treatment can be initiated. According to the calculation results, prediction models of age, UA, LDL, HDL, TG, and BMI can be established for clinical practice.

However, some limitations need to be mentioned in this study. First, the study was conducted retrospectively in a single-center,

Figure 5. Receiver operating characteristic curves for the accuracy of the nonalcoholic fatty liver disease (NAFLD). Nomogram in patients with NAFLD. The overall predictive accuracy of the nomogram for NAFLD was 0.8135 (95% CI:0.7921–0.8349), and the sensitivity and specificity were 0.8108 and 0.6960, respectively. AUC = area under the curve, CI = confidence interval, ROC = receiver operating characteristic.

so some inevitable bias must be noted. Second, abdominal ultrasonography is not the perfect gold standard examination for NAFLD and may be missed diagnosis. Third, this patient cohort was from 7 companies, thus these findings may not be applicable to the general population in the community. Fourth, during the

follow-up, the patient's medication status was unclear, which would affect the final prediction results. Fifth, most of the subjects included were male, and therefore, gender factors may be a potential bias leading to the nomogram score developed.

Despite these limitations, our study produced some strengths. In particular, to limit heterogeneity in our study, we used the following methods. First, the selection of all participants was done by 2 independent researchers according to predefined inclusion and exclusion criteria and discrepancies were resolved through discussion between these reviewers. Second, in the existing literature on predicting the risk of NAFLD, most of the statistical methods used are logistic or Cox proportional hazard model analysis.[48–50] However, in our study, we not only used logistic and Cox proportional hazard models, the generalized additive model was also used to clarify the non-linear relationship, which may reduce the heterogeneity caused by statistical methods.

5. Conclusions

The combination of age, BMI, TG, HDL, LDL, and UA translated into a nomogram can reliably estimate the incidence of NAFLD within 5 years. It may serve as a decision support tool to determine whether to intervene at an early stage.

Acknowledgments
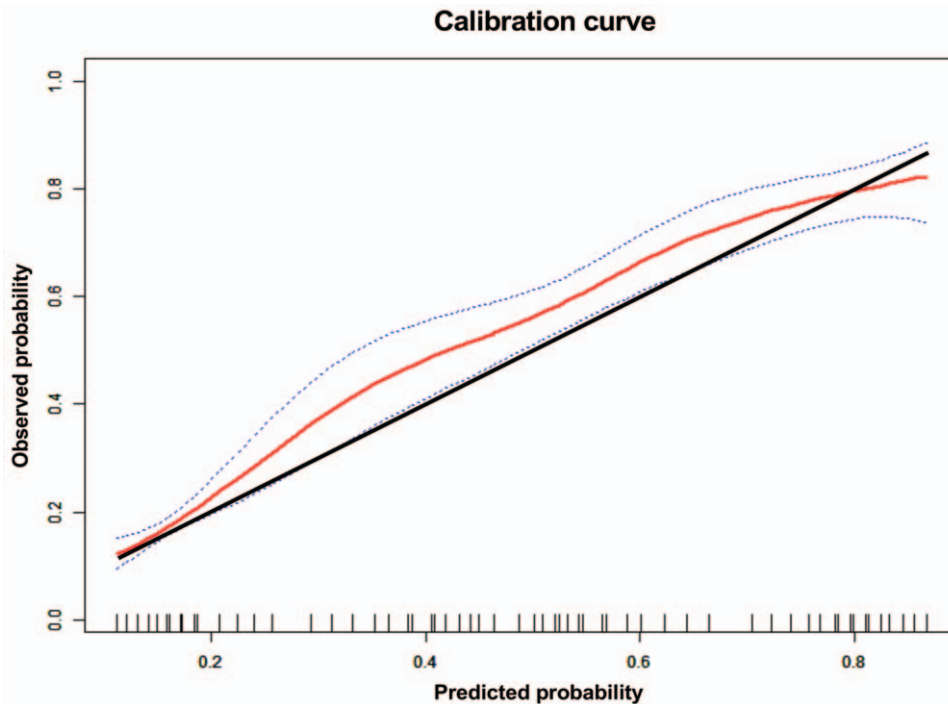
Figure 6. Calibration curves of the nomogram in the cohort. The black line indicates perfect prediction. The red line means the predictive performance of the nomogram. The blue region represents a 95% confidence interval of the nomogram.

Author contributions

References

[1] Angulo P. Nonalcoholic fatty liver disease. N Engl J Med 2002;346: 1221–31.
[2] Clark JM, Brancati FL, Diehl AM. The prevalence and etiology of elevated aminotransferase levels in the United States. Am J Gastroenterol 2003;98:960–7.
[3] Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 2011;34:274–85.
[4] Fan JG, Farrell GC. Epidemiology of non-alcoholic fatty liver disease in China. J Hepatol 2008;50:204–10.
[5] Caldwell S, Argo C. The natural history of non-alcoholic fatty liver disease. Dig Dis 2010;28:162–8.
[6] Farrell GC, Larter CZ. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. Hepatology 2006;43:99–112.
[7] Matteoni CA, Younossi ZM, Gramlich T, et al. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology 1999;116:1413–9.
[8] Fassio E, Alvarez E, Domínguez N, et al. Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. Hepatology 2004;40:820–6.
[9] Ekstedt M, Franzén LE, Mathiesen UL, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology 2006; 44:865–73.
[10] Söderberg C, Stål P, Askling J, et al. Decreased survival of subjects with elevated liver function tests during a 28-year follow-up. Hepatology 2010;51:595–602.
[11] Starley BQ, Calcagno CJ, Harrison SA. Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. Hepatology 2010;51:1820–32.
[12] Targher G, Day CP, Bonora E. Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. N Engl J Med 2010;363: 1341–50.
[13] Alexander M, Loomis AK, Lei JVD, et al. Non-alcoholic fatty liver disease and risk of incident acute myocardial infarction and stroke: findings from matched cohort study of 18 million European adults. BMJ 2019;367:l5367.
[14] Sung KC, Kim SH. Interrelationship between fatty liver and insulin resistance in the development of type 2 diabetes. J Clin Endocrinol Metab 2011;96:1093–7.
[15] Targher G, Bertolini L, Rodella S, et al. Non-alcoholic fatty liver disease is independently associated with an increased prevalence of chronic kidney disease and proliferative/laser-treated retinopathy in type 2 diabetic patients. Diabetologia 2008;51:444–50.
[16] Zelber-Sagi S, Lotan R, Shlomai A, et al. Predictors for incidence and remission of NAFLD in the general population during a seven-year prospective follow-up. J Hepatol 2012;56:1145–51.
[17] Kistler KD, Brunt EM, Clark JM, et al. Physical activity recommendations, exercise intensity, and histological severity of nonalcoholic fatty liver disease. Am J Gastroenterol 2011;106:460–8.
[18] Ryan MC, Itsiopoulos C, Thodis T, et al. The Mediterranean diet improves hepatic steatosis and insulin sensitivity in individuals with non-alcoholic fatty liver disease. J Hepatol 2013;59:138–43.
[19] Yang SH, Du Y, Li XL, et al. Triglyceride to high-density lipoprotein cholesterol ratio and cardiovascular events in diabetics with coronary artery disease. Am J Med Sci 2017;354:117–24.
[20] Wu L, Parhofer KG. Diabetic dyslipidemia. Metabolism 2014;63: 1469–79.
[21] Carr RM, Ahima RS. Pathophysiology of lipid droplet proteins in liver diseases. Exp Cell Res 2016;340:187–92.
[22] Mashek DG, Khan SA, Sathyanarayan A, et al. Hepatic lipid droplet biology: getting to the root of fatty liver. Hepatology 2015;62: 964–7.
[23] Sahini N, Borlak J. Recent insights into the molecular pathophysiology of lipid droplet formation in hepatocytes. Prog Lipid Res 2014;54:86–112.
[24] Almeda-Valdes P, Altamirano-Barrera A, Méndez-Sánchez N. Insights in non-alcoholic fatty liver disease pathophysiology with lipidomic analyses. Ann Hepatol 2015;14:567–9.
[25] Ress C, Kaser S. Mechanisms of intrahepatic triglyceride accumulation. World J Gastroenterol 2016;22:1664–73.
[26] Xu C, Yu C, Xu L, et al. High serum uric acid increases the risk for nonalcoholic fatty liver disease: a prospective observational study. PLoS One 2010;5:e11578.
[27] Ryu S, Chang Y, Kim SG, et al. Serum uric acid levels predict incident nonalcoholic fatty liver disease in healthy Korean men. Metabolism 2011;60:860–6.
[28] Li Y, Xu C, Yu C, et al. Association of serum uric acid level with non-alcoholic fatty liver disease: a cross-sectional study. J Hepatol 2009;50:1029–34.
[29] Afzali A, Weiss NS, Boyko EJ, et al. Association between serum uric acid level and chronic liver disease in the United States. Hepatology 2010;52:578–89.
[30] Zheng X, Gong L, Luo R, et al. Serum uric acid and non-alcoholic fatty liver disease in non-obesity Chinese adults. Lipids Health Dis 2017;16:202.
[31] Speliotes EK, Massaro JM, Hoffmann U, et al. Fatty liver is associated with dyslipidemia and dysglycemia independent of visceral fat: the Framingham heart study. Hepatology 2010;51:1979–87.
[32] Nseir W, Shalata A, Marmor A, et al. Mechanisms linking nonalcoholic fatty liver disease with coronary artery disease. Dig Dis Sci 2011; 56:3439–49.
[33] Defilippis AP, Blaha MJ, Martin SS, et al. Nonalcoholic fatty liver disease and serum lipoproteins: the multi-ethnic study of atherosclerosis. Atherosclerosis 2013;227:429–36.
[34] Bedogni G, Miglioli L, Masutti F, et al. Incidence and natural course of fatty liver in the general population: the Dionysos study. Hepatology 2007;46:1387–91.
[35] Neuschwander-Tetri BA. Hepatic lipotoxicity and the pathogenesis of nonalcoholic steatohepatitis: the central role of non-triglyceride fatty acid metabolites. Hepatology 2010;52:774–88.
[36] Jian GF, Ji DJ, You ML, et al. Guidelines for the diagnosis and management of nonalcoholic fatty liver disease: Update 2010. J Dig Dis 2011;12:38–44.
[37] Marchesini G, Bugianesi E, Forlani G, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. Hepatology 2003;37:917–23.
[38] Shimada M, Hashimoto E, Taniai M, et al. Hepatocellular carcinoma in patients with non-alcoholic steatohepatitis. J Hepatol 2002;37:154–60.
[39] Hamaguchi M, Kojima T, Takeda N, et al. The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. Ann Intern Med 2005; 143:722–8.
[40] Akbar DH, Kawther AH. Nonalcoholic fatty liver disease in Saudi type 2 diabetic subjects attending a medical outpatient clinic: prevalence and general characteristics. Diabetes Care 2003;26:3351–2.
[41] Donati G, Stagni B, Piscaglia F, et al. Increased prevalence of fatty liver in arterial hypertensive patients with normal liver enzymes: role of insulin resistance. Gut 2004;53:1020–3.
[42] Miyake T, Kumagi T, Hirooka M, et al. Metabolic markers and ALT cutoff level for diagnosing nonalcoholic fatty liver disease: a community-based cross-sectional study. J Gastroenterol 2012;47:696–703.
[43] Prati D, Taioli E, Zanella A, et al. Updated definitions of healthy ranges for serum alanine aminotransferase levels. Ann Intern Med 2002; 137:1–0.
[44] Lee JK, Shim JH, Lee HC, et al. Estimation of the healthy upper limits for serum alanine aminotransferase in Asian populations with normal liver histology. Hepatology 2010;51:1577–83.
[45] Kim HC, Nam CM, Jee SH, et al. Normal serum aminotransferase concentration and risk of mortality from liver diseases: prospective cohort study. BMJ 2004;328:983.
[46] Nugent C, Younossi ZM. Evaluation and management of obesity-related nonalcoholic fatty liver disease. Nat Clin Pract Gastroenterol Hepatol 2007;4:432–41.
[47] Fujikawa K, Ohata K, Honda T, et al. Nonalcoholic steatohepatitis with improved hepatic fibrosis after weight reduction. Intern Med 2004; 43:289–94.

[48] Yu C, Xu C, Xu L, et al. Serum proteomic analysis revealed diagnostic value of hemoglobin for nonalcoholic fatty liver disease. J Hepatol 2012;56:241–7.

[49] Chang Y, Ryu S, Sung E, et al. Higher concentrations of alanine aminotransferase within the reference interval predict nonalcoholic fatty liver disease. Clin Chem 2007;53:686–92.

[50] Chen Z, Qin H, Qiu S, et al. Correlation of triglyceride to high-density lipoprotein cholesterol ratio with nonalcoholic fatty liver disease among the non-obese Chinese population with normal blood lipid levels: a retrospective cohort research. Lipids Health Dis 2019;18:162.

[51] Chen Y, Cao D, Li C, et al. A nomogram for discrimination of non-alcoholic fatty liver disease in patients with chronic hepatitis B. Eur J Gastroenterol Hepatol 2020;Epub ahead of print.

[52] Eng K, Lopez R, Liccardo D, et al. A non-invasive prediction model for non-alcoholic steatohepatitis in pediatrics patients with non-alcoholic fatty liver disease. Dig Liver Dis 2014;46:1008–13.

[53] Zhao X, Zhang X, Ran X, et al. Simple-to-use nomogram for evaluating the incident risk of moderate-to-severe LEAD in adults with type 2 diabetes: a cross-sectional study in a Chinese population. Sci Rep 2020;10:3182.