Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

A Non-Invasive Prediction Model for Non-Alcoholic Fatty Liver Disease in Adults with Type 2 Diabetes Based on the Population of Northern Urumqi, China

This article was published in the following Dove Press journal:
Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy

Mingyue Xue[1,2]
Xiaoping Yang[3]
Yuan Zou[3]
Tao Liu[3]
Yinxia Su[3]
Cheng Li[4]
Hua Yao[3]
Shuxia Wang[3]

[1]Hospital of Traditional Chinese Medicine Affiliated to the Fourth Clinical Medical College of Xinjiang Medical University, Urumqi 830011, People's Republic of China; [2]College of Public Health, Xinjiang Medical University, Urumqi, Xinjiang 830011, People's Republic of China; [3]Health Management Institute, Xinjiang Medical University, Urumqi 830011, People's Republic of China; [4]The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830011, People's Republic of China

Background: High prevalence of non-alcoholic fatty liver disease (NAFLD) occurs in type 2 diabetes mellitus (T2DM), and about 13% of diabetic patients eventually die of liver cirrhosis or liver cancer. The purpose of our research was to develop a non-invasive predictive model of NAFLD in adults with T2DM.

Patients and Methods: Adult patients diagnosed with T2DM during physical examination in 2018 in Urumqi were recruited, in total 40,921 cases. We chose questionnaire and physical measurement variables to build a simple, low-cost model. Variables were selected by the least absolute shrinkage and selection operator regression (LASSO). The features chosen by LASSO were used to build the nomogram prediction model of NAFLD. The receiver operating curve (ROC) and calibration were used for model validation.

Results: Determinants in the nomogram included age, ethnicity, sex, exercise, smoking, dietary ratio, heart rate, systolic blood pressure (SBP), BMI, waist circumference, and atherosclerotic vascular disease (ASCVD). The area under ROC of developing group and validation group was 0.756 (95% confidence interval 0.750–0.761) and 0.755 (95% confidence interval 0.746–0.763), respectively, and the *P* values of the two calibration curves were 0.694 and 0.950, suggesting that the nomogram had good disease recognition ability and calibration.

Conclusion: A nomogram constructed with accuracy can calculate the possibility of NAFLD in adults with T2DM. If validated externally, this tool could be utilized as a non-invasive method to diagnose non-alcoholic fatty liver in adults with T2DM.

Keywords: type 2 diabetes mellitus, non-alcoholic fatty liver disease, screening tool, nomogram

Introduction

Nonalcoholic fatty liver disease (NAFLD) is the main cause of chronic liver disease, the incidence rate of which has been increasing rapidly in the world. NAFLD is usually closely related to obesity and type 2 diabetes mellitus (T2DM).[1,2] According to different reports, the prevalence of fatty liver in patients with T2DM is between 29.6%[3] and 87.1%.[4] In some patients, simple steatosis can lead to nonalcoholic steatohepatitis (NASH), and over time, NASH can lead to cirrhosis and associated sequelae, which includes hepatocellular carcinoma (HCC).[5] NAFLD is a systemic disease associated with

Correspondence: Hua Yao; Shuxia Wang
Email yaohua01@sina.com; 2724443591@qq.com

increased risk of various diseases, such as acute ischemic stroke,[6] which has been found in 42.5% acute ischemic stroke patients. Studies have found there were relations between helicobacter pylori infection and NAFLD.[7] Lacking effective treatment for NAFLD, more irreversible effects on the patient's metabolic system occur.[8] In China, elderly diabetic patients cannot go for timely, regular liver pathological examinations due to inconvenience and difficulty in getting medical treatment, and other factors.[9] Therefore, early identification of NAFLD patients with T2DM and timely change of their diet and lifestyle are crucial to the prevention of liver diseases and controlling the complications of diabetes.

The prevalence of metabolic syndrome in Xinjiang is at a high level. According to data from the national health examination in 2008, the detection rates of T2DM, NAFLD and obesity in Urumqi, the capital of Xinjiang, were 13.96%, 22.28% and 8.95%, respectively, while the prevalence rates of these diseases in the whole country were 11%, 29.2% and 7.9%,[10,11] respectively. Therefore, enough attention should be paid to metabolic syndrome in Xinjiang.

In clinics, most of the diagnoses of NAFLD are made via radiology. The most common diagnostic method is abdominal ultrasound, as liver ultrasonography has high accuracy.[12] In addition, semi-quantitative ultrasonography indices are correlated with metabolic derangements and liver histology changes. The use of indices is encouraged, such as the ultrasonographic fatty liver indicator, but the accuracy of ultrasound diagnosis depends on the level of the fabricator, and for people with abdominal obesity, the accuracy of ultrasound diagnosis of fatty liver is more limited.[13] Liver biopsy is the gold standard for diagnosis and staging of NAFLD.[14] However, liver biopsy is invasive, which may result in severe complications and is limited by sampling error.[15] Considering these inherent defects of liver ultrasonography and liver biopsy, an ideal detection method should be non-invasive, repeatable, simple, easy to obtain and low cost. Our research met these requirements, using questionnaire and physical measurement indicators to predict the risk of non-alcoholic fatty liver in patients with T2DM, and is especially useful in developing countries with high rates of disease and poor health care.

A non-invasive model is already being used to predict occurrence of liver cell inflammation and cardiovascular complications in NAFLD patients.[16] Nomogram is a graphical representation of the individual's risk classification. Although it has been widely used in other diseases,[17,18] it is rarely used in the prediction of fatty liver. To the best of our knowledge, only one study had established a risk prediction model for NAFLD in T2DM,[9] although it was limited by the small sample size and demand for blood indicators. Our aim was to establish a non-invasive model of nomogram based on large sample size, simple questionnaire and physical measurement data (without blood indicators), so as to accurately predict the risk of NAFLD in patients with T2DM.

Patients and Methods

Patient Selection

The physical examination data in the present study came from 643,439 samples of Urumqi residents in 2018. The government provides free physical examination to all residents in Xinjiang. The included samples cover all administrative areas in Urumqi, making it the largest and most formal examination in history for all the people. The examinations were carried out by medical teams in each administrative region, and all medical personnel received rigorous training, which ensures the accuracy of the examination results. After the investigation, all the data were summarized at the Health Management Hospital of Xinjiang Medical University, and we got the data from the Health Management Hospital.

As mentioned in the previous paragraph, the physical examination data used to support the findings of this study are stored and managed by the health management institute of Xinjiang medical university, and the data can only be used after applying for approval to the health management institute. Since I am an employee of the health management institute, my students and I participated in the collection of this medical examination data, and I have obtained approval to use these data for all my research.

Among 643,439 physical examination subjects, 52,365 adult patients with type 2 diabetes were screened out. 11,443 self-reported frequent drinkers were excluded from these patients. Totally, 40,922 T2DM cases were included in the final study, and all the patients in the included study signed written informed consent for this study. This study was conducted in

accordance with the principles outlined in the Declaration of Helsinki and approved by the Ethics Committee and institutional Review Committee of the Xinjiang Uygur Autonomous Region CDC.

Diagnosis of Diabetes

The diagnostic criteria for diabetes recommended by the American Diabetes Association (ADA) in 1997 were used in this study:[19] fasting blood glucose ≥ 7.0mmol/l, or people who self-reported having T2DM and took hypoglycemic drugs; the final incidence of diabetes was 13.96%, which was equivalent to the incidence of previous studies,[20] excluding gestational diabetes and type 1 diabetes.

Diagnosis of NAFLD

Diagnostic criteria and exclusion criteria for NAFLD:[21] i) diagnostic criteria: liver blips with fine density, enhanced near-field echo, and mild or obvious attenuation of far-field echo; liver and kidney echo contrast increased; the structure of intrahepatic lumen is blurred, which may be accompanied by mild to moderate liver enlargement. ii) Exclusion criteria: heavy drinking, that is, weekly drinking was converted into alcohol intake, >140g for male and b> 70g for female; viral hepatitis, mainly hepatitis B and C; autoimmune hepatitis; drug-induced hepatitis; some genetic diseases, including Wilson's disease, hemochromatosis, etc; anemia; long-term use of corticosteroids; patients with fatty liver caused by endocrine diseases. After summarizing all the results of physical examination, two doctors from hepatology checked the diagnosis results of fatty liver, which were consistent with the preliminary diagnosis. Finally, the prevalence of non-alcoholic fatty liver in patients with T2DM in Urumqi was 37.39%.

Variable Definitions

Co-variables included in this study: 1) demographic information: age, gender, ethnicity and career; 2) living habits, exercise, smoking and dietary habits; 3) indicators of body measurement: body mass index (BMI), blood pressure, heart ratio, waist circumference (WC), abdominal examination, fundus examination and B-ultrasound; 4) other relative diseases.

In the study, the age was divided into: "18–29", "30–39", "40–49", "50–59", "60–69", "70–79" and "over 80" group; ethnicity consisted of: "Han", "Uygur", "Kazak", "Hui", "Mongolian" and "other nationalities" group. Smoking condition was divided into non, ever or current group; exercise: someone who exercised more than three times a week for the previous six months was defined as "yes"[22] others were defined as "no"; dietary habits: "meat based", "meat balanced", and "vegetarian based".

The baseline comorbidities considered in this study were atherosclerotic vascular disease (ASCVD), heart failure, bronchitis, renal failure, chronic nephritis, precordial pain, eye diseases, psychosis, diabetic nephropathy and anemia, defined as yes and no. In this study, ASCVD included cerebrovascular, coronary or peripheral disease; participants were asked to answer the question "have you ever had a heart attack, or do you have a stent or a bridge?", if the answer was affirmative, it was considered as ASCVD. Diabetic nephropathy is one of the most important complications in diabetic patients and has become the second cause of end-stage renal disease at present, the diagnosis of which is traditionally based on microalbuminuria (MA);[23] and renal failure is a pathological condition in which partial or complete loss of renal function is caused by various chronic kidney diseases in the later stages;[24] chronic nephritis is also called chronic glomerulonephritis, this form of nephritis develops slowly and causes few symptoms in its early stages. But as the condition gets worse, this condition can cause severe kidney damage and kidney failure, which may develop after a sudden disease;[25] heart failure, also known as congestive heart failure and congestive cardiac failure, is a decline in cardiac function caused by various heart diseases, including shortness of breath, fatigue, palpitation, edema of lower limbs, etc;[26] precordial pain is the primary cardiac-related complaint, and it is a symptom that can be caused by many diseases;[27] eye diseases include retinal hemorrhage, papilledema and cataract; psychosis is defined as a condition that affects the way our brain processes information, which causes the patient to lose touch with reality; psychosis is a symptom, not an illness, mental or physical illness, substance abuse, or extreme stress or trauma can cause it;[28] anemia refers to a common clinical symptom when the volume of red blood cells in human peripheral blood is decreased below the lower limit of the normal range. In our research, the diagnosis of all of these diseases was based on the main complaint of the patient.

Statistical Analysis

Xinjiang physical examination data contained a large scale of samples, data cleaning is of great importance. We first deleted more than 200 variables that were not relevant to the study (such as the time of physical examination, the information of inputator, etc.) or had a large number of missing values, and finally 23 variables were included in our study. The mode was used to fill the null values in categorical variables, and the mean value was used to fill the null values of continuous variables.

The samples were randomly assigned to the development group (n=28,645) and the validation group (n=12,276) on a ratio of 7:3. When comparing the baseline characteristics of the modeling group and the verification group, for the continuous variables, the normality test was performed, when meeting normal distribution, the *t*-test would be used or the rank sum test was used. Chi-squared or Fisher's exact test was used for categorical variables. Differences with a two-sided P-value of <0.05 were deemed statistically significant. The appendix lists the results of logistic regression analysis of risk factors for NAFLD in T2DM by gender (Table S1; Figure S1 and S2).

The open source Python software Version 3.7.2 (https://www.python.org) was used for data cleaning; the open source R software Version 3.6.1 (http://www.r-project.org) was used for statistical analysis and modeling: the data of development and validation group were randomly divided by CARET-package; variable selection of LASSO regression was performed by glmnet-packages; establishing and calibrating nomogram model were done with RMS-package; the score of nomogram was calculated by nomogramEx-package; the ROCR-package and PROC-package was used to plot ROC curve.

Results

The Missing Data of Covariates

Figure 1 showed the missing data of all covariates, the bar chart described the missing proportion of 23 variables (career 0.02%, exercise 0.81%, smoking 0.88%, dietary ratio 7.91%, heart rate 0.79%, DBP 4.42%, SBP 4.41%, BMI 0.30%, waist 2.49%, and no loss for other variables). The graph was sorted by samples missing components, and the red represents missing data. As shown in the figure, the proportion of samples without missing values is the largest among all samples, and the proportion of samples only missing dietary ratio was second.

Baseline Situation (Table 1)

Among the 40,921 diabetic patients in this study, 28,645 and 12,276 were randomly assigned to the development group and validation group, respectively. The comparison of baseline data showed that there was no significant difference between the development group and the validation group ($P > 0.05$).

Characteristics' Selection

As shown in Figure 2A, Figure 3 and B, LASSO screened out 11 variables from 23 covariates, including age, ethnicity, sex, exercise, smoking, dietary ratio, heart rate, systolic blood pressure (SBP), BMI, waist circumference, and ASCVD (Table 2).

Establishment of the Nomogram

Logistic multiple regression analysis showed that the 11 variables screened by LASSO were significant. In addition, among the 11 covariates, the maximum variance expansion factor (VIF) was 1.979 and the minimum was 1.004, indicating that there was no multicollinearity among them.

Factors with statistically significant differences were determined by the multiple logistic regression model. Nomogram prediction model was constructed to predict the possibility of NAFLD in T2DM patients(Figure 3). If each factor corresponded to the upper scale, the score of that factor could be obtained. By adding the scores of each factor, the total score could be calculated. According to total scores, the probability of each individual developing NAFLD could be calculated. The higher the score, the more likely the individual was to develop NAFLD.

Take an example of nomogram usage: a sample was randomly selected from the developing set, a 45 year old man, Han nationality, no exercise, balanced diet, current smoking, heart rate was 90, systolic blood pressure (SBP) was 140 (mmHg), BMI was 28 (kg/m^2), waist circumference was 85 (cm), no ASCVD history, so total score was 168.405, and the corresponding possibility of NAFLD was close to 80%.

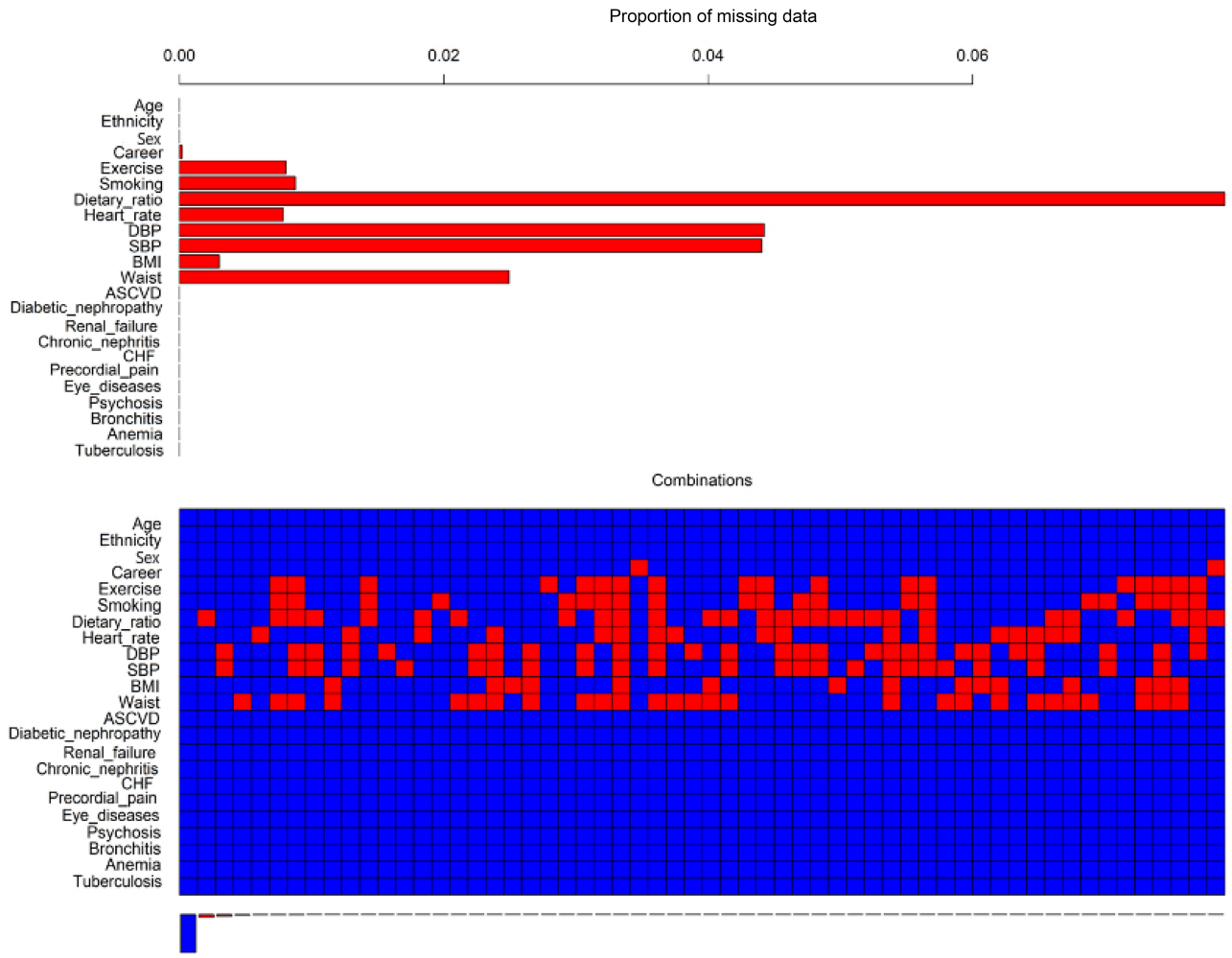

Figure 1 Missing data of covariates.
Abbreviations: DBP, diastolic blood pressure; SBP, systolic blood pressure; BMI, body mass index; ASCVD, atherosclerotic vascular disease; CHF, congestive heart failure.

Validation of the Nomogram

The accuracy of the model was evaluated in terms of AUC values and Hosmer-Lemeshow goodness of fit test. The sensitivity and specificity of cut-off in development group were 69.7% and 67.7% and in validation group were 67.0% and 69.8%, respectively. And the AUC value of development group and validation group were 0.756 (95% confidence interval 0.750–0.761) and 0.755 (95% confidence interval 0.746–0.763) respectively (Figures 4A and B and 5A and B).The calibration chart also showed good agreement between nomogram's predicted value and the actual probability of disease. When $p>0.05$, the calibration ability of the model is good. The calibration curve of the development group and validation group showed $p=0.654$ and $p=0.950$, respectively. All of which were greater than 0.05, indicating that the model had good calibration ability.

Discussion

In this study, we aimed to establish an NAFLD screening model only based on questionnaire indicators and physical measurement. The model had the following advantages: low cost, repeatable, graphic display and easy to use, which was particularly suitable for areas with high epidemiological risk and low socioeconomic status, such as Xinjiang, China. Risk factors in this nomogram included age, ethnicity, sex, exercise, smoking, dietary ratio, heart rate, systolic blood pressure (SBP), BMI, waist circumference, and ASCVD.

Previously recognized results showed that poor living habits such as lack of exercise and poor diet were the main

Table 1 Characteristics of the Development Set and Validation Set

Characteristics	Development Set (n = 28,645)	Validation Set (n = 122,76)	*P*-value
Age (years), n(%)			0.380
18–29	200(0.70)	72(0.59)	
30–39	563(1.97)	241(1.96)	
40–49	2061(7.19)	920(7.49)	
50–59	4600(16.06)	1990(16.21)	
60–69	8683(30.31)	3604(29.36)	
70–79	9503(33.18)	4149(33.80)	
≥ 80	3035(10.60)	1300(10.59)	
Ethnicity, n(%)			0.692
Han	20,809 (72.64)	8908 (72.56)	
Uygur	3263 (11.39)	1393 (11.35)	
Kazak	401 (1.40)	173 (1.41)	
Hui	3825 (13.35)	1639 (13.35)	
Mongolian	36 (0.13)	24 (0.20)	
Other nationalities	311 (1.09)	139 (1.13)	
Sex, n(%)			0.623
Male	12,085(42.19)	5212(42.46)	
Female	16,560(57.81)	7064(57.54)	
Career, n(%)			0.602
Trader or service people	12,684(44.28)	5408(44.05)	
Agricultural workers	13,148(45.90)	5638(45.93)	
Factory workers	1582(5.52)	723(5.89)	
Soldiers	359(1.25)	146(1.19)	
Others	872(3.04)	361(2.94)	
Exercise, n (%)			0.703
No	12,740(44.48)	5434(44.27)	
Yes	15,905(55.52)	6842(55.73)	
Smoking, n (%)			0.193
Non	25,467(88.91)	10,871(88.55)	-
Ever	639(2.23)	256(2.09)	
Current	2539(8.86)	1149(9.36)	
Dietary ratio, n (%)			0.246
Vegetarian based	1160(4.05)	495(4.03)	
Meat balanced	24,630(85.98)	10,623(86.53)	
Meat based	2855(9.97)	1158(9.43)	
ASCVD, n(%)			0.735
No	23,623(82.47)	10,106(82.32)	
Yes	5022(17.53)	2170(17.68)	
Diabetic nephropathy, n (%)			0.230
No	28,325(98.88)	12,156(99.02)	
Yes	320(1.12)	120(0.98)	
Renal failure, n(%)			0.534
No	28,465(99.37)	12,206(99.43)	
Yes	180(0.63)	70(0.57)	
Chronic nephritis, n(%)			0.665
No	28,545(99.65)	12,229(99.62)	
Yes	100(0.35)	47(0.38)	
Congestive heart failure, n(%)			0.693
No	28,632(99.55)	12,270(99.51)	
Yes	130(0.45)	60(0.49)	
Precordial pain, n(%)			0.812
No	28,202(98.05)	12,095(98.09)	
Yes	560(1.95)	235(1.91)	
Eye diseases, n(%)			0.461
No	28,303(98.81)	12,118(98.71)	
Yes	342(1.19)	158(1.29)	
Psychosis, n(%)			0.935
No	28,518(99.56)	12,223(99.57)	
Yes	127(0.44)	53(0.43)	
Bronchitis, n(%)			0.433
No	27,383(95.59)	11,713(95.41)	
Yes	1262(4.41)	563(4.59)	
Anemia, n(%)			0.322
No	24,990(87.24)	10,665(86.88)	
Yes	3655(12.76)	1611(13.12)	
Tuberculosis, n(%)			0.749
No	28,537(99.62)	12,233(99.65)	
Yes	108(0.38)	43(0.35)	
BMI (kg/m^2)	26.03±3.58	26.02±3.55	0.801
SBP (mmHg)	130(120–141)	130 (120–141)	0.314
DBP (mmHg)	78(70–82)	78(70–82)	0.416
Heart rate (counts per minute)	76(70–81)	76(70–81)	0.614
Waist circumference (cm)	90.41±10.45	90.35±10.40	0.617

Abbreviations: BMI, body mass index; ASCVD, atherosclerotic vascular disease; DBP, diastolic blood pressure; SBP, systolic blood pressure.

risk factors for T2DM, fatty liver, obesity and other chronic diseases,[29–31] which was the same as the results of our study. In addition, this paper proved that smoking was a risk factor of NAFLD and at present, there are few studies about the relation between smoking and NAFLD. Smoking was confirmed to be associated with advanced fibronecrosis in a large, multi-center cohort clinical research network of non-alcoholic steatohepatitis.[32] Considering the direct influence of smoking on the liver, patients should be encouraged to quit smoking to reduce

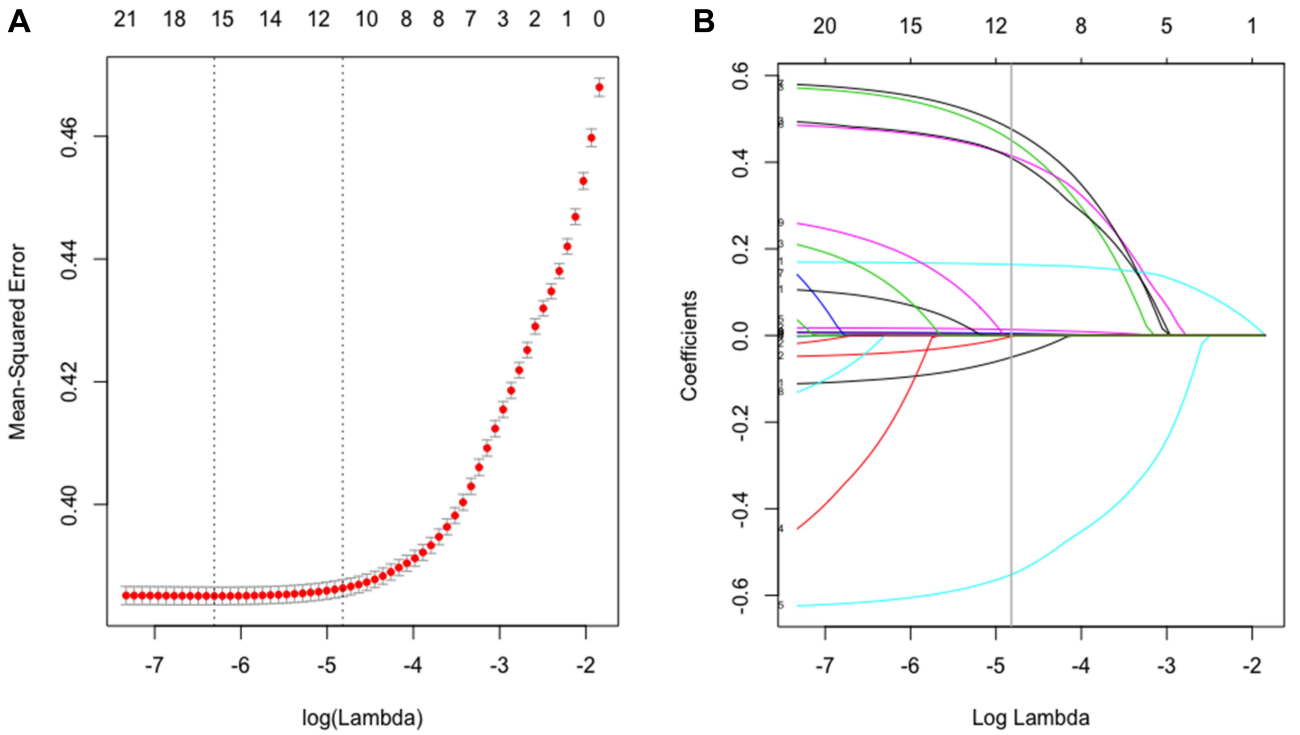

Figure 2 Texture feature selection using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model.
Notes: (**A**) Optimization parameters (lambda) of lasso model were selected by 10 times cross-validation. The Mean-Squared Error was plotted versus log (Lambda). Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). Minimum criteria refers to the one among all λ values to get the mean value of the minimum target parameter. 1-SE criteria refers to the λ value of the simplest model in a variance range of minimum criteria. (**B**) LASSO coefficient profiles of the 23 features. A coefficient profile plot was produced against the log (Lambda) sequence. Vertical line was drawn at the value selected using 10 times cross-validation, where optimal lambda resulted in 11 features with nonzero coefficients.

the risk of cardiovascular disease and hypertension in this high-risk population. As in previous studies, the prevalence of NAFLD was highest in those who were middle aged, and men compared to women.[33]

Xinjiang is the largest minority region in China. Different ethnic groups have different living habits, which may be the cause of the difference in the prevalence of non-alcoholic fatty liver. This study found that the prevalence among Kazakhs and Uyghurs was low, and that it was highest among Hans, which is consistent with previous studies.[34]

There is a strong association between liver diseases and T2DM, which is higher than expected by a correlation between two very common diseases.[35] Overexpression of hepassocin in diabetic patients with NAFLD may facilitate increased hepatic lipid accumulation.[36] NAFLD was also closely related to other diseases such as arteriosclerosis, cerebral and cardiovascular strokes, and colon tumor.[37,38] NAFLD is considered as the "hepatic manifestation of the metabolic syndrome", and increases the risk of ASCVD through effects on synthesis of lipoproteins, coagulation proteins, and inflammatory factors.

In view of the gender differences in NAFLD, we established the NAFLD screening models for both men and women in the appendix. In the female model, smoking was not selected by LASSO, but ASCVD and eye diseases had greater influence on females than males. Zhang et al reported that the presence of diabetes conferred the potential for a greater risk of vascular complications and inflammatory disease in females compared with males due to some factors, including the contribution of sex hormones and sex-specific risk factors.[9] Models based on different sexes have a good degree of discrimination and prediction ability in the development set (AUC = 0.750 for females, AUC = 0.764 for males) and

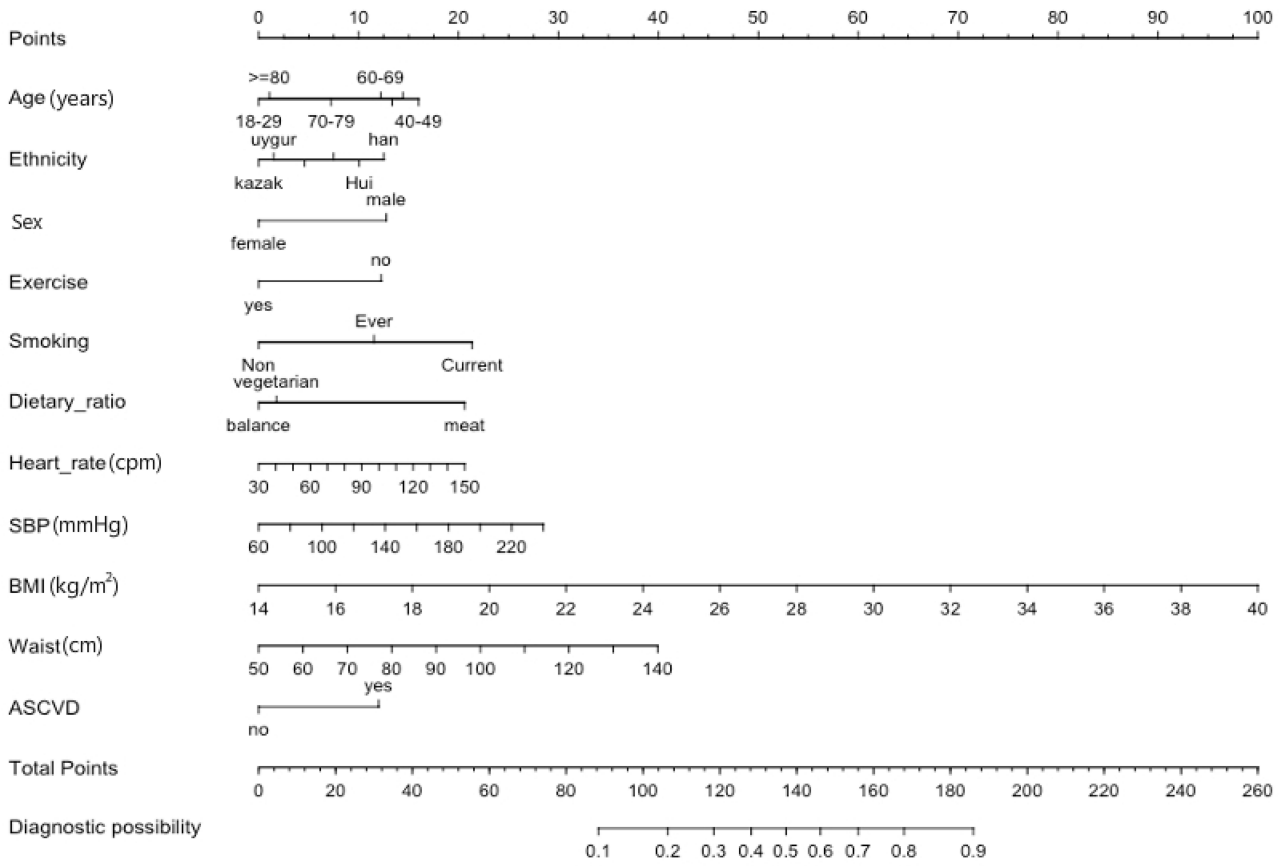

Figure 3 Nomogram to predict the risk of T2DM developing into NAFLD.
Notes: To use the nomogram, an individual participant's value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value; the sum of these numbers is located on the Total Points axis to determine the risk of NAFLD.
Abbreviations: BMI, body mass index; SBP, systolic blood pressure; ASCVD, atherosclerotic vascular disease; cpm, counts per minute.

validation set (AUC = 0.754 for females, AUC = 0.764 for males).

There are several advantages to this study. First, as far as we know, this was the first study in China to develop and evaluate a model for predicting NAFLD prevalence in patients with T2DM. Second, the data for this study came from a large, representative national health examination. Third, the variables we used were easy to get from the physical examination data, which is especially useful in less developed regions. Fourth, the nomogram that we developed uses non-invasive variables and it includes four adjustable variables: waistline, BMI, eating habits and smoking. Thus, it meant that we could control the occurrence and development of NAFLD by suggesting balanced diet, strengthening exercise, and quitting smoking.

Because our study came from a large census rather than a special survey on diabetes or fatty liver disease, we did not have the family history of diabetes. In fact, a flaw in China's health system in the last century has led to a large number of undiagnosed disease histories. Therefore, family history of diabetes is not an accurate variable and cannot be adjusted to improve the disease.

This study also had some limitations. First of all, our study was based on a cross-sectional survey and cannot predict long-term trends. Second, there was a lack of some important potential risk factors, NAFLD is strongly related to bacteria in the gut, and some tumors are also related to NAFLD.[39] Our physical examination data did not contain these detection indicators. Third, in addition to the measurement of blood pressure, all other measurement indicators were

Table 2 Multivariate Logistic Regression Analysis for Risk Factors Associated with Hypertension in the Development Group (N =28,645)

Variables	β	Odds Ratio	95% CI	Z value	P value
Age(years), n(%)					
18–29	Ref	1	Ref	—	—
30–39	0.601	1.824	1.154–2.943	2.522	0.012
40–49	0.720	2.056	1.346–3.220	3.245	0.001
50–59	0.651	1.917	1.263–2.986	2.971	0.003
60–69	0.551	1.734	1.144–2.697	2.522	0.012
70–79	0.327	1.386	0.913–2.158	1.492	0.136
≥ 80	0.048	1.049	0.686–1.643	0.215	0.830
Ethnicity, n(%)					
Han	Ref	1	Ref	—	—
Uygur	−0.497	0.608	0.554–0.668	−10.368	<0.001
Kazak	−0.564	0.569	0.448–0.720	−4.654	<0.001
Hui	−0.111	0.895	0.824–0.972	−2.639	0.008
Mongolian	−0.226	0.797	0.366–1.657	−0.592	0.554
Other nationalities	−0.357	0.700	0.526–0.924	−2.486	0.013
Sex, n(%)					
Female	Ref	1	Ref	—	—
Male	0.576	1.779	1.676–1.888	19.005	<0.001
Exercise					
No	Ref	1	Ref	—	—
Yes	−0.552	0.576	0.544–0.610	−18.983	<0.001
Smoking, n(%)					
Non	Ref	1	Ref	—	—
Ever	0.520	1.683	1.408–2.009	5.740	<0.001
Current	0.963	2.619	2.375–2.890	19.222	<0.001
Dietary ratio, n(%)					
Vegetarian based	Ref	1	Ref	—	—
Meat balanced	−0.081	0.923	0.808–1.054	−1.191	0.234
Meat based	0.848	2.334	2.003–2.722	10.839	<0.001
Heart rate (counts per minute)	0.008	1.008	1.006–1.009	5.699	<0.001
SBP (mmHg)	0.007	1.007	1.006–1.009	9.088	<0.001
BMI (kg/m^2)	0.173	1.189	1.177–1.202	31.485	<0.001
Waist circumference (cm)	0.020	1.020	1.016–1.024	10.798	<0.001
ASCVD, n(%)					
No	Ref	1	Ref	—	—
Yes	0.540	1.716	1.599–1.843	14.891	<0.001

Abbreviations: β, the regression coefficient; CI, confidence interval; BMI, body mass index; SBP, systolic blood pressure; ASCVD, atherosclerotic vascular disease.

only a single measurement, which might cause deviation of the results.

Conclusion

With the great progress of scientists in understanding the pathogenesis and treatment of NAFLD, the prediction and early recognition of NAFLD have become more meaningful. We have developed a nomogram model with high accuracy and recognition ability. It could calculate the risk of NAFLD quickly and simply. Our model makes up for shortcomings of missed detection of traditional detection methods. The

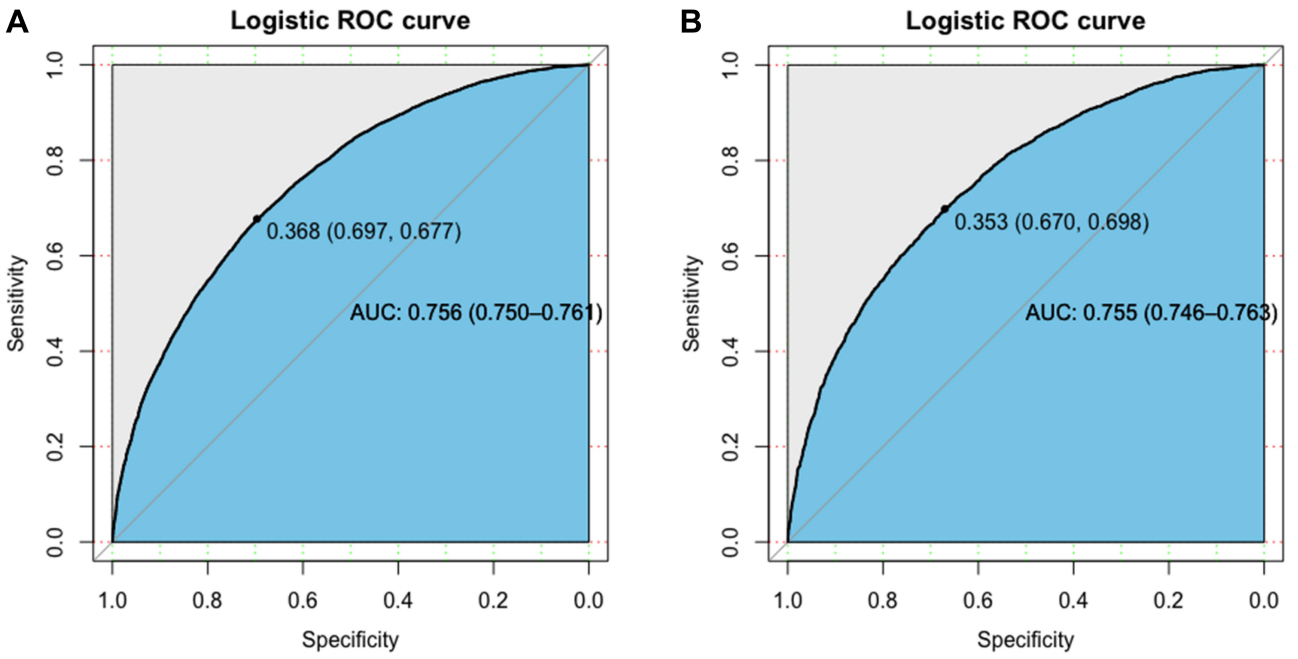

Figure 4 ROCs for validating the discrimination power of the nomogram. (**A**) Development group. (**B**) Validation group (AUC = 0.756 vs 0.755).

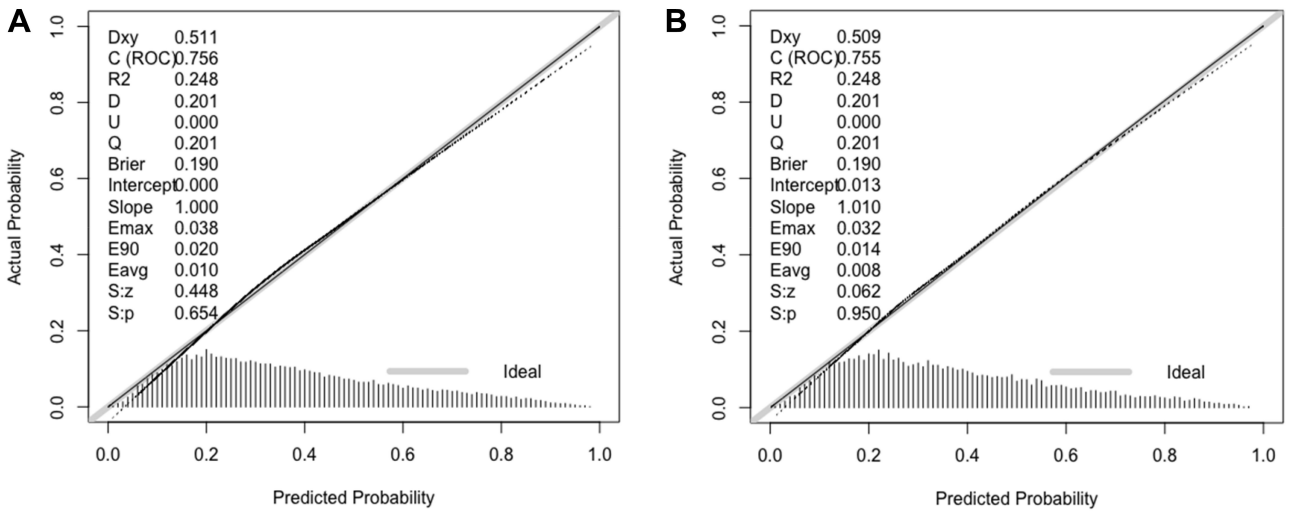

Figure 5 Calibration curves for the prediction and training group models (P = 0.654 vs 0.950).
Notes: The diagonal solid line represents a perfect prediction by an ideal model; the dotted line represents the performance of the nomogram, of which a closer fit to the diagonal dotted line represents a better prediction.

prediction model has been shown in a nomogram, which consists of five adjustable variables for guidance to a healthy life. However, the incidence of children's NAFLD is increasing year by year, and it attracts more and more attention of society. In the future, we hope to have accurate prediction of children's disease model, and provide support for children's NAFLD prevention guidelines and interventions.

Data Sharing Statement

The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgments

The study was supported by Xinjiang science and technology department who set up the project (NO.2018E02057) for Hua Yao. We would like to thank all the health

professionals who participated in this cross-sectional survey for their efforts.

Author Contributions

All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

The authors declare no conflicts of interest.

References

1. European Association for the Study of the L, European Association for the Study of D, European Association for the Study of O. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol*. 2016;64(6):1388–1402. doi:10.1016/j.jhep.2015.11.004
2. Doycheva I, Watt KD, Alkhouri N. Nonalcoholic fatty liver disease in adolescents and young adults: the next frontier in the epidemic. *Hepatology*. 2017;65(6):2100–2109. doi:10.1002/hep.29068
3. Fan N, Zhang L, Xia Z, Peng L, Wang Y, Peng Y. Sex-specific association between serum uric acid and nonalcoholic fatty liver disease in type 2 diabetic patients. *J Diabetes Res*. 2016;2016:3805372. doi:10.1155/2016/3805372
4. Sima A, Timar R, Vlad A, et al. Nonalcoholic fatty liver disease: a frequent condition in type 2 diabetic patients. *Wien Klin Wochenschr*. 2014;126(11–12):335–340. doi:10.1007/s00508-014-0530-8
5. Cai S, Li Z, Yu T, Xia M, Peng J. Serum hepatitis B core antibody levels predict HBeAg seroconversion in chronic hepatitis B patients with high viral load treated with nucleos(t)ide analogs. *Infect Drug Resist*. 2018;11:469–477. doi:10.2147/IDR.S163038
6. Abdeldyem SM, Goda T, Khodeir SA, Abou Saif S, Abd-Elsalam S. Nonalcoholic fatty liver disease in patients with acute ischemic stroke is associated with more severe stroke and worse outcome. *J Clin Lipidol*. 2017;11(4):915–919. doi:10.1016/j.jacl.2017.04.115
7. Abo-Amer EE, Sabal A, Ahmed R, Hasan NFE, Abd-Elsalam S. Relationship between helicobacter pylori infection and nonalcoholic fatty liver disease (NAFLD) in a developing country: a cross-sectional study. *Diabetes Metab Syndr Obes*. 2020;13:619–625. doi:10.2147/DMSO.S237866
8. Schmidt MI, Duncan BB, Bang H, et al. Identifying individuals at high risk for diabetes: the atherosclerosis risk in communities study. *Diabetes Care*. 2005;28(8):2013–2018. doi:10.2337/diacare.28.8.2013
9. Zhang Y, Shi R, Yu L, Ji L, Li M, Hu F. Establishment of a risk prediction model for non-alcoholic fatty liver disease in type 2 diabetes. *Diabetes Ther*. 2020;11(9):2057–2073. doi:10.1007/s13300-020-00893-z
10. National Bureau of Statistics of China. Population census of People's Republic of China. 2018. Available from: http://www.stats.gov.cn/tjsj/ndsj/2018/indexeh.htm. Accessed January 12, 2021.
11. Zhou F, Zhou J, Wang W, et al. Unexpected rapid increase in the burden of NAFLD in China from 2008 to 2018: a systematic review and meta-analysis. *Hepatology*. 2019;70(4):1119–1133. doi:10.1002/hep.30702
12. Kromrey ML, Ittermann T, Berning M, et al. Accuracy of ultrasonography in the assessment of liver fat compared with MRI. *Clin Radiol*. 2019;74(7):539–546. doi:10.1016/j.crad.2019.02.014
13. Hernaez R, Lazo M, Bonekamp S, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. *Hepatology*. 2011;54(3):1082–1090. doi:10.1002/hep.24452
14. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology*. 2012;55(6):2005–2023. doi:10.1002/hep.25762
15. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology*. 2005;128(7):1898–1906. doi:10.1053/j.gastro.2005.03.084
16. Hanafy AS, Seleem WM, El-Kalla F, AbdAlkhalik Basha M, Abd-Elsalam S. Efficacy of a non-invasive model in predicting the cardiovascular morbidity and histological severity in non-alcoholic fatty liver disease. *Diabetes Metab Syndr*. 2019;13(3):2272–2278. doi:10.1016/j.dsx.2019.05.032
17. Xue M, Su Y, Feng Z, Wang S, Yao H. A nomogram model for screening the risk of diabetes in a large-scale Chinese population: an observational study from 345,718 participants. *Sci Rep*. 2020;10(1):11600.
18. Xue M, Liu L, Wang S, et al. A simple nomogram score for screening patients with type 2 diabetes to detect those with hypertension: a cross-sectional study based on a large community survey in China. *PLoS One*. 2020;15(8):e0236957. doi:10.1371/journal.pone.0236957
19. Gavin JR, Alberti KG, Davidson MB, et al. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2000;23(Suppl 1):S4–19.
20. Gao Y, Xie X, Wang S-X, et al. Effects of sedentary occupations on type 2 diabetes and hypertension in different ethnic groups in North West China. *Diab Vasc Dis Res*. 2017;14(4):372–375. doi:10.1177/1479164117696050
21. Fan JG, Jia JD, Li YM, et al. Guidelines for the diagnosis and management of nonalcoholic fatty liver disease: update 2010: (published in Chinese on Chinese Journal of Hepatology 2010; 18: 163–166). *J Dig Dis*. 2011;12(1):38–44. doi:10.1111/j.1751-2980.2010.00476.x
22. Yang L, Yan K, Zeng D, et al. Association of polycyclic aromatic hydrocarbons metabolites and risk of diabetes in coke oven workers. *Environ Pollut*. 2017;223:305–310. doi:10.1016/j.envpol.2017.01.027
23. Ioannou K. Diabetic nephropathy: is it always there? Assumptions, weaknesses and pitfalls in the diagnosis. *Hormones*. 2017;16(4):351–361. doi:10.14310/horm.2002.1755
24. Herrera AS Use of nicotine, analogues, thereof, precursors thereof or derivatives thereof in the treatment of various pathological processes capable of improvement with a-msh administered in prophylactic or therapeutic form. 2009.
25. Oliva-Damaso N, Oliva-Damaso E, Payan J. Acute and chronic tubulointerstitial nephritis of rheumatic causes. *Rheum Dis Clin North Am*. 2018;44(4):619–633. doi:10.1016/j.rdc.2018.06.009
26. National Clinical Guideline C. National Institute for Health and Clinical Excellence: Guidance. Chronic heart failure: national clinical guideline for diagnosis and management in primary and secondary care. London: Royal College of Physicians (UK) Copyright © 2010, National Clinical Guideline Centre; 2010.
27. Ocampo-Vázquez CM, Hernández-Benítez R, Iglesias-Leboreiro J, Bernardez-Zapata I, Luna MM, Cervantes-Izaguirre AP. Pediatric precordial pain, Hospital Español de México's experience. *Arch Cardiol Mex*. 2019;89(1):31–37. doi:10.24875/acm.M19000004
28. Reinhardt MM, Cohen CI. Late-life psychosis: diagnosis and treatment. *Curr Psychiatry Rep*. 2015;17(2):1. doi:10.1007/s11920-014-0542-0
29. Brody H. Fatty liver disease. *Nature*. 2017;551(7681):S85.

30. Booth FW, Roberts CK, Laye MJ. Lack of exercise is a major cause of chronic diseases. *Compr Physiol*. 2012;2(2):1143–1211. doi:10.1002/cphy.c110025
31. Suárez M, Boqué N, Del Bas JM, Mayneris-Perxachs J, Arola L, Caimari A. Mediterranean diet and multi-ingredient-based interventions for the management of non-alcoholic fatty liver disease. *Nutrients*. 2017;9(10):1052. doi:10.3390/nu9101052
32. Zein CO, Unalp A, Colvin R, Liu Y-C, McCullough AJ. Nonalcoholic steatohepatitis clinical research N. smoking and severity of hepatic fibrosis in nonalcoholic fatty liver disease. *J Hepatol*. 2011;54(4):753–759. doi:10.1016/j.jhep.2010.07.040
33. Lonardo A, Bellentani S, Argo CK, et al. Epidemiological modifiers of non-alcoholic fatty liver disease: focus on high-risk groups. *Dig Liver Dis*. 2015;47(12):997–1006. doi:10.1016/j.dld.2015.08.004
34. Cai W, Fangping HE, Xuan WU, Yao H. The association of serum uric acid level with prevalence of non-alcoholic fatty liver in Uyghur and Han ethnicities. *Chin J Endocrinol Metabolism*. 2012;11:890–894.
35. Hamed AE, Elwan N, Naguib M, et al. Diabetes association with liver diseases: an overview for clinicians. *Endocr Metab Immune Disord Drug Targets*. 2019;19(3):274–280. doi:10.2174/1871530318666181116111945
36. Abdelmoemen G, Khodeir SA, Zaki AN, Kassab M, Abou-Saif S, Abd-Elsalam S. Overexpression of hepassocin in diabetic patients with nonalcoholic fatty liver disease may facilitate increased hepatic lipid accumulation. *Endocr Metab Immune Disord Drug Targets*. 2019;19(2):185–188. doi:10.2174/1871530318666180716100543
37. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. *JAMA*. 2015;313(22):2263–2273. doi:10.1001/jama.2015.5370
38. Oni ET, Agatston AS, Blaha MJ, et al. A systematic review: burden and severity of subclinical cardiovascular disease among those with nonalcoholic fatty liver; should we care? *Atherosclerosis*. 2013;230(2):258–267. doi:10.1016/j.atherosclerosis.2013.07.052
39. de Groot PF, Frissen MN, de Clercq NC, Nieuwdorp M. Fecal microbiota transplantation in metabolic syndrome: history, present and future. *Gut Microbes*. 2017;8(3):253–267. doi:10.1080/19490976.2017.1293224