ORIGINAL ARTICLE

Analysis of risk factors and their interactions in type 2 diabetes mellitus: A cross-sectional survey in Guilin, China

Disha Zou[1], Yao Ye[1], Nina Zou[2], Jian Yu[1]*

[1]Department of Endocrinology, The Affiliated Hospital of Guilin Medical University, Guilin, Guangxi, and [2]The Nursing Group of Loudi Health School, Loudi, Hunan, China

ABSTRACT

Aims/Introduction: Type 2 diabetes is a metabolic disease characterized by insulin resistance, and is associated with the effects of genetic and environmental factors. The present study aimed to not only analyze the influence of a single factor for type 2 diabetes, but also to investigate the interaction effects between risk factors.

Materials and Methods: A total of 6,660 individuals selected by the method of cluster random sampling accepted a cross-sectional survey (questionnaire investigation, physical measurement, laboratory examination and liver ultrasound examination). The classification tree was used to analyze the risk factors and their interactions in type 2 diabetes. The clinical and metabolic characteristics were compared between type 2 diabetes patients and controls, and the non-conditional logistic regression model was used to quantitatively analyze the interactions.

Results: A total of 338 participants were classified as type 2 diabetes (217 men and 121 women), the classification tree model showed three variables with close associations with type 2 diabetes: age, triglycerides (TG) and non-alcoholic fatty liver disease (NAFLD). Type 2 diabetes patients had higher age and incidences of high TG, NAFLD, hypertension, high body mass index, high uric acid, high total cholesterol, high low-density lipoprotein cholesterol and low high-density lipoprotein cholesterol. The multivariate logistic regression analysis showed that the following factors had interactions in type2 diabetes: high TG × advanced age (odds ratio 2.499, 95% confidence interval 1.868–3.344, *P* = 0.000), NAFLD × advanced age (odds ratio 1.250, 95% confidence interval 1.048–1.491, *P* = 0.013) and NAFLD × high TG (odds ratio 1.349, 95% confidence interval 1.144–1.590, *P* = 0.000).

Conclusions: The present study showed that type 2 diabetes resulted from the interactions of many factors; the interactions among age, TG and NAFLD are important risk factors for type 2 diabetes.

INTRODUCTION

Diabetes mellitus is characterized by elevated blood glucose, and is a metabolic disease caused by insulin insufficiency and/ or dysfunction associated with the long-term effects of genetic and environmental factors[1]. With rapid socioeconomic development, rising living standards and the accelerated aging of society, the prevalence of diabetes mellitus is trending upward in both developed and developing countries[2,3]. Consequently, diabetes mellitus has become the third most common chronic non-communicable disease after cardiovascular and cerebrovascular diseases, and cancers. According to the 2013 statistics of the International Diabetes Federation, the number of diabetic patients worldwide has reached 382 million, which is an increase of nearly 34% from 2010 (285 million)[4], and the number is expected to reach 592 million in 2035[5].

Diabetes mellitus is classified into type 1 diabetes mellitus, type 2 diabetes mellitus, gestational diabetes mellitus and other special types of diabetes mellitus. Type 2 diabetes accounts for 90% of all cases, and is primarily associated with insulin resistance and subsequent lipid metabolism disorders[5]. Type 2

diabetes not only results in high glucose, but also affects the metabolisms of other tissues and organs, leading to a series of complications with severe consequences for the quality of life (i.e., vascular disease[6], chronic kidney disease[7], neuropathy and retinopathy). The 2012 International Diabetes Federation data show that diabetic patients in China spend up to 107.4 billion yuan on diabetes mellitus treatment, and the World Health Organization estimates that by 2030, diabetes mellitus will become the seventh leading cause of death worldwide[8]. As a result, type 2 diabetes places a heavy burden on families and society.

The pathogenesis of type 2 diabetes is complex and unclear[1], certain clinical and metabolic factors are closely associated with the development and progression of type 2 diabetes, and changes in the environment or lifestyle might improve the risk of type 2 diabetes. We should be concerned not only about the influence of a single factor for type 2 diabetes, but also about whether there are interaction effects between risk factors, which are meaningful for understanding the pathogenesis and the basic care of type 2 diabetes. The classification and regression tree (CRT) analysis has been applied to study type 2 diabetes, including discriminating predictors of type 2 diabetes and risk stratification[9], identifying central variables predicting the albuminuria in type 2 diabetes patients[10], and so on. However, to our knowledge, this method has not been used to investigate the interactions of risk factors for type 2 diabetes, which might be a new strategy to visually display the interactions between risk factors in the form of a tree diagram. Accordingly, in the present study, we used a classification tree to investigate the risk factors and their interactions in type 2 diabetes, and a non-conditional logistic regression model to quantitatively analyze the interactions.

MATERIALS AND METHODS

Study population

The study protocol obtained approval from the ethics committee of the Affiliated Hospital of Guilin Medical University. We enrolled individuals who visited the Center of Health Examination, Affiliated Hospital of Guilin Medical University, Guilin, Guangxi, China, for routine check-ups from July to November 2012, and excluded individuals with severe cardiovascular and cerebrovascular disease, type 1 diabetes mellitus, gestational diabetes mellitus, other special types of diabetes mellitus, and liver or kidney impairment. Type 2 diabetes was defined according to the 1999 World Health Organization criteria[11]. A total of 6,660 individuals with complete data were finally included in the statistical analysis by the method of cluster random sampling, and all participants signed the informed consent form and accepted the cross-sectional survey. The data were analyzed to investigate the risk factors and their interactions in type 2 diabetes.

Measurements

Using the standard questionnaire to collect the basic information of participants, including the basic situation of respondents (name, sex, age, nationality, marital status etc.), living habits (diet, smoking history, history of alcohol intake etc.), medical histories, medications, family histories of disease and physical activity. The training qualified medical examiners measured height to the nearest 0.1 cm, and bodyweight with a calibrated balance scale to the nearest 0.1 kg. Body mass index (BMI) was calculated as bodyweight (kg)/height^2 (m^2). The participants rested for 5–10 min before their right brachial artery blood pressure was measured with the mercury sphygmomanometer. The blood pressure was measured twice with a 1-min interval, the average of the two measurements was used for the subsequent analysis and the numerical precision to 1 mmHg.

All participants were asked to fast overnight for at least 10 h, and two cubital venous blood samples (5 mL each) were collected the following day: one sample in a fasting state and the other sample 120 min after an oral glucose load (75 g). The blood samples were analyzed within 2 h of collection using the Roche Cobas C501 (Basel, Switzerland) automatic biochemical analyzer to determine blood biochemical indicators. An enzymatic assay was carried out to determine the uric acid (UA) concentration, a glycerol phosphate oxidase-peroxidase assay was carried out to measure triglycerides (TG), the cholesterol oxidase assay was used to measure the total cholesterol (TC), the surfactant assay was used to measure low-density lipoprotein cholesterol (LDL-C), the phosphotungstic acid-magnesium assay was used to measure high-density lipoprotein cholesterol (HDL-C), and the glucose oxidase assay was used to measure the fasting plasma glucose (FPG) and 2-h plasma glucose (2hPG) levels. Furthermore, the Mindray DC-6 Expert type II ultrasound system (Shenzhen, China) was used by trained qualified specialists to examine the liver.

Quality control

The trained staff administered a detailed questionnaire survey and recorded the responses; the sphygmomanometer and scale were centrally calibrated, the blood pressure, height and weight were centrally measured; the EpiData software, version 3.0.2 (EpiData Association, Odense, Denmark), was used to enter the data and carry out logistic testing; and the laboratory testing and quality control were carried out by the laboratory analysis center of our hospital.

Definitions

BMI was classified based on the Chinese criteria[12]: normal range (BMI < 24.0 kg/m^2) and overweight (BMI ≥ 24.0 kg/m^2). Hypertension was defined as systolic blood pressure ≥140 mmHg and/or the diastolic blood pressure ≥90 mmHg, or having a history of taking antihypertensive drugs[13]. The definition of hyperuricemia in men and women was UA >420 μmol/L and UA >360 μmol/L, respectively. According to the 1999 World Health Organization criteria[11], FPG ≥7.0 mmol/L and/or 2hPG ≥11.1 mmol/L in the oral glucose tolerance test could be diagnosed as diabetes. The diagnostic criteria for non-alcoholic fatty liver disease (NAFLD) was as

follows: no history of heavy drinking; no other NAFLD-inducing diseases, such as drug-induced liver disease or viral hepatitis; and liver ultrasound imaging findings suggestive of a diffuse fatty liver[14].

Statistical analysis

All analyses were carried out with the Spss software, version 18.0 (SPSS, Chicago, IL, USA). Categorical variables were expressed as percentages and analyzed with the χ^2-test. The CRT model is a non-parametric analysis that is applied to analyze potential interactions between continuous or categorical variables. The minimum sample size was 100 for the parent nodes of the classification tree and 50 for the child nodes. CRT analysis identified optimal cut-off points for continuous discriminating variables in the sets of growth limits and significance levels of splitting nodes, and merging categories at each stage. The CRT model beginning with a root node included all cases, the best discriminating variable with the maximum χ^2-value was selected as the first layer, and then subsequent variables were selected into the model if they contributed significantly to subtyping variables as the second layer, and so on. The tree branched and grew iteratively until a termination criterion was met or no further significant improvement in the classification of study participants was possible. The CRT selected outcome variables with close associations with type 2 diabetes from all influencing factors, and visually displayed their interactions in the form of a tree diagram. The effect of the classification tree-based model was evaluated based on the overall accuracy rate of the predicted categories. Non-conditional logistic regression analysis was carried out to analyze the interactions between influencing factors. The univariate logistic regression analysis used the additive model to quantitatively analyze, and the interactions were presented as the product term in the multiple logistic regression analysis. A *P*-value <0.05 was considered statistically significant.

RESULTS

Analysis of the classification tree model

To analyze factors affecting the prevalence of type 2 diabetes, participants with type 2 diabetes ($n = 338$) were assigned a value of '1', and participants without type 2 diabetes ($n = 6,322$) were assigned a value of '0'. CRT was used to analyze the relevant factors affecting type 2 diabetes. In the present study, the classification tree model had three layers of nine nodes including five terminal nodes. The results showed that nine factors (age, hypertension, NAFLD, BMI, UA, TG, TC, LDL-C and HDL-C) affected type 2 diabetes. However, after the classification tree was grown and pruned based on the restriction of the minimum sample size of the root and child nodes; as seen in Figure 1, the results showed just three variables with close associations with type 2 diabetes: age, TG and NAFLD. The three variables interacted with each other.

For the classification tree model, the estimated risk was 0.051 and the standard error was 0.003, showing that the model was

capable of accurately estimating 94.9% of type 2 diabetes patients. Thus, the fitting effect was good.

Comparison of the clinical and metabolic characteristics of the two groups

Based on the results from the classification tree model, type 2 diabetes was used as the dependent variable, and the independent variables with the closest relationship with type 2 diabetes were divided into categorical variables based on the cut-off points of the classification tree. The type 2 diabetes incidence was higher in participants aged >51.5 years, and thus the use of 1.775 mmol/L as the cut-off point of TG level, which was consistent with the diagnostic criteria of hypertriglyceridaemia[15]. The remaining six influencing factors were assigned a value according to professional guidelines, as shown in Table 1.

After assigning the value of categorical variables, type 2 diabetes (value = '1') was included in the case group, including 217 men (64.2%) and 121 women (35.8%). The remaining participants (value = '0') were divided into the control group, including 3,263 men (51.6%) and 3,059 women (48.4%). We used the χ^2-test to compare the clinical and metabolic characteristics of the two groups, and significant differences in the composition ratios of sex and age were observed between the two groups, and the incidences of high TG, NAFLD, hypertension, high BMI, high UA, high TC and high LDL-C were significantly higher in participants with type 2 diabetes than in participants without type 2 diabetes, whereas the incidence of high HDL-C was significantly lower in the participants with type 2 diabetes (Table 1).

Quantitative analysis of the interactions of risk factors for type 2 diabetes

An additive model was used for univariate logistic regression to quantitatively analyze the interactions of the three variables that had close associations with type 2 diabetes. The results showed that the risk of type 2 diabetes was 22.552-fold higher in participants with advanced age and high TG than in participants without advanced age and high TG. The risk of type 2 diabetes was 7.276-fold higher in participants with NAFLD and high TG than in participants without NAFLD and high TG. The risk of type 2 diabetes was 17.909-fold higher in participants with advanced age and NAFLD than in participants without advanced age and NAFLD (Table 2). In addition, hypertension, high BMI, high UA, high TC, high LDL-C and low HDL-C also were influencing factors for type 2 diabetes (Table 2). To adjust the confounding factors, the between-factor interaction presented as the product term was incorporated into the multivariate logistic regression analysis by the forward stepwise method of introduction ($\alpha_{in} = 0.10$, $\alpha_{out} = 0.15$), and the results showed that the following factors had a significant influence on type 2 diabetes: hypertension, high BMI, high TC, advanced age × NAFLD, advanced age × high TG and NAFLD × high TG (Table 3).
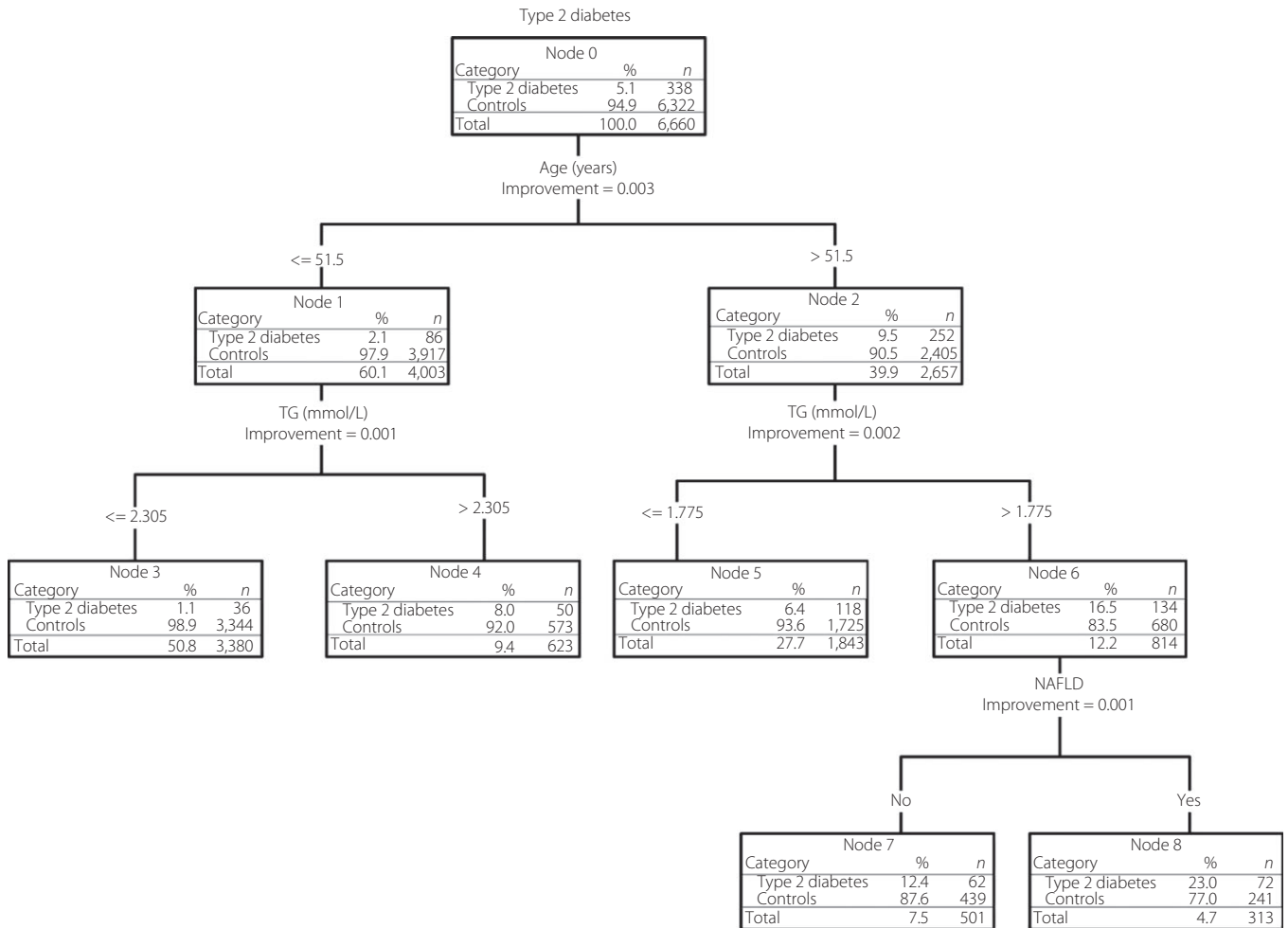
Type 2 diabetes

Figure 1 | The classification tree model for type 2 diabetes. Data are shown as the percentage (%) and number (n). Continuous variables were divided into categorical variables after the classification tree was grown and pruned based on the restriction of the minimum sample size of the root (100) and child (50) nodes. The figure shows just three variables with close associations with type 2 diabetes.

DISCUSSION

Identifying the factors that affect the prevalence of type 2 diabetes and interactions between these factors have important implications for elucidating the clinical and metabolic pathogenesis of type 2 diabetes, and for taking preventive measures.

The decision tree model is used to quickly and accurately identify specific populations, study the correlations between populations and predict future events. The classification tree model is primarily used to analyze discrete dependent variables, whereas the regression tree is primarily used to predict continuous dependent variables[9], and to select optimal branch thresholds by growing and pruning the tree[16]. Compared with the traditional analysis method, the classification tree model is not subject to the potential colinear effect of variables during the analysis. The classification tree model visually displays the results in the form of a tree diagram, processes the interactions between variables well[17], analyzes specific populations affected by each variable, shows the characteristics of populations prone

to disease[16] and enables interventions for high-risk populations[18].

The first layer of the classification tree was age, therefore the age was the most important factor affecting type 2 diabetes in the present study. In the Chinese population, type 2 diabetes tends to occur in individuals aged >50 years with a prevalence of approximately 9.7%[19]. Studies carried out in the USA[2] and Africa[20] have also shown that advanced age is a risk factor for type 2 diabetes, and that the prevalence of type 2 diabetes increases with age. In the present study, the cut-off age was 51.5 years. The results showed that the incidence of type 2 diabetes was significantly higher in participants aged >51.5 years (9.5%) compared with participants aged ≤51.5 years (2.1%). This result suggested that the type 2 diabetes incidence was higher in middle-aged and elderly populations, which was consistent with previous studies[19,20]. A pancreatic metabolic disorder might occur as a result of the aging of tissues and organs in the body, resulting in elevated blood glucose levels. Hence,

Table 1 | Clinical and metabolic characteristics of two groups

Variables	Without type 2 diabetes (n = 6,322) n (%)	With type 2 diabetes (n = 338) n (%)	χ^2-value	P-value
Sex				
Male	3,263 (51.6)	217 (64.2)	20.377	0.000
Age (years)				
>51.5	2,405 (38.0)	252 (74.6)	178.401	0.000
TG (mmol/L)				
>1.775	1,623 (25.7)	193 (57.1)	159.797	0.000
NAFLD				
Yes	1,122 (17.7)	151 (44.7)	150.469	0.000
Hypertension				
Yes	1,458 (23.1)	165 (48.8)	115.466	0.000
BMI (kg/m^2)				
≥24.0	3,051 (48.3)	244 (72.2)	73.496	0.000
UA (μmol/L)				
Male >420, Female >360	845 (13.4)	62 (18.3)	6.756	0.009
TC (mmol/L)				
≥5.18	2,235 (35.4)	187 (55.3)	55.308	0.000
LDL-C (mmol/L)				
≥3.37	2,386 (37.7)	204 (60.4)	69.040	0.000
HDL-C (mmol/L)				
<1.04	807 (12.8)	77 (22.8)	27.962	0.000

The sample size of each variable is shown as number (*n*) and percentage (%). *P*-values compare clinical and metabolic characteristics between 338 participants with type 2 diabetes and 6,322 participants without type 2 diabetes, using the χ^2-test. BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; NAFLD, non-alcoholic fatty liver disease; TC, total cholesterol; TG, triglyceride; UA, uric acid.

Table 2 | Univariate logistic regression analysis for type 2 diabetes

Variables		Controls n (%)	Type 2 diabetes n (%)	OR (95% CI)
TG	Age			
−	−	2,967 (46.9)	26 (7.7)	1.000
−	+	1,732 (27.4)	119 (32.5)	7.840 (5.108–12.034)*
+	−	950 (15.0)	60 (17.8)	7.207 (4.523–11.485)*
+	+	673 (10.6)	133 (39.3)	22.552 (14.692–34.617)*
NAFLD	TG			
−	−	4,184 (66.2)	108 (32.0)	1.000
−	+	1,016 (16.1)	79 (23.4)	3.012 (2.236–4.059)*
+	−	515 (8.1)	37 (10.9)	2.783 (1.895–4.088)*
+	+	607 (9.6)	114 (33.7)	7.276 (5.518–9.594)*
NAFLD	Age			
−	−	3,284 (51.9)	39 (11.5)	1.000
−	+	1,916 (30.3)	148 (43.8)	6.504 (4.550–9.297)*
+	−	633 (10.0)	47 (13.9)	6.252 (4.055–9.640)*
+	+	489 (7.7)	104 (30.8)	17.909 (12.246–26.190)*
Hypertension	−	4,864 (76.9)	173 (51.2)	1.000
	+	1,458 (23.1)	165 (48.8)	3.182 (2.551–3.969)*
BMI	−	3,271 (51.7)	94 (27.8)	1.000
	+	3,051 (48.3)	244 (72.2)	2.783 (2.183–3.548)*
TC	−	4,087 (64.4)	151 (44.7)	1.000
	+	2,235 (35.4)	187 (55.3)	2.265 (1.816–2.823)*
UA	−	5,477 (86.6)	276 (81.7)	1.000
	+	845 (13.4)	62 (18.3)	1.456 (1.095–1.936)*
LDL-C	−	3,936 (62.3)	134 (39.6)	1.000
	+	2,386 (37.7)	204 (60.4)	2.511 (2.008–3.141)*
HDL-C	+	5,515 (87.2)	261 (77.2)	1.000
	−	807 (12.8)	77 (22.8)	2.016 (1.547–2.627)*

The univariate logistic regression model used the additive model to quantitatively analyze the interactions among age, non-alcoholic fatty liver disease (NAFLD), triglycerides (TG) and other influencing factors. Compared with the reference, *P < 0.05. −, 0; +, 1; BMI, body mass index; CI, confidence interval; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; OR, odds ratio; TC, total cholesterol; UA, uric acid.

efforts should be made to educate middle-aged and elderly populations about diabetes mellitus to reduce the incidence of type 2 diabetes.

The second layer of the classification tree was TG. The present study showed that participants with high TG (>1.775 mmol/L) had a significantly higher incidence of type 2 diabetes. Multivariate logistic regression analysis found that high TC was also a risk factor for type 2 diabetes. Both high TG and high TC are manifestations of dyslipidemia, and studies have shown that many type 2 diabetes patients have concurrent dyslipidemia[21]. Plasma glucose and lipid metabolisms are important components of energy metabolism, and these two factors are interrelated and influence each other[22]. Dyslipidemia is a link between type 2 diabetes and cardiovascular disease[23], and increases the risk of vascular disease in type 2 diabetes patients. Therefore, dyslipidemia control and the understanding of potential influencing mechanisms might provide new strategies for the prevention and treatment of type 2 diabetes.

The third layer of the classification tree was NAFLD. NAFLD is a clinical and pathological syndrome characterized by diffuse liver parenchymal cell steatosis and TG deposits[24]. Patients with NAFLD have no history of heavy drinking or other specific causes of liver damage. The present study found that the presence of NAFLD significantly increased the incidence of type 2 diabetes for participants whose age was >51.5 years and whose TG was >1.775 mmol/L. A series of studies have shown that NAFLD is closely related to type 2 diabetes and its complications, and NAFLD-associated oxidative stress, lipid peroxidation and inflammation are key factors leading to insulin resistance[25], dyslipidemia and subsequently diabetes mellitus[26], metabolic syndrome[27], and other cardiovascular diseases[28]. Furthermore, type 2 diabetes might promote the development of NAFLD. Studies have shown that a low-fat diet, lifestyle change, and more exercise can improve liver function and mitigate insulin resistance in liver, thereby

Table 3 | Multivariate logistic regression analysis for type 2 diabetes

Variables	B	SE	Wald	P-value	OR	95% CI
High BMI	0.479	0.135	12.595	0.000	1.615	1.239~2.104
Hypertension	0.648	0.123	27.787	0.000	1.912	1.503~2.434
High TC	0.407	0.120	11.618	0.001	1.503	1.189~1.900
High TG × advanced age	0.916	0.149	38.031	0.000	2.499	1.868~3.344
NAFLD × advanced age	0.223	0.090	6.164	0.013	1.250	1.048~1.491
NAFLD × high TG	0.299	0.084	12.689	0.000	1.349	1.144~1.590

Adjusted for age (≤51.5 or >51.5 years), hypertension (yes or no), non-alcoholic fatty liver disease (NAFLD; yes or no), body mass index (BMI; <24.0 or ≥24.0 kg/m^2), uric acid (UA; male ≤420, female ≤360 or male >420, female >360 μmol/L), triglycerides (TG; ≤1.775 or >1.775 mmol/L), total cholesterol (TC; <5.18 or ≥5.18 mmol/L), low-density lipoprotein cholesterol (LDL-C; <3.37 or ≥3.37 mmol/L) and high-density lipoprotein cholesterol (HDL-C; <1.04 or ≥1.04 mmol/L). B, correlation coefficient; CI, confidence interval; OR, odds ratio; SE, standard error; Wald, Wald statistics for logistic regression analysis.

treating NAFLD[29] and exerting positive effects on type 2 diabetes.

Based on the cut-off values of the classification tree and professional guidelines, we assigned the categorical variables, and counted the sample size of each variable to compare the clinical and metabolic characteristics of the two groups. The results showed that type 2 diabetes patients had higher age and incidences of high TG, NAFLD, hypertension, high BMI, high UA, high TC, high LDL-C, and low HDL-C, showing that patients with type 2 diabetes might easily suffer from metabolic disorders, such as hypertension, overweight, hyperuricemia, dyslipidemia, NAFLD and so on. Therefore, a further study by the non-conditioned logistic regression analysis on the meaningful factors was carried out.

We used the logistic regression model to quantitatively analyze the interactions to compensate for the inadequacy of the classification tree in the present study. OR (A) and OR (B) represented the odds ratio for only one variable existing, OR (AB) represented the odds ratio for the existence of both variables. The results of the univariate logistic regression model suggested that advanced age and high TG, advanced age and NAFLD, and NAFLD and high TG had positive interactions on the risk of type 2 diabetes (OR [AB] > OR [A] + OR [B] − 1), which were consistent with the results of the classification tree model. Thus, advanced age, high TG and NAFLD were independent risk factors for type 2 diabetes. Furthermore, the between-factor interaction exerted a synergistic effect on the incidence of type 2 diabetes; that is, the presence of any two risk factors was associated with a higher incidence of type 2 diabetes than any factor alone. After controlling for confounding factors, the multivariate logistic regression analysis showed interactive effects between advanced age × high TG, advanced age × NAFLD and NAFLD × high TG on the type 2 diabetes incidence. Additionally, hypertension, high BMI and high TC also were risk factors for type 2 diabetes, suggesting that type 2 diabetes resulted from the interactions of environmental, genetic and behavioral factors. Regardless of the uncontrollable factors, such as age, type 2 diabetes could be controlled by taking comprehensive measures, such as improving lifestyle, alleviating hyperlipidemia, reducing the prevalence of NAFLD and so on.

The present study had some limitations. First, all included participants were from Guilin, China, and thus they might not represent the whole Chinese population; second, prospective studies are required to investigate the causal relationship between the influencing factors and type 2 diabetes; and third, this was an epidemiological survey, and thus the exact mechanism underlying the between-factor interaction requires further investigation.

In the present study, we conclude that type 2 diabetes results from the interactions of many factors, and the interactions among age, TG and NAFLD are important risk factors for type 2 diabetes. These findings have important implications for researchers studying the etiology of type 2 diabetes, who should take targeted preventive measures and clarify the relevant mechanisms of action in future research.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet* 2005; 365: 1333–1346.
2. Menke A, Casagrande S, Geiss L, *et al.* Prevalence of and trends in diabetes among adults in the United States, 1988–2012. *JAMA* 2015; 314: 1021–1029.
3. Yu S, Sun Z, Zheng L, *et al.* Prevalence of diabetes and impaired fasting glucose in hypertensive adults in rural

China: far from leveling-off. *Int J Environ Res Public Health* 2015; 12: 14764–14779.

4. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract* 2010; 87: 4–14.

5. Guariguata L, Whiting DR, Hambleton I, et al. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 2014; 103: 137–149.

6. Tanaka S, Tanaka S, Iimuro S, et al. Predicting macro- and microvascular complications in type 2 diabetes: the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine. *Diabetes Care* 2013; 36: 1193–1199.

7. Rigalleau V, Cougnard-Gregoire A, Nov S, et al. Association of advanced glycation end products and chronic kidney disease with macroangiopathy in type 2 diabetes. *J Diabetes Complications* 2015; 29: 270–274.

8. World Health Organization. Global Status Report on Noncommunicable Diseases: 2010. Geneva: World Health Organization, 2011.

9. Xie J, Hu D, Yu D, et al. A quick self-assessment tool to identify individuals at high risk of type 2 diabetes in the Chinese general population. *J Epidemiol Community Health* 2010; 64: 236–242.

10. Omurlu IK, Ture M, Unubol M, et al. Comparing Performances of logistic regression, classification & regression trees and artificial neural networks for predicting albuminuria in Type 2 diabetes mellitus. *Int J Sci Basic Appl Res* 2014; 16: 173–187.

11. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; 15: 539–553.

12. Zhou BF. Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults–study on optimal cut-off points of body mass index and waist circumference in Chinese adults. *Biomed Environ Sci* 2002; 15: 83–96.

13. World Health Organization. International Society of Hypertension guidelines for the management of hypertension. Guidelines Sub-Committee. *Blood Press Suppl* 1999; 1: 9–43.

14. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. *Gastroenterology* 2012; 142: 1592–1609.

15. Hegele RA, Ginsberg HN, Chapman MJ, et al. The polygenic nature of hypertriglyceridaemia: implications for definition, diagnosis, and management. *Lancet Diabetes Endocrinol* 2014; 2: 655–666.

16. Henrard S, Speybroeck N, Hermans C. Classification and regression tree analysis vs. multivariable linear and logistic regression methods as statistical tools for studying haemophilia. *Haemophilia* 2015; 21: 715–722.

17. Camp NJ, Slattery ML. Classification tree analysis: a statistical tool to investigate risk factor interactions with an example for colon cancer (United States). *Cancer Causes Control* 2002; 13: 813–823.

18. Chin MP, Wrolstad D, Bakris GL, et al. Risk factors for heart failure in patients with type 2 diabetes mellitus and stage 4 chronic kidney disease treated with bardoxolone methyl. *J Card Fail* 2014; 20: 953–958.

19. Yang W, Lu J, Weng J, et al. Prevalence of diabetes among men and women in China. *N Engl J Med* 2010; 362: 1090–1101.

20. Werfalli M, Musekiwa A, Engel ME, et al. The prevalence of type 2 diabetes mellitus among older people in Africa: a systematic review study protocol. *BMJ Open* 2014; 4: e004747.

21. Parhofer KG. Interaction between Glucose and Lipid Metabolism: more than Diabetic Dyslipidemia. *Diabetes Metab J* 2015; 39: 353–362.

22. ACCORD Study Group, Ginsberg HN, Elam MB, et al. Effects of combination lipid therapy in type 2 diabetes mellitus. *N Engl J Med* 2010; 362: 1563–1574.

23. Li Y, Mu Y, Ji Q, et al. Hypoglycaemia, Abnormal Lipids, and Cardiovascular Disease among Chinese with Type 2 Diabetes. *Biomed Res Int* 2015; 2015: 862896.

24. Schreuder TC, Verwer BJ, van Nieuwkerk CM, et al. Nonalcoholic fatty liver disease: an overview of current insights in pathogenesis, diagnosis and treatment. *World J Gastroenterol* 2008; 14: 2474–2486.

25. Bugianesi E, Moscatiello S, Ciaravella MF, et al. Insulin resistance in nonalcoholic fatty liver disease. *Curr Pharm Des* 2010; 16: 1941–1951.

26. Jia G, Di F, Wang Q, et al. Non-Alcoholic Fatty Liver Disease Is a Risk Factor for the Development of Diabetic Nephropathy in Patients with Type 2 Diabetes Mellitus. *PLoS One* 2015; 10: e0142808.

27. Fabbrini E, Sullivan S, Klein S. Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology* 2010; 51: 679–689.

28. Bonapace S, Valbusa F, Bertolini L, et al. Nonalcoholic fatty liver disease is associated with aortic valve sclerosis in patients with type 2 diabetes mellitus. *PLoS One* 2014; 9: e88371.

29. Al-Jiffri O, Al-Sharif FM, Abd El-Kader SM, et al. Weight reduction improves markers of hepatic function and insulinresistance in type-2 diabetic patients with non-alcoholic fatty liver. *Afr Health Sci* 2013; 13: 667–672.