


Analysis of the effect of glutamyltransferase on hyperlipidemia based on decision tree

DIGITAL HEALTH
Volume 9: 1–12
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076231185441
journals.sagepub.com/home/dhj



Tingting Zhang^{1,2} , Dantong Ouyang^{1,2}, Chenglin Sun^{3,4}, Yaru Bi³, Lili He^{1,2} and Hongtao Bai²

Abstract

Objectives: This study is designed to analyze the potential influencing factors of hyperlipidemia, and to explore the relationship between liver function indicators such as gamma-glutamyltransferase (GGT) and hyperlipidemia.

Methods: Data were derived from 7599 outpatients who visited the Department of Endocrinology of the First Hospital of Jilin University (2017–2019). A multinomial regression model is used to identify related factors of hyperlipidemia and the decision tree method is used to explore the general rules in hyperlipidemia patients and non-hyperlipidemia patients on these factors.

Results: The average of age, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure, aspartate aminotransferase, alanine aminotransferase (ALT), GGT and glycosylated hemoglobin (HbA1c) in the hyperlipidemia group are higher than those in the non-hyperlipidemia group. In multiple regression analysis, SBP, BMI, fasting plasma glucose, 2-h postprandial blood glucose, HbA1c, ALT, GGT are associated with triglyceride. For people with HbA1c less than 6.0%, controlling GGT within 30 IU/L reduces the prevalence of hypertriglyceridemia by 4%, and for people with metabolic syndrome with impaired glucose tolerance controlling GGT within 20 IU/L reduces the prevalence of hypertriglyceridemia by 11%.

Conclusions: Even when GGT is in the normal range, the prevalence of hypertriglyceridemia increases with its gradual increase. Controlling GGT in people with normoglycemia and impaired glucose tolerance can reduce the risk of hyperlipidemia.

Keywords

Glutamyltransferase, hyperlipidemia, public health, prevention, decision tree

Submission date: 18 October 2022; Acceptance date: 14 June 2023

Introduction

With the changes of people's lifestyle, the incidence of hyperlipidemia is increasing year by year. Hyperlipidemia is a major risk factor for cardiovascular and cerebrovascular diseases,^{1,2} type 2 diabetes,³ obesity,⁴ some cancers,⁵ etc. It can increase the risk of all-cause death from cardiovascular diseases by 1.7–8.2 times,^{6,7} and a 10% reduction in serum cholesterol level can reduce the risk of cardiovascular diseases by 40% at the age of 50.⁸ Therefore, it is particularly important to identify the indexes related to blood lipid and pay attention to the changes of the indexes.

Hyperlipidemia is mainly affected by diet, exercise, genetics, smoking, blood glucose, hypertension, obesity,

¹College of Computer Science and Technology, Jilin University, Changchun, China

²Symbol Computation and Knowledge Engineer of Ministry of Education, Jilin University, Changchun, China

³Department of Endocrinology and Metabolism, First Hospital of Jilin University, Changchun, China

⁴Department of Clinical Nutrition, First Hospital of Jilin University, Changchun, China

Corresponding author:

Hongtao Bai, College of Computer Science and Technology, Jilin University, Changchun 130021, China; Symbol Computation and Knowledge Engineer of Ministry of Education, Jilin University, Changchun 130021, China.
Email: baiht@jlu.edu.cn



chronic kidney disease, hyperuricemia,⁹ fatty liver and other factors. Among them, high-fat diet,¹⁰ lack of exercise, obesity, type 2 diabetes are the main risk factors. The effects of high-fat diet,¹⁰ obesity^{11,12} and blood glucose¹³ on the risk of hyperlipidemia are relatively clear. In addition to the above risk factors, Suhad M's study suggests that elevated liver transaminase is not only closely related to type 2 diabetes, but also a risk factor for hyperlipidemia.¹⁴ Alanine aminotransferase (ALT), aspartate aminotransferase (AST) and γ -GGT are often used as indicators of liver damage. Gamma-glutamyltransferase (GGT) exists in the epithelial cells of the liver and biliary tract and participates in the degradation of the extracellular antioxidant glutathione,¹⁵ and is also considered as a marker of oxidative stress.¹⁶ A growing body of evidence suggests that GGT levels are closely associated with stroke,¹⁷ atherosclerotic cardiovascular disease,^{18–20} type 2 diabetes²¹ and obesity.²² Kazuhiko Kotani, Tang and Yuli et al. prove that GGT is closely associated with lipid metabolism, and GGT is significantly correlated with triglyceride and high-density lipoprotein.^{23–25} However, some studies have not been clear about the correlation between GGT and blood lipid. Different studies have different conclusions on the relationship between GGT and blood lipid, which may be caused by regional differences or the limited number of people included in the study.

This study incorporates the laboratory testing data of over 7000 outpatients for cross-sectional analysis, the objectives of this study are to identify the related factors of hyperlipidemia and to explore the influence of GGT and other liver function indicators on the prevalence of hyperlipidemia. This study also attempt to discuss the control goals needed to prevent hyperlipidemia.

Methods

Data sources

Data were obtained from 7599 outpatients who visited the Department of Endocrinology of the First Hospital of Jilin University (2017–2019). Basic information of the patients such as height and weight were collected through questionnaires, and blood samples were collected to obtain biochemical indicators and glycosylated hemoglobin (HbA1c).

Data cleaning

Due to a large number of missing values and outliers in the original data set, the data set is cleaned first to ensure the effectiveness and robustness of the subsequent analysis. For missing values, we delete the samples missing key factors, so as to obtain a relatively complete sample set. Outliers are those points that are far away from high-density groups in the sample space. From the perspective of statistics, if there are only a few sparsely distributed samples in a

certain region in the sample space, it means that the probability of data falling in this region is very low in reality. Moreover, the proportion of such data in the whole data set is too small to be universal, so it is not within the scope of this study. The isolated forest method²⁶ is used to detect outliers.

The idea of isolated forest is to use a random hyperplane to cut the data space in two, then use a random hyperplane to cut each subspace, and perform the above steps recursively until there is only one data point in each subspace. Since outliers are usually far from high-density groups in the data space, these outliers are quickly divided into a separate subspace, while non-outliers require more cuts. In this way, outliers can be quickly distinguished. Isolated forest method can be roughly divided into two stages: (i) Construct t isolated trees in the form of binary trees to form isolated forest; (ii) Calculate the average height of every sample point in each tree, and calculate the score of abnormality.

To ensure that the outliers in each feature dimension can be fully detected, we choose to combine the features in pairs (age-BMI, SBP-Cre, ALT-AST, GGT-HbA1c) to construct four two-dimensional data spaces. In one two-dimensional space, we use random forest method to identify the outliers and delete the samples containing the outliers. Then the remaining data is detected and processed in the data space composed of the other two features until all the features are filtered.

The construction process of an isolated tree is described as follows: Randomly select φ sample points in the data set and put them into the root node of the tree; Randomly generate a straight line in the current two-dimensional space to divide the current data space into two subspaces. The points above the line are drawn to the left child node of the current node, and the points below the line are drawn to the right child node; Perform the above steps recursively in each child node until there is only one sample point in the child node or the isolated tree has reached the specified height.

After the isolated forest is constructed, it can be used to evaluate the anomaly degree of each sample point in the data space. The calculation formula is:

$$s(x, \psi) = 2 \frac{E(h(x))}{c(\psi)} \quad (1)$$

where $h(x)$ represents the average height of x in all isolated trees, and $c(\psi)$ is defined as

$$c(\psi) = \begin{cases} 2H(\psi - 1) - \frac{2(\psi - 1)}{\psi}, & \psi > 2 \\ 1, & \psi = 2 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$H(i)$ is the harmonic number, which can be estimated by $\ln(i) + 0.5772156649$.

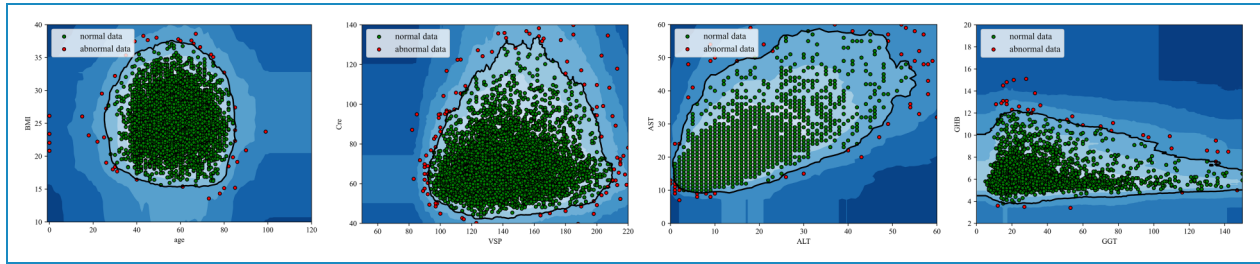


Figure 1. Outlier detection. Isolated forest acts on the two-dimensional data space composed of age-BMI, VSP-CRP, etc. The red points are scattered outside the green cluster and are the abnormal points identified by isolated forest.

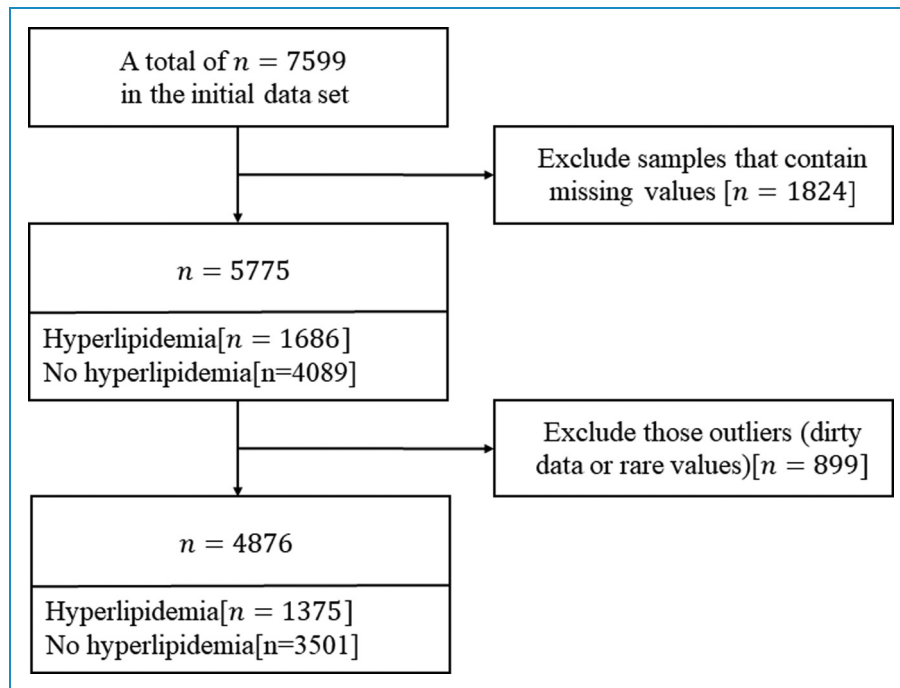


Figure 2. Flowchart of study population.

The schematic diagram of random forest anomaly detection is shown in Figure 1. The overall processing flow of the data set is shown in Figure 2.

Decision tree

Since the prevalence of hypertriglyceridemia is the highest in the study population, and the relationship between hypertriglyceridemia and GGT and HbA1c is stronger, we explore the general rules of the above indicators in the non-hypertriglyceridemia population and try to give advice on maintaining health. We use the binary classification method of decision tree²⁷ to extract different rules of the two types of samples, and then screen and artificially modify the extracted rules to find the general rules of the non-hyperlipidemia population.

Decision tree is a common classification and regression method with tree structure. The classification process can be

regarded as a set of if-then rules. Each node of the tree is a judgment condition, and the classification process is equivalent to starting from the root node of the tree model and judging at every node on the path until reaching the leaf node, that is, the classification process is completed. The judgment conditions on the path are combined to form a rule. Compared with the black-box characteristic of deep learning models such as neural networks, decision tree algorithm is still commonly used in medical, financial and other fields requiring high interpretability.

The decision tree generation process mainly includes the following two steps: (i) Feature selection: select one from the feature set as the split criterion for the current node, and there are multiple evaluation criteria for feature selection; (ii) Decision tree construction: according to the selected feature evaluation criteria, recursively generate child-nodes from top to bottom until the data set is inseparable.

The C4.5 algorithm is used here.²⁸ The algorithm uses the information gain rate as the feature selection criterion. The information gain rate is defined as:

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)} \quad (3)$$

Among them, *Gain* represents the information gain, which represents the degree to which a certain feature is known to reduce the uncertainty of the information of class *X*. The information gain of feature *A* to data set *S* *Gain*(*S*, *A*) is the difference between the entropy *H*(*S*) of set *S* and the conditional entropy of *S* under the conditions given by feature *A*. *SplitInfo*(*S*, *A*) is used to measure the breadth and uniformity of the attribute *A* split data set *S*.

$$Gain(S, A) = H(S) - H(S|A) \quad (4)$$

$$SplitInfo(S, A) = - \sum_{i=1}^C \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (5)$$

Results

Study population

Of the remaining 4876 samples, 1375 (Figure 2) has hyperlipidemia with a prevalence of 28.2%, of which the prevalence rate of hypertriglyceridemia is the highest (20.8%). There is no statistical significance between the hyperlipidemia group and the non-hyperlipidemia group in terms of gender ($P > 0.05$) (Table 1). Compared with the non-hyperlipidemia group, the hyperlipidemia group has higher body mass index (BMI), systolic blood pressure

Table 1. Baseline characteristics of the study population.

Variables	Overall (n = 4876)	Hyperlipidemia (n = 1375)	Non- hyperlipidemia (n = 3501)	P-value
Age(years) at survey	57.0 ± 9.2	57.8 ± 8.8	56.6 ± 9.3	<0.001
Age > 60 years, N(%)	1712(35.1)	518(37.7)	1194(34.1)	<0.001
Gender	Male, N(%)	3350(68.7)	961(69.9)	0.944
	Female, N(%)	1526(31.3)	414(30.1)	
BMI (kg/m ²)	25.0 ± 3.0	25.7 ± 3.0	24.8 ± 3.0	<0.001
Obesity, N(%)	803(16.5)	302(22.0)	501(14.3)	<0.001
SBP (mmHg)	140.1 ± 19.7	143.4 ± 19.6)	138.8 ± 19.6	<0.001
DBP (mmHg)	80.2 ± 11.5	82.1 ± 11.7	79.5 ± 11.4	<0.001
FPG (mmol/L)	6.13(5.22–6.34)	6.5(5.25–6.82)	5.98(5.28–6.18)	<0.001
2hPG (mmol/L)	8.58(5.56–9.9)	9.59(6.41–11.6)	8.18(5.83–9.25)	<0.001
GGT (IU/L)	21(15–31)	26.5(19–40)	19.0(14–28)	<0.001
ALT (IU/L)	12(9–16)	14(10–19)	11(8–16)	<0.001
AST (IU/L)	19(17–23)	20(17–24)	19(16–22)	<0.001
Glycemic status	Normoglycemia, N(%)	1686(34.6)	333(24.2)	<0.001
	Prediabetes, N(%)	2236(45.8)	656(47.7)	
	Diabetes, N(%)	954(19.5)	386(28.1)	

Data for continuous variables are expressed as mean ± standard deviation or mean (quartile range), and categorical variables are expressed as numbers(percentages). Obesity is defined as a body mass index ≥ 28 kg/m². According to the diagnostic criteria for diabetes released by WHO in 2011, HbA1c can be used to diagnose diabetes, and the diagnostic cut-off point is HbA1c ≥ 6.5%. According to ADA standard, 5.7–6.4% HbA1c can be used as one of the diagnostic criterias for prediabetes.

Abbreviations: BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; FPG: fasting plasma glucose; 2hPG: 2-h postprandial blood glucose; GGT: glutamyltransferase; ALT: alanine aminotransferase; AST: aspartate aminotransferase.

(SBP), diastolic blood pressure (DBP), fasting plasma glucose (FPG), 2-h postprandial blood glucose (2hPG), ALT, AST, GGT, and the difference between the two groups in these indicators is statistically significant ($P < 0.05$). In the GGT quintile,²⁹ BMI, SBP, DBP, triglyceride (TG), and total cholesterol (TC) increased from the lowest quintile to the highest quintile, and high-density lipoprotein (HDL) showed an overall downward trend (Table 2). The higher the GGT, the higher the FPG and 2hPG, which is statistically significant ($P < 0.05$).

Effect of GGT on the prevalence of hyperlipidemia

To explore the relationship between GGT and various types of hyperlipidemia, we divide the data set into ten roughly equal parts according to GGT and compare the prevalence of hyperlipidemia among each group (Figure 3(a)). With the progressive increase of GGT, the prevalence of hypertriglyceridemia, hypercholesterolemia and low-HDL increased, and the prevalence is statistically significant among the ten groups ($P < 0.05$). Compared with hypercholesterolemia and low-HDL, GGT has a more significant effect on the prevalence of hypertriglyceridemia, with the prevalence of hypertriglyceridemia in the highest decile of GGT 11 times higher as in the lowest decile (3.5% vs 38.9%). The effect of GGT on the prevalence of hypercholesterolemia and low-HDL is not significant, but the overall prevalence increases with increasing GGT, with the prevalence of hypercholesterolemia in the highest decile three times (6.9% vs 19.6%) and twice (9.4% vs 21.8%) as in the lowest decile, respectively. In addition to GGT, we also analyze the effects of ALT and AST on the prevalence of hyperlipidemia. Consistent with the above, the data set is divided into ten groups roughly equal according to ALT and AST respectively, and the prevalence of the three types of hyperlipidemia is statistically significant among the ten groups ($P < 0.05$). With the progressive elevation of ALT, the prevalence of hypertriglyceridemia and hypercholesterolemia gradually increases, and the prevalence of the highest decile of ALT is four times (10.8% vs 41.5%) and twice (7.4% vs 16.2%) of the lowest decile (Figure 3(b)). ALT is U-shaped associated with the prevalence of low-HDL. With the progressive elevation of AST, the prevalence of hypertriglyceridemia and hypercholesterolemia generally increases, with local fluctuations. The prevalence of the highest AST decile is about twice that of the lowest AST decile (15.2% vs 32%, 6.8% vs 15.1%) (Figure 3(c)). AST is U-shaped associated with the prevalence of low-HDL.

The combined effect of GGT and blood glucose on the prevalence of hyperlipidemia

Compared with the population with normoglycemia and the lowest GGT level, the prevalence of hyperlipidemia in the

population with the highest GGT level and diabetes is significantly increased (Figure 4). This phenomenon is most pronounced in hypertriglyceridemia, and the prevalence has increased by about 22 times (2.3% vs 50%). The prevalence of hypertriglyceridemia corresponding to the quintile of GGT is observed under the same blood glucose state, and it is found that the effect of GGT on hypertriglyceridemia is independent of blood glucose, and the prevalence always increases with the increase of GGT, with statistical significance ($P < 0.05$). Under the same GGT level, the prevalence of hypertriglyceridemia in the diabetic group is always higher than that in the normoglycemia group, and the prevalence of hypertriglyceridemia in the prediabetic group is between them. With the same blood glucose status, the overall prevalence of hypercholesterolemia increases with the increase of GGT level, and fluctuates in the diabetic group, with statistical significance ($P < 0.001$ in the normoglycemia group; $P = 0.002$ in the prediabetic group; $P = 0.003$ in the diabetic group). In normoglycemia group and prediabetes group, the prevalence of low-HDL increases with the increase of GGT level, and the difference is statistically significant ($P < 0.05$). The prevalence of low-HDL is not statistically significant in the diabetic group.

Correlation analysis of various variables and triglyceride levels

We conduct a Spearman correlation analysis to quantify the correlation between various indicators and triglyceride levels, so as to select indicators with higher correlation for subsequent analysis.

Table 3 shows the results of Spearman's correlation analysis. Among them, SBP ($r = 0.11$, $P < 0.001$), DBP ($r = 0.14$, $p < 0.001$), BMI ($r = 0.178$, $P < 0.001$), FPG ($r = 0.154$, $P < 0.001$), 2hPG ($r = 0.178$, $P < 0.001$), HbA1c ($r = 0.179$, $P < 0.001$), ALT ($r = 0.19$, $P < 0.001$), AST ($r = 0.108$, $P < 0.001$), GGT ($r = 0.276$, $P < 0.001$) and triglyceride are significantly correlated.

Table 4 shows the results of multiple regression analyses. SBP, BMI, FPG, 2hPG, HbA1c, ALT, GGT are significantly correlated with triglyceride.

General rules in non-hypertriglyceridemia population

Figure 5 shows the classification decision tree constructed with the data of the study population as input. In order to measure the classification accuracy of decision tree, the ROC curve is drawn (Figure 6), and the model finally achieved an accuracy of 73.2%, indicating that the classification results and the rules formed by decision tree have certain reference value.

Table 5 shows the general rules of the non-triglyceride population extracted from the decision tree and manually

Table 2. Characteristics of the study population by categories of GGT levels.

Variables	GGT					P-value
	Quartile 1 (n = 775)	Quartile 2 (n = 1146)	Quartile 3 (n = 907)	Quartile 4 (n = 1058)	Quartile 5 (n = 990)	
GGT (IU/L)	11.5(11-13)	15.9(15-17)	20.9(20-22)	28.1(25-31)	56.5(39-63)	
GGT (IU/L), cutoff in men	<13	13-17	17-22	22-30	≥30	
GGT (IU/L), cutoff in women	<18	18-24	24-30	30-45	≥45	
Age (years) at survey	55.3 ± 9.8	57.6 ± 9.3	58.2 ± 9.0	57.3 ± 8.6	56.0 ± 8.8	0.824
Age > 60 years, N(%)	272(35.1)	402(35.3)	318(35.1)	372(35.2)	348(35.2)	0.775
Gender						
Male, N(%)	687(88.6)	903(78.8)	632(69.7)	627(59.3)	501(50.6)	<0.001
Female, N(%)	88(11.4)	243(21.2)	275(30.3)	431(40.7)	489(49.4)	
BMI (kg/m ²)	23.7 ± 2.8	24.4 ± 2.9	25.3 ± 2.8	25.5 ± 2.9	26.0 ± 3.0	<0.001
Obesity, N(%)	53(6.9)	126(11.0)	152(16.8)	204(19.3)	268(27.1)	<0.001
SBP (mmHg)	132.5 ± 19	138.4 ± 19.3	142 ± 19.2	142.3 ± 19.7	144.0 ± 19.2	<0.001
DBP (mmHg)	76.2 ± 10.6	78.6 ± 11.3	80.7 ± 11.0	81.2 ± 11.4	83.7 ± 11.8	<0.001
FPG (mmol/L)	5.6(5.04-5.8)	5.9(5.17-6.08)	6.2(5.27-6.46)	6.3(5.28-6.57)	6.6(5.25-7.03)	<0.001
2hPG (mmol/L)	7.3(5.6-8.04)	8.0(5.8-8.86)	8.8(6.11-10.2)	9.0(6.24-10.6)	9.6(6.23-11.8)	<0.001
ALT (IU/L)	9(7-12)	11(8-14)	12(9-15)	13(10-18)	17(12-23)	<0.001
AST (IU/L)	18(15-20)	18(16-21)	19(16-22)	20(17-23)	22(19-26)	<0.001
HDL (mmol/L)	1.38 ± 0.29	1.33 ± 0.29	1.27 ± 0.28	1.25 ± 0.28	1.26 ± 0.29	<0.001
TC (mmol/L)	4.84 ± 0.89	5.06 ± 0.9	5.13 ± 0.92	5.19 ± 0.92	5.36 ± 0.97	<0.001
TG (mmol/L)	1.18(0.81-1.37)	1.52(0.94-1.78)	1.76(1.07-2.11)	2(1.18-2.29)	2.3(1.31-2.79)	<0.001

(continued)

Table 2. Continued.

Variables	GGT					P-value
	Quartile 1 (n = 775)	Quartile 2 (n = 1146)	Quartile 3 (n = 907)	Quartile 4 (n = 1058)	Quartile 5 (n = 990)	
Glycemic status						
Normoglycemia, N(%)	348(44.9)	440(38.4)	264(29.1)	316(29.9)	318(32.1)	<0.001
Prediabetes, N(%)	353(45.6)	548(47.8)	441(48.6)	496(46.9)	398(40.2)	
Diabetes, N(%)	74(9.5)	158(13.8)	202(22.3)	246(23.2)	274(27.7)	
Hyperlipidemic type						
Hypertriglyceridemia, N(%)	36(4.6)	143(12.5)	190(21)	273(25.8)	373(37.7)	<0.001
Hypercholesterolemia, N(%)	56(7.2)	108(9.4)	98(10.8)	133(12.6)	175(17.7)	<0.001
Low HDL, N(%)	74(9.6)	146(12.7)	178(19.6)	220(20.8)	229(23.1)	<0.001

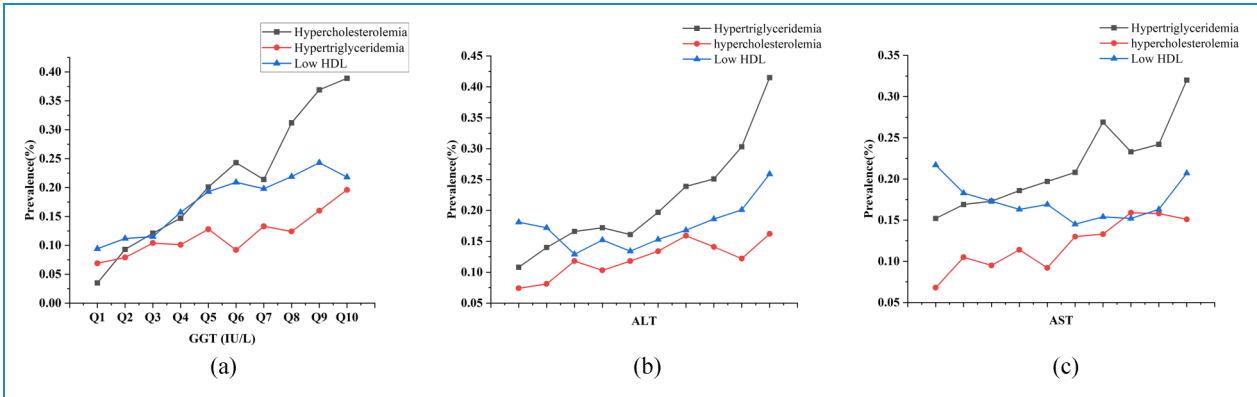


Figure 3. The influence of liver enzymes on the prevalence of three types of hyperlipidemia.

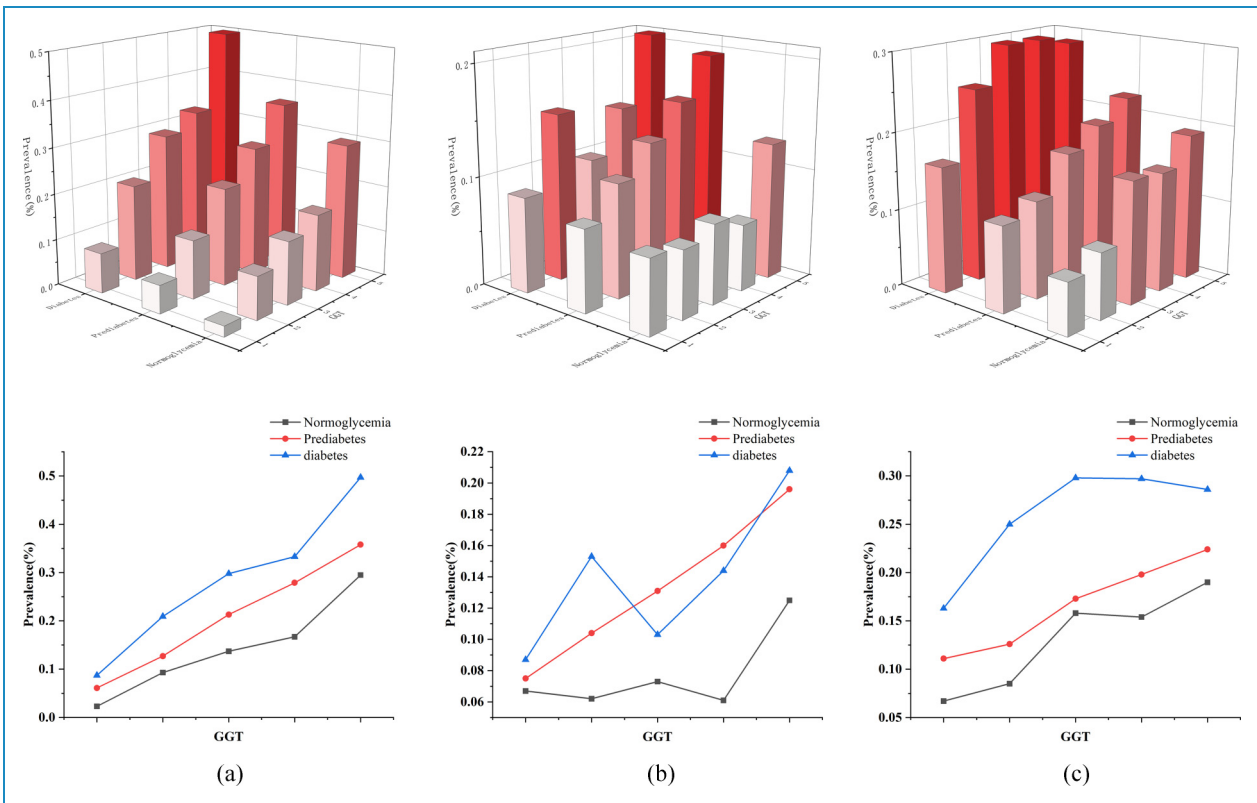


Figure 4. The prevalence of hyperlipidemia under the combined influence of GGT and blood glucose.

modified. It can be seen that for people with HbA1c less than 6.0%, controlling GGT within 30 IU/L reduces the prevalence of hypertriglyceridemia by 4% compared with not controlling GGT (0.85 vs 0.89). For people with HbA1c between 6.0% and 6.5%, the effect of controlling GGT is more obvious. Controlling GGT within 30 IU/L reduces the prevalence of hypertriglyceridemia by 6%, and controlling GGT within 20 IU/L, the prevalence of hypertriglyceridemia decreases by 11%. Although the normal range of GGT is generally accepted to be below 54 IU/L, it can be seen from the

previous analysis that when GGT is greater than 35 IU/L, the prevalence of hypertriglyceridemia has increased significantly, and the control of GGT is often neglected. Therefore, it is meaningful to control GGT for different groups with different blood glucose states.

Discussion

Hyperlipidemia, like diabetes, obesity, hypertension and non-alcoholic fatty liver disease, belongs to the category

Table 3. The correlation analysis results of triglycerides.

Variables	Correlation coefficient	P-value
Sex	-0.038	0.003
Age	0.026	0.045
SBP	0.11	<0.001
DBP	0.14	<0.001
BMI	0.178	<0.001
FPG	0.154	<0.001
2hPG	0.178	<0.001
HbA1c	0.179	<0.001
ALT	0.19	<0.001
AST	0.108	<0.001
GGT	0.276	<0.001

Table 4. Results of multiple regression analysis of triglycerides.

Variables	B	SE	Wald	EXP(B)(95% CI)	P-value
SBP	0.00	0.00	0.17	1.00(1.00-1.00)	0.68
DBP	0.02	0.00	20.17	1.02(1.01-1.02)	<0.001
BMI	0.09	0.01	55.11	1.09(1.07-1.12)	<0.001
FPG	-0.08	0.04	4.74	0.93(0.86-0.99)	0.03
2hPG	0.05	0.01	14.57	1.06(1.03-1.09)	<0.001
HbA1c	0.22	0.06	15.20	1.24(1.11-1.39)	<0.001
ALT	0.05	0.01	49.82	1.05(1.04-1.06)	<0.001
AST	-0.02	0.01	3.62	0.98(0.97-1.00)	0.06
GGT	0.01	0.00	73.98	1.01(1.01-1.02)	<0.001

of metabolic syndrome and is a major risk factor for atherosclerotic cardiovascular disease, whose pathogenesis includes insulin resistance, oxidative stress, endothelial cell damage, inflammation, etc.³⁰⁻³² We observe that BMI, blood glucose, blood pressure, ALT, AST and GGT in the hyperlipidemia group are higher than those in the non-hyperlipidemia group, and the prevalence of diabetes in the hyperlipidemia group is 1.7 times higher than that in the non-hyperlipidemia group, indicating that hyperlipidemia increases the risk of diseases related to metabolic syndromes, such as diabetes, obesity and hypertension.

In this study, the effects of ALT, AST and GGT on the prevalence of hypertriglyceridemia are analyzed. The prevalence of hypertriglyceridemia in the highest decile of ALT and GGT is 4 times and 11 times as in the lowest decile, indicating that GGT is more significant than ALT in the risk of hypertriglyceridemia. In the study, although GGT is at a normal range in most people, BMI, blood glucose and triglyceride gradually increased in the quintile of GGT, which is consistent with Giacomo Zoppini's study.³³ This study also confirms that GGT is significantly correlated with triglyceride level, and GGT has a significant impact on the prevalence of hypertriglyceridemia. GGT is significantly positively correlated with the prevalence of hypertriglyceridemia in both diabetic and non-diabetic populations, and further analysis proves that the effect of GGT on the prevalence of hypertriglyceridemia is independent of blood glucose. More and more evidence shows that the activity of GGT is closely related to metabolic disorders and is an independent predictor of metabolic syndrome and insulin resistance.^{34,35} In this study, GGT levels are related to the levels of metabolic factors such as blood glucose and blood lipids, which further proves the correlation between GGT and metabolic disorders. The underlying mechanism of the relationship between GGT and hypertriglyceridemia is not clear and may be related to the following reasons. First, GGT is a marker of oxidative stress.³⁶ Oxidative stress can occur in metabolic disorders such as hyperglycemia and hyperlipidemia, and the free radicals produced by oxidative stress consume glutathione, which induces the expression of GGT in the liver. Secondly, liver insulin resistance is an important pathogenesis of metabolic syndrome such as hyperlipidemia, diabetes mellitus and nonalcoholic fatty

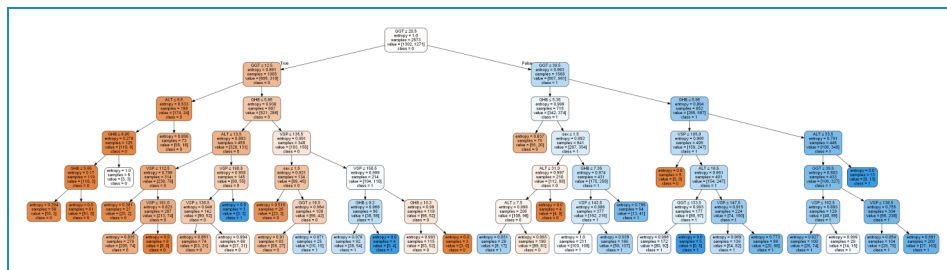


Figure 5. The decision tree. Orange nodules indicate hyperlipidemia and blue nodules indicate non-hyperlipidemia.

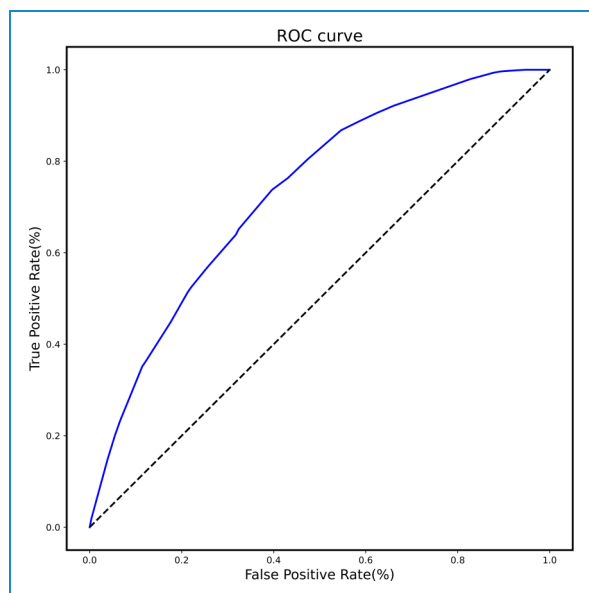


Figure 6. The ROC curve of decision tree classification model.

Table 5. General rules of the non-hypertriglyceridemia population extracted from the decision tree and manually modified.

Rules	Number of positive examples	Number of negative examples	Proportion of positive examples
GGT < 30, HbA1c < 6	2395	287	0.89
GGT < 20, HbA1c < 6	1600	124	0.93
HbA1c < 6	3011	544	0.85
GGT < 20, 6 ≤ HbA1c < 6.5	365	61	0.86
GGT < 30, 6 ≤ HbA1c < 6.5	489	119	0.81
6 ≤ HbA1c < 6.5	744	242	0.75

liver disease. GGT is a liver enzyme index closely related to liver steatosis such as fatty liver, so hyperlipidemia may increase GGT level. More studies are needed to clarify the mechanisms by which GGT levels are associated with an increased prevalence of hypertriglyceridemia.

Based on statistical methods and decision tree, this study find that GGT is an important biochemical indicator closely related to the prevalence of hypertriglyceridemia. Even if

the GGT level is within the normal range, the prevalence of hyperlipidemia will increase gradually with the increase of GGT. Controlling GGT may reduce the prevalence of hypertriglyceridemia, thereby reducing the risk of cardiovascular diseases. Controlling GGT concentration within 30 IU/L in the normoglycemia group and 20 IU/L in the impaired glucose tolerance group can significantly reduce the prevalence of hypertriglyceridemia, suggesting that the incidence of hyperlipidemia can be prevented by controlling GGT level in the sub-health state.

Overall, this study demonstrates that even GGT levels in the normal range increase the risk of hypertriglyceridemia. At present, GGT detection is relatively standard and belongs to the detection scope of liver function, which is economical, simple and feasible. Therefore, GGT detection should be included in routine laboratory testing, and GGT level should be appropriately controlled for different populations to reduce the risk of hyperlipidemia.

Acknowledgements: The authors wish to thank the study participants for their contribution to the research, as well as current and past investigators and staff. The authors would specifically be grateful to the Science Technology Department of Jilin Province for its support.

Contributorship: All authors contributed substantially to the design of the study and the review of the data and the writing of the manuscript.

Declaration of conflicting interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: The Institutional Ethics Committee at the First Hospital of Jilin University approved the study protocol. Written informed consent was obtained from each participant before data collection.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study is supported in part by the Science and Technology Development Plan Project of Jilin Province (grant number 20210101181JC).

Guarantor: Hongtao Bai.

ORCID ID: Tingting Zhang  <https://orcid.org/0000-0002-0193-058X>

References

- Kopin L and Lowenstein C. In the Clinic® dyslipidemia. *Ann Intern Med* 2017; 167: ITC81–ITC95.
- Ye X, Kong W, Zafar MI, et al. Serum triglyceride as a risk factor for cardiovascular diseases in type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies.

- Cardiovasc Diabetol* 2019; 18. BioMed Central Ltd. <https://doi.org/10.1186/s12933-019-0851-z>
- Zheng D, Dou J, Liu G, et al. Association between triglyceride level and glycemic control among insulin-treated patients with type 2 diabetes. *J Clin Endocrinol Metab* 2019; 104: 1211–1220.
 - Vekic J, Zeljkovic A, Stefanovic A, et al. Obesity and dyslipidemia. *Metabolism* 2019; 92: 71–81. W.B. Saunders.
 - Johnson CB, Davis MK, Law A, et al. Shared risk factors for cardiovascular disease and cancer: implications for preventive health and clinical care in oncology patients. *Can J Cardiol* 2016; 32: 900–907. Pulsus Group Inc.
 - Séguro F, Bongard V, Bérard E, et al. Dutch lipid clinic network low-density lipoprotein cholesterol criteria are associated with long-term mortality in the general population. *Arch Cardiovasc Dis* 2015; 108: 511–518.
 - Ferrières J. Hypercholesterolaemia and coronary artery disease: A silent killer with several faces. *Arch Cardiovasc Dis* 2019; 112: 75–78.
 - Law MR, Wald NJ, Thompson SG, et al. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? *Br Med J* 1994; 308: 367–372.
 - Su P, Hong L, Zhao Y, et al. Relationship between hyperuricemia and cardiovascular disease risk factors in a Chinese population: a cross-sectional study. *Med Sci Monit* 2015; 21: 2707–2717.
 - Bozzetto L, della Pepa G, Vetrani C, et al. Dietary impact on postprandial lipemia. *Front Endocrinol* 2020; 11, Frontiers Media S.A. <https://doi.org/10.3389/fendo.2020.00337>.
 - Mozaffarian D, Benjamin EJ, Go A, et al. Heart disease and stroke statistics-2016 update: a report from the American heart association. *Circulation* 2016; 133: e38–e48. Lippincott Williams and Wilkins.
 - Karr S. Epidemiology and management of hyperlipidemia. *Am J Manag Care*. 2017; 23(9 Suppl): S139–S148.
 - Bjornstad P and Eckel RH. Pathogenesis of lipid disorders in insulin resistance: a brief review. *Curr Diab Rep* 18
 - Bahijri SM, Ahmed M, Al-Shali K, et al. The relationship of management modality in Saudi patients with type 2 diabetes to components of metabolic syndrome, γ glutamyl transferase and highly sensitive C-reactive protein. *Therapeutic Adv Chronic Dis* 7(5): 246–254. <https://doi.org/10.1177/2040622316658459>.
 - Turgut O and Tandogan I. Gamma-glutamyltransferase to determine cardiovascular risk: shifting the paradigm forward. *J Atheroscler Thromb* 2011; 18: 177–181.
 - Lee DH, Blomhoff R and Jacobs DR. Is serum gamma glutamyltransferase a maker of oxidative stress? *Free Radical Res* 38: 535–539. <https://doi.org/10.1080/107157604100>.
 - Kalirawna TR, Rohilla J, Bairwa SS, et al. Increased concentration of serum gamma-glutamyl transferase in ischemic stroke patients. *Brain Circ* 7: 71–76. https://doi.org/10.4103/bc.bc_47_20.
 - Park EO, Bae EJ, Park BH, et al. The associations between liver enzymes and cardiovascular risk factors in adults with mild dyslipidemia. *J Clin Med* 9. <https://doi.org/10.3390/jcm9041147>.
 - Ndrepepa G, Collieran R and Kastrati A. Gamma-glutamyl transferase and the risk of atherosclerosis and coronary heart disease. *Clin Chim Acta* 476: 130–138. Elsevier B.V. <https://doi.org/10.1016/j.cca.2017.11.026>.
 - Morling JR, Fallowfield JA, Williamson RM, et al. γ -Glutamyltransferase, but not markers of hepatic fibrosis, is associated with cardiovascular disease in older people with type 2 diabetes mellitus: the Edinburgh type 2 diabetes study. *Diabetologia* 58: 1484–1493. <https://doi.org/10.1007/s00125-015-3575-y>.
 - Lee D-H, Jacobs DR, Gross M, et al. Glutamyltransferase Is a Predictor of Incident Diabetes and Hypertension: The Coronary Artery Risk Development in Young Adults (CARDIA) Study (2.7-11.1) for 15-year incident diabetes and 1.0, 1.2, 1.7 (1.2-2.2), 2.3 (1.7-3.2), and 2.3 (1.7-3.2) for hypertension. Additional adjustment for year 0 alcohol consumption, body mass, <https://academic.oup.com/clinchem/article/49/8/1358/5641890>
 - Yoon HE, Mo EY, Shin SJ, et al. Serum gamma-glutamyltransferase is not associated with subclinical atherosclerosis in patients with type 2 diabetes. *Cardiovasc Diabetol* 15. <https://doi.org/10.1186/s12933-016-0426-1>.
 - Kotani K and Sakane N. Association between remnant-like particle cholesterol and gamma-glutamyltransferase in dyslipidemic patients. *Med Princ Pract* 2013; 22: 189–193.
 - Tang L, Yuan B, Zhang F, et al. Visceral fat is associated with elevation of serum alanine aminotransferase and gamma glutamyltransferase in middle-aged Chinese adults. *Postgrad Med J* 94: 641–646. <https://doi.org/10.1136/postgradmedj-2018-135644>.
 - Huang Y, Luo J, Liu X, et al. Gamma-glutamyltransferase and risk of acute coronary syndrome in young Chinese patients: a case-control study. *Dis Markers* 2018. <https://doi.org/10.1155/2018/2429160>
 - Liu FT, Ting KM and Zhou ZH. Isolation forest. In: 2008 8th IEEE international conference on data mining, Pisa, 15–19 December 2008, pp. 413–422. <https://doi.org/10.1109/ICDM.2008.17>
 - Hunt EB, Marin J and Stone PJ. *Experiments in induction*. New York and London: Academic Press, 1966.
 - Salzberg SL. C4.5: programs for machine learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Mach Learn* 1994; 16: 235–240.
 - Nishina K, Li W, Sato M, et al. Serum c-glutamyltransferase and development of impaired fasting glucose or type 2 diabetes in middle-aged Japanese men. *J Intern Med*. 2003; 254: 287–295.
 - Kassi E, Pervanidou P, Kaltsas G, et al. Metabolic syndrome: definitions and controversies. *BMC Med* 2011; 9: 48.
 - Huang PL. A comprehensive definition for metabolic syndrome. *Dis Model Mech* 2009; 2: 231–237.
 - Dandona P, Aljada A, Chaudhuri A, et al. Metabolic syndrome: a comprehensive perspective based on interactions between obesity, diabetes, and inflammation. *Circulation* 2005; 111: 1448–1454.
 - Zoppini G, Targher G, Trombetta M, et al. Relationship of serum γ -glutamyltransferase to atherogenic dyslipidemia and glycemic control in type 2 diabetes. *Obesity* 2009; 17: 370–374.
 - Lonardo A, Lombardini S, Scaglioni F, et al. Hepatic steatosis and insulin resistance: does etiology make a difference? *J Hepatol* 2006; 44: 190–196.

35. Rantala AO, Lilja M, Kauma H, et al. Gamma-glutamyl transpeptidase and the metabolic syndrome. *J Intern Med* 2000; 248: 230–238.
36. Lee DH, Blomhoff R and Jacobs DR. Is serum gamma glutamyltransferase a marker of oxidative stress? *Free Radical Res* 2004; 38: 535–539.
-