

# Non-Linear Relationship Between High-Density Lipoprotein Cholesterol and Incident of Diabetes Mellitus: A Secondary Retrospective Analysis Based on A Japanese Cohort Study

**Changchun Cao**

Shenzhen Dapeng New District Nan'ao People's Hospital

**Haofei Hu** (✉ [huhaofei0319@126.com](mailto:huhaofei0319@126.com))

Shenzhen Second People's Hospital <https://orcid.org/0000-0001-6061-6796>

**Xiaodan Zheng**

Shenzhen Second People's Hospital

**xinyu Wang**

Shenzhen Second People's Hospital

**Yulong Wang**

Shenzhen Dapeng New District Nao'ao People's Hospital

**Yongcheng He**

Shenzhen Hengsheng Hospital

---

## Research

**Keywords:** High-density lipoprotein cholesterol, incident of diabetes mellitus, nonlinearity

**Posted Date:** January 20th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-146864/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Background and Objective** High-density lipoprotein cholesterol (HDL-C) may be directly involved in glucose metabolism by enhancing insulin sensitivity and insulin secretion. This current study aimed to explore the association between HDL-C and the risk of diabetes mellitus (DM) in Japanese population.

**Methods** This retrospective cohort study was based on a publicly available DRYAD dataset. A total of 15388 study participants from Murakami Memorial Hospital in Japan included all medical records for participants who received medical examinations from 2004 to 2015. The target-independent variable and the dependent variable were HDL-C measured at baseline, and DM incident appeared during follow-up, respectively. Covariates involved in this study included gender, age, ethanol consumption, smoking status, regular exerciser, systolic blood pressures, diastolic blood pressures, body mass index, waist circumference, fatty liver, total cholesterol, triglycerides, hemoglobin A1c, fasting plasma glucose. Cox proportional-hazards regression was used to investigate the association between HDL-C and DM, generalized additive models to identify non-linear relationships.

**Results** After adjusting for gender, age, ethanol consumption, smoking status, regular exerciser, systolic blood pressures, diastolic blood pressures, body mass index, waist circumference, total cholesterol, triglycerides, hemoglobin A1c, fasting plasma glucose, the result showed HDL-C was negatively associated with incident of DM (HR = 0.54, 95%CI (0.35, 0.82)). The results remained robust in a series of sensitive analysis. A non-linear relationship was detected between HDL-C and incident of DM, which had an inflection point of HDL-C was 1.72mmol/L. The effect sizes and the confidence intervals on the left and right sides of the inflection point were 0.36 (0.21, 0.59) and 2.90 (0.96, 8.75). Subgroup analysis showed a stronger association could be found in the population with the ex-smoker and current-smoker. The same trend was also seen in the community with hypertension (P for interaction = 0.010, HR = 1.324).

**Conclusion:** HDL-C is negatively associated with DM risk. The relationship between HDL-C and incident of DM is also non-linear. HDL-C was strongly negatively related to the incident of diabetes when HDL-C is less than 1.72mmol/L.

## Introduction

Diabetes has become one of the most critical public health problems in the world. The International Diabetes Federation reported that the global prevalence of age-standardized diabetes, estimated at 9.3% in 2019, will rise to 10.2% by 2030 and 10.9% by 2045 [1]. As one of the most common chronic diseases, diabetes has brought a substantial economic burden to patients and the country[2]. Diabetes, a metabolic disorder, is characterized by chronic and long-term hyperglycemia caused by genetic and environmental factors[3]. However, the pathogenesis of diabetes is unclear. Therefore, it is imperative to clarify the risk factors of diabetes and carry out early screening, intervention, and prevention and treatment of high-risk groups.

Diabetes mellitus (DM) and prediabetic states are often accompanied by abnormal blood lipid metabolism, characterized by decreased serum high-density lipoprotein cholesterol (HDL-C) and increased triglycerides (TG)[4–7]. Dyslipidemia associated with diabetes is the primary factor contributing to increased cardiovascular risk[8]. HDL has been recognized as a cardiovascular protective factor[9]. Some studies showed that higher HDL-C levels are related to a lower risk of DM[10]. Evidence suggests that HDL-C may be directly involved in glucose metabolism by enhancing insulin sensitivity and insulin secretion[11]. However, some studies have demonstrated elevated blood HDL levels to increase the risk of diabetes[12, 13]. Some clinical trials also have shown that some drugs that increase HDL-C levels, such as cholesteryl ester transfer protein (CETP) inhibitors and niacin, cannot decrease the risk of cardiovascular and cerebrovascular diseases[14]. Therefore, this current study aimed to explore the association between HDL-C and the risk of incident DM in Japanese population through a public database from Japan.

Our study conducted a secondary analysis according to a previous public database from a longitudinal study that showed ectopic obesity increased the risk of diabetes. The author explored the relationship between Ectopic fat obesity and the risk of incident diabetes in original research[15]. In our study, we chose the baseline HDL-C level as the independent variable and incident DM as the dependent variable. The other covariates were the same as the original research.

## Research Objects And Methods

### Data source

This study conducted a secondary analysis from the DRYAD database([www.Datadryad.org](http://www.Datadryad.org)). We have free access to the raw data from this site and analyzed the data. (Dryad data package: Okamura T, Hashimoto Y, Hamaguchi M, Obora A, Kojima T, Fukui M (2019) Data from Ectopic fat obesity presents the greatest risk for incident diabetes: a population-based longitudinal study. Dryad Digital Repository.<https://datadryad.org/stash/dataset/doi:10.5061/dryad.8q0p192>) [15]. Variables contained in the public database were as follows: age, gender, body mass index (BMI), waist circumference (WC), ethanol consumption, smoking status, exercise habit, fatty liver, and baseline levels of systolic blood pressure (SBP), diastolic blood pressure (DBP), alanine aminotransferase(ALT), aspartate aminotransferase(AST), total cholesterol(TC), TG, HDL-C, fasting blood glucose (FPG), glycosylated hemoglobin (HbA1c), day of follow-up and incident diabetes at follow up. The authors of the original research waive all copyrights to these data. Therefore, we can conduct a secondary analysis by using these databases without compromising the authors' rights.

### Study participants

We obtained data from a database which was provided by the Murakami Memorial Hospital in Japan. This database contained 20,944 participants who received medical examinations from 2004 to 2015. The exclusion criteria of the original data were as follows: (1) alcoholic fatty liver disease, (2) viral hepatitis (detection of hepatitis B antigen and hepatitis C antibody at baseline), (3) using any medication at baseline, (4) diabetes at baseline, (5) missing data of covariates, (6) FPG  $\geq$  6.1 mmol/L. Okamura T et al.

selected 15,464 participants finally. The original retrospective cohort study described inclusion/exclusion criteria and trial outcome measures[15]. The ethics committee approved the original research of Murakami Memorial Hospital and informed consent was obtained from all participants. For further analysis, 126 participants with outliers of HDL-C, including HDL-C less than the mean minus three standard deviations (SD) or greater than the mean plus three SD, were excluded.[18]. Eventually, 15338 subjects (8397 male and 6941 female) were included in our study's data analysis.

## **Study design and measurement of variables**

The original study obtained medical history and lifestyle factors of all participants through a standardized self-management questionnaire. The trained staff conducted the clinical measurements, including body weight, waist circumference height, blood pressure. Laboratory inspection results are collected under standardized conditions and conducted in accordance with uniform procedures. The original study defined regular exercise as playing any type of exercise  $> 1 \times / \text{week}$ [16]. The original study assessed ethanol consumption by asking the participants about their alcohol consumption in the previous month, then estimating the mean ethanol intake per week[15]. The original study defined visceral fat obesity as a waist circumference  $\geq 90$  cm in men or  $\geq 80$  cm in women[15]. The target independent variable is HDL-C obtained at baseline. The dependent variable is incident diabetes obtained in the follow up. As a retrospective cohort study, our study reduced the possibility of selection bias and observation bias.

## **Diagnosis of incident diabetes**

DM was defined as participants whose HbA1c was not lower than 6.5%, FPG was not lower than 7 mmol/l[17] or self-reported during the follow-up period.

## **Statistical analysis**

We used EmpowerStats software ([www.empowerstats.com](http://www.empowerstats.com), X&Y solutions, Inc., Boston, MA, USA) and R (<http://www.R-project.org>) for data analysis.

Participants were divided into four groups according to the quartile of HDL-C. Continuous variables were presented as mean  $\pm$  standard deviation when the data obeyed normal distribution and expressed as the median or interquartile range when the data obeyed skewed distribution. We used frequencies and percentages to present categorical variables. We also analyzed differences among groups by performing the Kruskal-Wallis H test, the one-way ANOVA and the chi-square test. Univariate and multivariate Cox regression analysis were applied to evaluate the association between HDL-C and diabetes risk. The Cox regression model was used to calculate the hazard ratio (HR) and 95% confidence interval (95% CI) of diabetes caused by HDL-C. According to the STROBE statement's recommendation[19], we used three models to describe HDL-C and DM's relationship. The crude model included only HDL-C as the independent variable. We adjusted for gender, age, smoking, ethanol consumption, regular exerciser, SBP, DBP, BMI and WC in the model I. Model II included all the variables in the model I and further adjusted for TC, TG, HbA1c, FPG. The confounding covariates with original effect size (hazard ratio) change  $> 10\%$  or P-value of regression coefficient  $< 0.1$  can be added to the model as confounding factors[20]. In order to ensure the robustness of the result, we conducted a series of sensitivity analysis. We transformed HDL-C into

categorical variables based on quartiles and calculated the P-value for trend. The aim was to evaluate a potential non-linear relationship of HDL-C with the risk of DM. We performed a weighted generalized additive model (GAM) model to adjust for the covariates in Model III, because the generalized linear model has limitations in dealing with nonlinearities. We used a GAM to evaluate the non-linear relationship between HDL-C levels and the incidence of diabetes, and further used a two-piece linear regression model to find the inflection point. The most appropriate model for describing the relationship between HDL-C and DM risk was determined by using log likelihood ratio test. The Cox proportional hazard model was used to conduct subgroup (age, gender, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, fatty liver) analysis to further verify the robustness of the results. In addition, a likelihood ratio test was performed to evaluate the interaction between subgroups. We used Kaplan-Meier method by using the time-to-first event for each endpoint to compare survival estimates and cumulative event rates. Kaplan-Meier HRs for adverse events and their corresponding 95% CIs were compared by the log-rank test. A two-sided  $P < 0.05$  was considered significant.

## Results

Our study included 15388 participants, of whom 54.7% were men and 45.3% were women. The mean  $\pm$  SD age of the participants was  $43.7 \pm 8.9$  years old. After a maximum of 4732 days of follow-up (median 1967 days), 373 people finally developed diabetes. The mean  $\pm$  SD HDL-C was  $1.5 \pm 0.4$  mmol/L. The mean  $\pm$  SD FPG, BMI, and WC were  $5.2 \pm 0.4$  mmol/L,  $22.1 \pm 3.1$  kg/m<sup>2</sup>, and  $76.5 \pm 9.1$  cm.

## Baseline demographic and clinical characteristics

Table 1 showed the baseline demographic and clinical characteristics of participants. Participants were divided into subgroups according to HDL-C quartiles ( $\leq 1.164$ ,  $1.164-1.407$ ,  $1.407-1.696$ ,  $> 1.696$ ). In the lowest HDL-C group, participants commonly had higher age, BMI, WC, SBP, DBP, FPG, HbA1c, TG, ethanol consumption, higher proportion of men, fatty liver and current smoker. In the highest HDL-C group, we found that participants generally had higher TC and regular exerciser.

Table 1  
The Baseline Characteristics of participants

HDL-C(mmol/L)	Q1(≤ 1.16)	Q2(1.16 to ≤ 1.41)	Q3(1.41 to ≤ 1.70)	Q4 > 1.70	P-value
Participants	3752	3895	3845	3846	
Gender					< 0.001
Women	568 (15.14%)	1419 (36.43%)	2122 (55.19%)	2832 (73.63%)	
Men	3184 (84.86%)	2476 (63.57%)	1723 (44.81%)	1014 (26.37%)	
Age(years)	44.49 ± 9.03	43.63 ± 8.96	43.31 ± 8.77	43.36 ± 8.78	< 0.001
Ethanol consumption(g/week)	4.20(0, 66.00)	1.00(0, 72.00)	1.00(0, 60.00)	1.00(0, 54.00)	< 0.001
Smoking status					< 0.001
Never-smoker	1436 (38.27%)	2114 (54.27%)	2484 (64.60%)	2906 (75.56%)	
Ex-smoker	863 (23.00%)	821 (21.08%)	694 (18.05%)	550 (14.30%)	
Current-smoker	1453 (38.73%)	960 (24.65%)	667 (17.35%)	390 (10.14%)	
Regular exerciser					< 0.001
No	3190 (85.02%)	3226 (82.82%)	3156 (82.08%)	3085 (80.21%)	
Yes	562 (14.98%)	669 (17.18%)	689 (17.92%)	761 (19.79%)	
SBP (mmHg)	118.88 ± 14.75	115.96 ± 14.97	112.67 ± 14.53	110.63 ± 14.31	< 0.001
DBP (mmHg)	75.00 ± 10.21	72.62 ± 10.41	70.29 ± 10.23	68.55 ± 10.03	< 0.001
BMI (kg/m <sup>2</sup> )	23.95 ± 3.11	22.61 ± 3.08	21.50 ± 2.71	20.52 ± 2.47	< 0.001
WC (cm)	82.37 ± 8.30	78.17 ± 8.52	74.47 ± 8.13	71.20 ± 7.40	< 0.001

Values are n(%) or mean ± SD

SBP systolic blood pressures, DBP diastolic blood pressures, BMI body mass index, WC waist circumference, ALT alanine aminotransferase, AST aspartate aminotransferase, HDL-C high-density lipoprotein cholesterol, TC total cholesterol, TG triglycerides, HbA1c hemoglobin A1c, FPG fasting plasma glucose

HDL-C(mmol/L)	Q1(≤ 1.16)	Q2(1.16 to ≤ 1.41)	Q3(1.41 to ≤ 1.70)	Q4 > 1.70	P-value
Fatty liver					< 0.001
No	2339 (62.34%)	3081 (79.10%)	3491 (90.79%)	3693 (96.02%)	
Yes	1413 (37.66%)	814 (20.90%)	354 (9.21%)	153 (3.98%)	
ALT (IU/L)	21(16.00,30.00)	17 (13.00,24.00)	15 (12.00,20.00)	14(11.00,19.00)	< 0.001
AST (IU/L)	18 (15.00,22.00)	17 (14.00,21.00)	17 (14.00,20.00)	17 (13.00,20.00)	< 0.001
HDL-C (mmol/L)	1.00 ± 0.12	1.29 ± 0.07	1.54 ± 0.08	1.97 ± 0.22	< 0.001
TG (mmol/L)	1.20 (0.83,1.73)	0.82 (0.56,1.17)	0.64 (0.45,0.88)	0.53 (0.38,0.72)	< 0.001
TC (mmol/L)	5.07 ± 0.89	5.07 ± 0.89	5.07 ± 0.84	5.27 ± 0.80	< 0.001
HbA1c (%)	5.19 ± 0.34	5.18 ± 0.33	5.15 ± 0.31	5.17 ± 0.30	< 0.001
FPG (mmol/L)	5.30 ± 0.38	5.20 ± 0.40	5.11 ± 0.41	5.04 ± 0.41	< 0.001
Values are n(%) or mean ± SD					
SBP systolic blood pressures, DBP diastolic blood pressures, BMI body mass index, WC waist circumference, ALT alanine aminotransferase, AST aspartate aminotransferase, HDL-C high-density lipoprotein cholesterol, TC total cholesterol, TG triglycerides, HbA1c hemoglobin A1c, FPG fasting plasma glucose					

## Univariate analysis

Table 2 showed the univariate analysis between HDL-C and incident of DM. The univariate analysis results demonstrated that age, ethanol consumption, smoking status, SBP, DBP, BMI, WC, fatty liver, TG, TC, HbA1c, and FPG were positively related to the incident of diabetes. It showed that HDL-C was negatively correlated with diabetes incident in univariate analysis. Compared with female participants, male participants were more likely to develop DM.

Table 2  
The results of univariate analysis

	Statistics	HR (95% CI)	P value
Gender			
Women	6941 (45.25%)	ref	
Men	8397 (54.75%)	2.51 (1.98, 3.20)	< 0.0001
Age(years)	43.69 ± 8.89	1.06 (1.04, 1.07)	< 0.0001
Ethanol consumption(g/week)	47.67 ± 82.24	1.00 (1.00, 1.00)	0.0011
Smoking status			
Never-smoker	8940 (58.29%)	ref	
Ex-smoker	2928 (19.09%)	1.65 (1.25, 2.18)	0.0004
Current-smoker	3470 (22.62%)	2.58 (2.05, 3.24)	< 0.0001
Regular exerciser			
No	12657 (82.52%)	ref	
Yes	2681 (17.48%)	0.76 (0.57, 1.02)	0.0686
SBP(mmHg)	114.51 ± 14.97	1.03 (1.03, 1.04)	< 0.0001
DBP(mmHg)	71.60 ± 10.50	1.05 (1.04, 1.06)	< 0.0001
BMI(kg/m <sup>2</sup> )	22.13 ± 3.13	1.24 (1.22, 1.27)	< 0.0001
WC(cm)	76.52 ± 9.10	1.09 (1.08, 1.10)	< 0.0001
Fatty liver			
No	12604 (82.17%)	ref	
Yes	2734 (17.83%)	6.99 (5.69, 8.60)	< 0.0001
ALT (IU/L)	20.00 ± 14.38	1.01 (1.01, 1.01)	< 0.0001
AST (IU/L)	18.39 ± 8.59	1.01 (1.01, 1.01)	< 0.0001
HDL-C (mmol/L)	1.45 ± 0.38	0.13 (0.09, 0.18)	< 0.0001
TG (mmol/L)	0.92 ± 0.66	1.80 (1.68, 1.92)	< 0.0001
TC (mmol/L)	5.12 ± 0.86	1.50 (1.35, 1.67)	< 0.0001
HbA1c (%)	5.17 ± 0.32	55.87 (40.56, 76.97)	< 0.0001
FPG (mmol/L)	5.16 ± 0.41	25.46 (18.77, 34.55)	< 0.0001

We stratified the Kaplan-Meier curves of the cumulative hazards of diabetes incident risk by HDL-C quartiles in Fig. 1. The diabetes incident risk was significantly different among the four HDL-C groups ( $p < 0.0001$ ). The cumulative diabetes incident risk gradually increased with decreased HDL-C. Therefore, the lowest HDL-C group has the greatest risk of diabetic events.

## The multivariate analysis of HDL-C with DM risk

In Table 3, multivariate analysis was applied to assess the relationships between HDL-C and DM risk, including crude and adjusted models. There was a negative correlation between HDL-C and DM risk in the unadjusted model (HR = 0.13, 95% CI: 0.09 to 0.18,  $P < 0.0001$ ). In model I (adjusted gender, age, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC), the result still existed (HR: 0.34, 95% CI: 0.23 to 0.50). After adjusting for gender, age, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG in the model II, the connection could also be detected (HR = 0.54, 95%CI: 0.35 to 0.82,  $P = 0.0044$ ). When converting HDL-C into a categorical variable, the relationship between HDL-C and diabetes in the highest quartile is not statistically significant, compared to the lowest quartile in model II (HR = 0.67, 95%CI: 0.43 to 1.05,  $P = 0.0791$ ). However, the trend across the quartiles was significant in model II ( $P = 0.0056$ ). The p trend became non-equidistant when HDL-C was converted to a categorical variable, indicating a non-linear relationship between HDL-C and the DM risk (Table 3). Then GAM was performed to insert the continuity covariate into the equation as a curve. It generally remained consistent with the GAM (HR: 0.47; 95% CI:0.30 to 0.75,  $P = 0.0015$ ), which demonstrated the robustness of the results (Table 3).

**Table3** Relationship between HDL-C and the incident of diabetes in different models

Variable	Crude model (H.R.,95% CI, P)	Model I (HR,95% CI, P)	Model II (HR,95% CI, P)	GAM (HR,95% CI, P)
HDL-C	0.13 (0.09, 0.18) <0.0001	0.34 (0.23, 0.50) <0.0001	0.54 (0.35, 0.82) 0.0044	0.47 (0.30, 0.75) 0.0015
HDL-C(quartile)				
Q1	ref	ref	ref	ref
Q2	0.49 (0.39, 0.63) <0.0001	0.68 (0.53, 0.87) 0.0023	0.83 (0.64, 1.07) 0.1575	0.82 (0.63, 1.06) 0.1343
Q3	0.24 (0.17, 0.33) <0.0001	0.44 (0.31, 0.63) <0.0001	0.58 (0.41, 0.83) 0.0031	0.54 (0.37, 0.79) 0.0014
Q4	0.18 (0.12, 0.27) <0.0001	0.45 (0.30, 0.70) 0.0003	0.67 (0.43, 1.05) 0.0791	0.58 (0.36, 0.94) 0.0266
P for trend	<0.0001	<0.0001	0.0056	0.0018

Crude model: we did not adjust for other covariants.

Model I: we adjusted for gender, age, ethanol consumption, smoking, regular exerciser, SBP, DBP, BMI, WC.

Model II: we adjusted for gender, age, ethanol consumption, smoking, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG.

GAM: All covariates listed in Table 1 were adjusted. However, continuous covariates were adjusted as nonlinearity.

CI Confidence interval, Ref Reference

## The analyses of non-linear relationship between HDL-C and DM risk

GAM was applied to assess whether there was a non-linear relationship between HDL-C and DM risk in our study (Fig. 2). A non-linear relationship between HDL-C and DM risk was discovered after adjusting for gender, age, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c and FPG (Log-likelihood ratio test  $P = 0.005$ ). The inflection point of HDL-C was detected to be 1.72 mmol/L by employing a two-piecewise linear regression model. When  $\text{HDL-C} \leq 1.72$  mmol/L, we observed a negative association between HDL-C and incident of diabetes (HR: 0.36, 95%CI: 0.21 to 0.59,  $P < 0.0001$ ). However, when  $\text{HDL-C} > 1.72$  mmol/L, their association tended to be saturated (HR: 2.90, 95%CI: 0.96 to 8.75,  $P = 0.0594$ ) (Table 4).

Table 4 The result of two-piecewise linear regression model

	Incident of diabetes	(HR,95%CI, P)
Fitting model by standard linear regression	0.54 (0.35, 0.82)	0.0044
Fitting model by two-piecewise linear regression		
Inflection point of HDL-C	1.72	
$\leq 1.72$	0.36 (0.21, 0.59)	<0.0001
$> 1.72$	2.90 (0.96, 8.75)	0.0594
P for log likelihood ratio test	0.005	

We adjusted for gender, age, ethanol consumption, smoking, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG

CI Confidence interval

## The results of subgroup analyses

According to the STROBE statement[19], subgroup analysis may reveal the underlying association between HDL-C and DM risk. As shown in Table 5, we chose age, gender, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, fatty liver as the stratification variables and further detected the trend of effect sizes in these stratification variables. Based on our a priori specification, we discovered most stratification variables did not affect the relationship between HDL-C and DM risk, except smoking status, SBP and DBP ( $P < 0.05$ ). A stronger association was observed in the participants with ex-smoker, current-smoker and hypertension( $SBP \geq 140$  mmHg or  $DBP \geq 90$  mmHg) in our study. In contrast, there was a weaker association in the people with never-smoker,  $SBP < 140$  mmHg, and  $DBP < 90$  mmHg.

Table 5  
Effect size of HDL-C on diabetes in prespecified and exploratory subgroups

Characteristic	No of patients	Effect size(95%CI)	P value	P for interaction
Age(years)				0.3924
< 60	14633	0.64 (0.41, 1.01)	0.0543	
≥ 60	705	0.33 (0.07, 1.49)	0.1497	
Gender				0.2075
Women	6941	0.84 (0.39, 1.81)	0.6515	
Men	8397	0.46 (0.27, 0.78)	0.0036	
Ethanol consumption (g/week)				0.4092
0	11718	0.70 (0.43, 1.14)	0.1508	
> 0	3620	0.51 (0.27, 0.98)	0.0442	
Smoking status				0.0062
Never-smoker	8940	1.28 (0.68, 2.42)	0.4420	
Ex-smoker	2928	0.29 (0.10, 0.83)	0.0212	
Current-smoker	3470	0.32 (0.16, 0.66)	0.0020	
Regular exerciser				0.9793
No	12657	0.54 (0.34, 0.84)	0.0070	
Yes	2681	0.54 (0.22, 1.34)	0.1860	
SBP(mmHg)				0.0160
< 140	14556	0.76 (0.48, 1.19)	0.2275	
≥ 140	782	0.14 (0.03, 0.56)	0.0055	
DBP(mmHg)				0.0055
< 90	14578	0.78 (0.50, 1.22)	0.2694	
≥ 90	760	0.09 (0.02, 0.42)	0.0023	
BMI(kg/m <sup>2</sup> )				0.2004
< 25	12817	0.71 (0.42, 1.20)	0.1952	
≥ 25	2521	0.40 (0.19, 0.82)	0.0131	

Note 1: Above model adjusted for gender, age, ethanol consumption, smoking, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG

Characteristic	No of patients	Effect size(95%CI)	P value	P for interaction
Visceral fat obesity				0.9829
No	13332	0.60 (0.37, 0.98)	0.0399	
Yes	2006	0.59 (0.29, 1.23)	0.1618	
Fatty liver				0.7475
No	12604	0.67 (0.39, 1.16)	0.1561	
Yes	2734	0.59 (0.32, 1.11)	0.1025	
Note 1: Above model adjusted for gender, age, ethanol consumption, smoking, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG				

## Discussion

In this retrospective cohort study, we found that low HDL-C was independently related to DM risk after adjusting for gender, age, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1c, FPG. Further analysis showed a non-linear relationship between HDL-C level and DM risk (Log-likelihood ratio test  $P = 0.005$ ). This result suggested that the HDL-C level was negatively associated with DM risk when the HDL-C level was  $\leq 1.72$ mmol/L (HR: 0.36, 95%CI: 0.21 to 0.59,  $P < 0.0001$ ). However, their relationship tended to be saturated when HDL-C was  $\geq 1.72$ mmol/L. (HR: 2.90, 95%CI: 0.96 to 8.75,  $P = 0.0594$ ). It could be better understood the trend of HDL-C and diabetes incidence in different populations by analysis of sub-groups. A stronger association between HDL-C and DM risk was discovered in the participants with ex-smoker, current-smoker and hypertension (SBP  $\geq 140$ mmHg or DBP  $\geq 90$ mmHg). In contrast, there was a weaker association in the people with never-smoker, SBP  $< 140$ mmHg, and DBP  $< 90$ mmHg.

Cardiovascular disease is a significant cause of morbidity and mortality in type 2 diabetes mellitus patients[8], and HDL-C plays an essential role in cardiovascular disease and diabetes. It was reported that low mean and high variability of HDL-C were independent predictors of diabetes with an additive effect[21]. Our study found that low HDL-C is an independent risk of DM, which was consistent with previous studies[22]. HDL-C affected the incident of diabetes through different mechanisms. Studies found that HDL also has anti-inflammatory and antioxidant activities[23]. HDL-C can inhibit apoptosis of islet  $\beta$ cell induced by oxidative stress [11]. Besides, HDL-C improves cell sensitivity to insulin and glucose uptake in skeletal muscle cells by activating Adenosine 5'-monophosphate (AMP)-activated protein kinase (AMPK) pathway[24]. Moreover, HDL-C can promote insulin secretion by increasing the outflow of cholesterol from pancreatic B cells[25]. These mechanisms can provide a reasonable explanation for reducing HDL-C and increasing the DM risk.

In the past, it was generally believed that higher HDL-C was more beneficial. However, lin et al. [12] found that high serum HDL-C levels increase the risk of DM after adjusting for the demographic and clinical covariates in a retrospective study of 9764 Chinese. Also, Chen et al.[13] found high serum HDL-C levels

increase the risk of DM after adjusting for potential confounders in a large retrospective cohort study. A population-based study in 2016 showed that plasma HDL-C levels were significantly increased by carriers of scavenger receptor BI P376L mutations have HDL-C levels, but the risk of coronary heart disease was increased[26]. Animal studies showed that SR-BI knockout mice increased HDL-C levels considerably, but the probability of atherosclerosis also increased[27]. A population-based study in 2016 showed that carriers of scavenger receptor BI P376L mutations have significantly increased plasma HDL-C levels, but the risk of coronary heart disease is increased[27]. Animal studies have also shown that SR-BI knockout mice have increased HDL-C levels considerably, but the probability of atherosclerosis has increased[28]. To clarify the association between HDL-C and the risk of diabetes, we did a smooth curve fitting. The results showed that when  $\text{HDL-C} \leq 1.72 \text{ mmol/L}$ , the risk of diabetes is inversely proportional to HDL-C. When  $\text{HDL-C} > 1.72 \text{ mmol/L}$ , the risk of diabetes is directly proportional to HDL-C, but it is not statistically significant (HR: 2.90, 95%CI: 0.96 to 8.75,  $P = 0.0594$ ). It suggests that HDL-C is not as high as possible in the occurrence of diabetes.

The current study has several following strengths. (1) We processed and further explored the non-linear relationship between HDL-C and diabetes in the present study; (2) Strict statistical adjustments were used to minimize residual confounding factors; (3) In order to decrease the contingency of data analysis and increase the reliability of the results, we divided HDL-C into continuous and categorical variables; (4) We used a GAM model and a smooth curve fitting (penalty curve method) to explore the non-linear relationship; therefore, our analysis has greater clinical value, which has not been explored in previous studies.

The current study has several following limitations. Firstly, the retrospective cohort study was based on the personal medical records of Murakami Memorial Hospital in Japan, and Takuro Okamura et al. screened the data. We could not conclude that our conclusion could be generalized to other races, areas, and some unique population because the participants were all from Japan. Similarly, we could not adjust some variables not included in the data because our study was based on the original data's reanalysis. Secondly, the incidence of diabetes may be underestimated because the original research lacked a 2 hours oral glucose tolerance test to diagnose DM. However, it was not feasible to perform 2 hours oral glucose tolerance tests for all participants due to a lack of financial and logistical support. Thirdly, only baseline HDL-C was measured in the original study, and the original study did not involve changes in HDL-C over time. Finally, our research could not exclude some residual or unmeasured confounding factors, such as dietary factors, which may bias the results. Further investigations are needed for longer-term follow-up and more population studies.

## Conclusion

HDL-C is negatively associated with DM risk. The relationship between HDL-C and DM risk is also non-linear. HDL-C is negatively related to diabetes incident when  $\text{HDL-C} \leq 1.72 \text{ mmol/L}$ . The stronger association of HDL-C and diabetes incidents is detected in the population with ex-smoker, current-smoker, and hypertension ( $\text{SBP} \geq 140 \text{ mmHg}$  or  $\text{DBP} \geq 90 \text{ mmHg}$ ).

# Abbreviations

BMI: Body mass index; WC: waist circumference; CVD: Cardiovascular diseases; SBP: systolic blood pressures; DBP: diastolic blood pressures; ALT: alanine aminotransferase; AST: aspartate aminotransferase; HDL-C: high-density lipoprotein cholesterol; TC: total cholesterol; TG: triglycerides; HbA1c: hemoglobin A1c; FPG: fasting plasma glucose; GAM: Generalized additive models; CETP: cholesteryl ester transfer protein, DM: diabetes mellitus; HR: hazard ratio; SD : standard deviations; CI : confidence interval

# Declarations

## Acknowledgments

Not applicable

## Authors' contributions

Changchun CAO and Haofei HU contributed to the study concept and design, researched and interpreted the data and drafted the manuscript. Xinyu WANG and Zhengxiao DAN researched data and reviewed the manuscript. Changchun CAO and Haofei HU oversaw the project's progress, contributed to the discussion and reviewed the manuscript. Yongcheng HE and Yulong WANG are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

## Funding

This study was supported in part by Discipline Construction Ability Enhancement Project of Shenzhen Municipal Health Commission (SZXJ2017031). This study was also supported in part by the International Cooperative Research Project of Shenzhen Municipal Science and Technology Innovation Council (accounts GJHZ2018041616481462).

## Availability of data and materials

The data are available from the 'DataDryad' database ([www.datadryad.org](http://www.datadryad.org)).

## Ethics approval and consent to participate

The ethics committee approved the original research of Murakami Memorial Hospital and informed consent was obtained from all participants

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

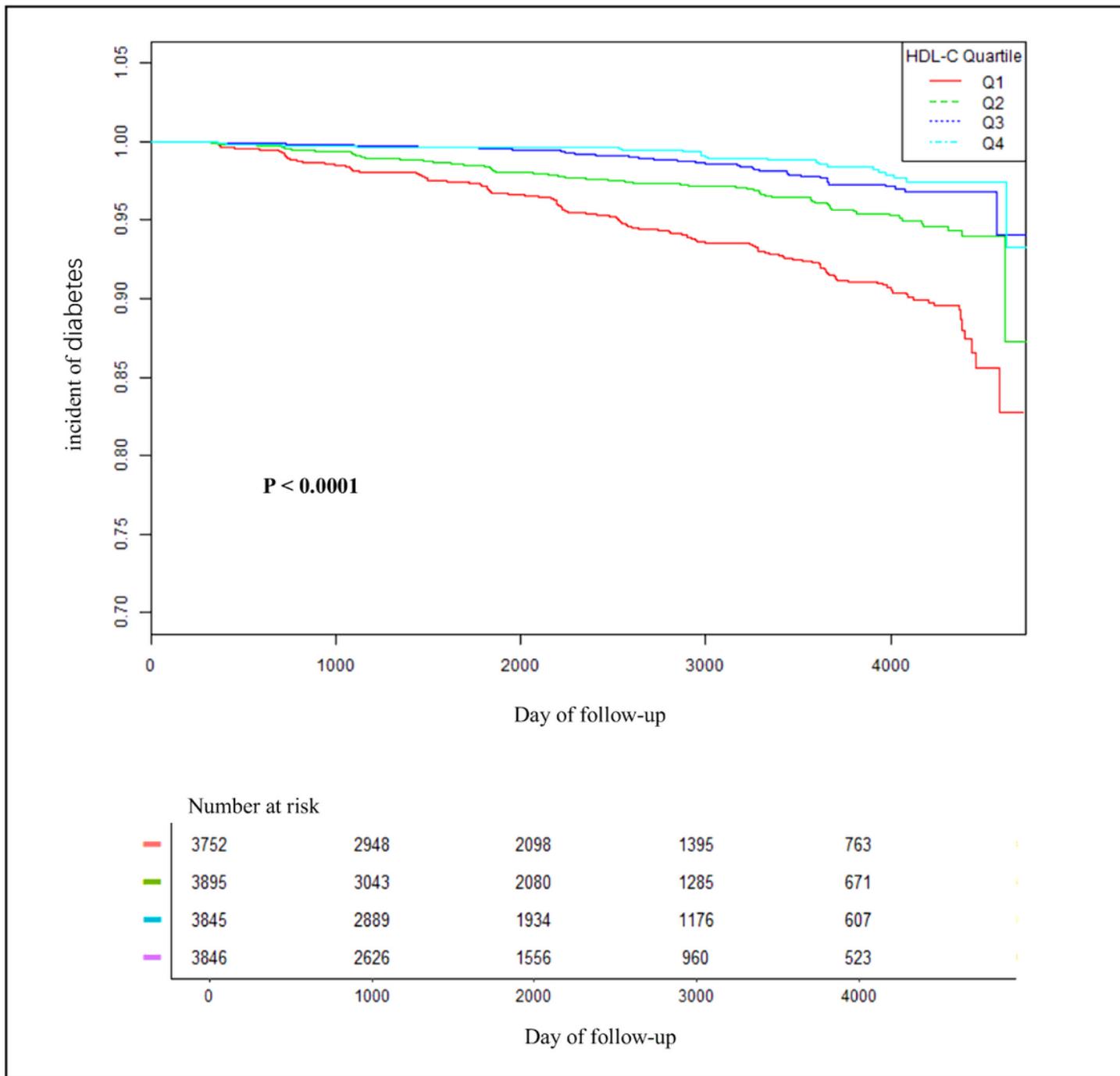
<sup>1</sup>Department of Rehabilitation, Shenzhen Dapeng New District Nan'ao People's Hospital, Shenzhen 518000, Guangdong Province, China. <sup>2</sup>Department of Nephrology, The First Affiliated Hospital of Shenzhen University, Shenzhen 518000, Guangdong Province, China. <sup>3</sup>Department of Endocrinology, The First Affiliated Hospital of Shenzhen University, Shenzhen 518000, Guangdong Province, China. <sup>4</sup>Department of Nephrology, Shenzhen Hengsheng Hospital, Shenzhen 518000, Guangdong Province, China.

## References

1. Saeedi P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract.* 2019;157:107843.
2. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care*, 2018. 41(5): p. 917–928.
3. Ali MK, et al., *Diabetes: An Update on the Pandemic and Potential Solutions.* 2017.
4. Quispe R, et al. Relationship of the triglyceride to high-density lipoprotein cholesterol (TG/HDL-C) ratio to the remainder of the lipid profile: The Very Large Database of Lipids-4 (VLDL-4) study. *Atherosclerosis.* 2015;242(1):243–50.
5. Ozder A. Lipid profile abnormalities seen in T2DM patients in primary healthcare in Turkey: a cross-sectional study. *Lipids Health Dis.* 2014;13:183.
6. Yan YZ, et al., Association of Insulin Resistance with Glucose and Lipid Metabolism: Ethnic Heterogeneity in Far Western China. *Mediators Inflamm*, 2016. 2016: p. 3825037.
7. Riserus U, Arnlov J, Berglund L. Long-term predictors of insulin resistance: role of lifestyle and metabolic factors in middle-aged men. *Diabetes Care.* 2007;30(11):2928–33.
8. Verges B. Pathophysiology of diabetic dyslipidaemia: where are we? *Diabetologia.* 2015;58(5):886–99.
9. Campbell S, Genest J. HDL-C: clinical equipoise and vascular endothelial function. *Expert Rev Cardiovasc Ther.* 2013;11(3):343–53.
10. Abbasi A, et al. Role of HDL cholesterol and estimates of HDL particle composition in future development of type 2 diabetes in the general population: the PREVENT study. *J Clin Endocrinol Metab.* 2013;98(8):E1352-9.
11. Drew BG, et al. The emerging role of HDL in glucose metabolism. *Nat Rev Endocrinol.* 2012;8(4):237–45.
12. Lin D, et al. Associations of lipid parameters with insulin resistance and diabetes: A population-based study. *Clin Nutr.* 2018;37(4):1423–9.
13. Chen Z, et al. Association of Triglyceride to high-density lipoprotein cholesterol ratio and incident of diabetes mellitus: a secondary retrospective analysis based on a Chinese cohort study. *Lipids Health Dis.* 2020;19(1):33.
14. Rader DJ, Tall AR. The not-so-simple HDL story: Is it time to revise the HDL cholesterol hypothesis? *Nat Med.* 2012;18(9):1344–6.

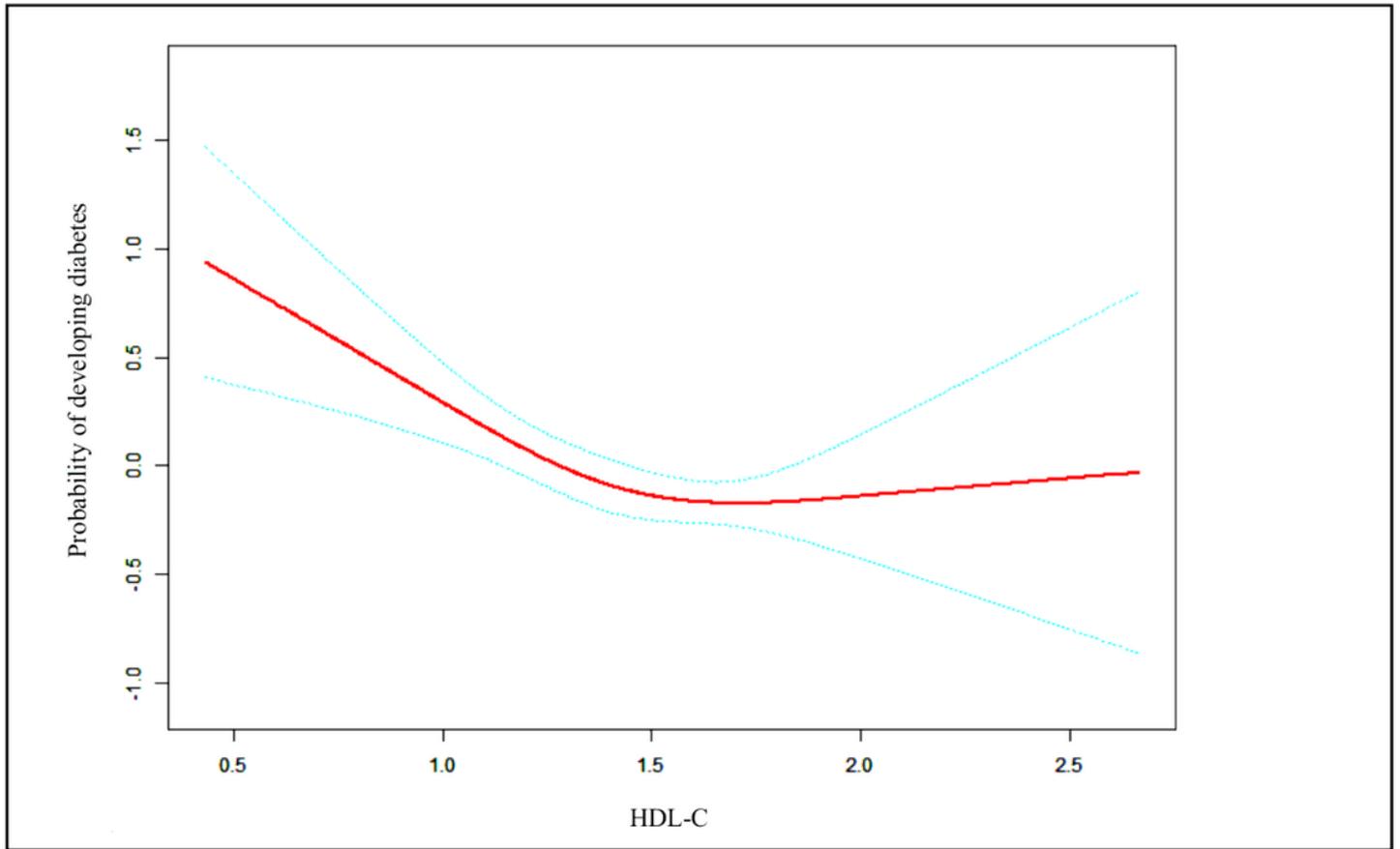
15. Okamura T, et al. Ectopic fat obesity presents the greatest risk for incident type 2 diabetes: a population-based longitudinal study. *Int J Obes (Lond)*. 2019;43(1):139–48.
16. Ryu S, et al. gamma-Glutamyltransferase as a predictor of chronic kidney disease in nonhypertensive and nondiabetic Korean men. *Clin Chem*. 2007;53(1):71–7.
17. Basevi V, et al., Comment on: American Diabetes Association. Standards of medical care in diabetes–2011. *Diabetes Care* 2011;34(Suppl. 1):S11-S61. *Diabetes Care*, 2011. 34(5): p. e53; author reply e54.
18. Zhang N, et al. Non-high-density lipoprotein cholesterol: High-density lipoprotein cholesterol ratio is an independent risk factor for diabetes mellitus: Results from a population-based cohort study. *J Diabetes*. 2018;10(9):708–14.
19. Yokoyama M, et al., Association of the Aspartate Aminotransferase to Alanine Aminotransferase Ratio with BNP Level and Cardiovascular Mortality in the General Population: The Yamagata Study 10-Year Follow-Up. *Dis Markers*, 2016. 2016: p. 4857917.
20. Vandembroucke JP, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg*. 2014;12(12):1500–24.
21. Lee SH, et al. HDL-Cholesterol, Its Variability, and the Risk of Diabetes: A Nationwide Population-Based Study. *J Clin Endocrinol Metab*. 2019;104(11):5633–41.
22. Wong N, et al., The Role of High-Density Lipoproteins in Diabetes and Its Vascular Complications. *Int J Mol Sci*, 2018. 19(6).
23. Podrez EA. Anti-oxidant properties of high-density lipoprotein and atherosclerosis. *Clin Exp Pharmacol Physiol*. 2010;37(7):719–25.
24. Ruan X, et al. Apolipoprotein A-I possesses an anti-obesity effect associated with increase of energy expenditure and up-regulation of UCP1 in brown fat. *J Cell Mol Med*. 2011;15(4):763–72.
25. Srivastava R. Dysfunctional HDL in diabetes mellitus and its role in the pathogenesis of cardiovascular disease. *Mol Cell Biochem*. 2018;440(1–2):167–87.
26. Zanoni P, et al. Rare variant in scavenger receptor BI raises HDL cholesterol and increases risk of coronary heart disease. *Science*. 2016;351(6278):1166–71.
27. Varban ML, et al. Targeted mutation reveals a central role for SR-BI in hepatic selective uptake of high density lipoprotein cholesterol. *Proc Natl Acad Sci U S A*. 1998;95(8):4619–24.

## Figures



**Figure 1**

Kaplan–Meier event-free survival curve. Kaplan–Meier event-free survival curve. Kaplan–Meier analysis of incident of diabetes based on HDL-C quartiles (log rank,  $P < 0.0001$ )



**Figure 2**

The non-linear relationship between HDL-C and incident of diabetes. A non-linear relationship between them was detected after adjusting for gender, age, ethanol consumption, smoking status, regular exerciser, SBP, DBP, BMI, WC, TC, TG, HbA1 c, FPG.