

RESEARCH

Open Access



Association of estimated small dense low-density lipoprotein cholesterol with diabetes mellitus: a large-scale multicenter retrospective cohort study

Yingkai Gao^{1*}

Abstract

Background Estimated small dense low-density lipoprotein cholesterol (sdLDL-C) was related to atherosclerosis, coronary heart disease and metabolic syndrome. Despite these findings, limited evidence is available on the relationship between sdLDL-C levels and the onset of diabetes mellitus (DM).

Methods The study analyzed data from 118,080 adults enrolled at the Rich Healthcare Group between 2010 and 2016. The relationship between sdLDL-C levels and the risk of DM was examined using Cox proportional hazards regression. In order to evaluate potential nonlinear associations, cubic spline functions and smooth curve fitting were incorporated into the Cox regression framework. Furthermore, a two-piecewise Cox proportional hazards regression was employed to pinpoint the inflection point of sdLDL-C regarding DM risk.

Results SdLDL-C was found to have a significant correlation with DM risk after controlling for confounders (HR: 1.04, 95% CI: 1.03–1.04, $P < 0.0001$). The inflection point for sdLDL-C was calculated to be 29.49 mg/dL. The HR was measured at 1.08 (95% CI: 1.06–1.10) when sdLDL-C was below 29.49 mg/dL, and it decreased to 1.03 (95% CI: 1.03–1.04) when above 29.49 mg/dL.

Conclusion This investigation reveals a nonlinear positive connection between sdLDL-C levels and the risk of developing DM in Chinese adults. Notably, sdLDL-C levels lower than 29.49 mg/dL were strongly associated with a greater risk of DM.

Clinical trial number not applicable.

Keywords Diabetes mellitus, Small dense low-density lipoprotein cholesterol, Non-linearity

*Correspondence:

Yingkai Gao
gaoyk8221@mails.jlu.edu.cn

¹The First Hospital of Jilin University, 1 Xinmin Street, Changchun, Jilin 130021, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Diabetes mellitus (DM) is identified as a chronic metabolic disease with persistent elevated glucose levels [1]. Estimates suggest that diabetes will impact 700 million people by 2045, up from 463 million in 2019 [2]. DM significantly heightens the likelihood of various complications, such as cardiovascular disease, chronic kidney disease and retinopathy [3–6]. Consequently, the early detection of individuals with a high susceptibility to DM is critically important for advancing public health effort.

Estimated small dense low-density lipoprotein cholesterol (sdLDL-C), defined as low-density lipoprotein cholesterol (LDL-C) with a density greater than 1.034 g/mL and an average particle diameter of less than 25.5 nm, exhibits distinct physical and chemical properties that make it more atherogenic than other LDL-C subclasses [7, 8]. Previous research has linked elevated sdLDL-C levels to conditions, such as atherosclerosis, metabolic syndrome and coronary heart disease [9–11]. It has been observed that sdLDL-C may exhibit heightened atherogenicity in individuals with type 2 diabetes mellitus (T2DM) due to its greater susceptibility to glycation compared to larger, more buoyant LDL particles [12]. Despite these findings, limited evidence is available on the relationship between sdLDL-C levels and the onset of DM. To address the current gap in research, the purpose of our research was to quantitatively evaluate the correlation between sdLDL-C levels and DM susceptibility in a large sample of Chinese individuals.

Methods

Dataset origin and study subjects

This study utilized a dataset from the Rich Healthcare Group, which can be accessed freely through the Dryad Data Platform, under the Dryad database's terms [13]. Ethical approval for this secondary analysis was not required, as the original study had already received approval from the Rich Healthcare Group Review Board.

The study initially included a cohort of 685,277 adults aged 20 years and above who underwent a minimum of two medical examinations from January 1, 2010, to December 31, 2016. These examinations were carried out at health-check centers located across 32 cities in China. Participants were not included if they met these criteria: (1) lack of baseline details regarding gender, fasting plasma glucose (FPG), height or weight ($n=135,317$); (2) body mass index (BMI) that are either less than 15 or greater than 55 kg/m² ($n=152$); (3) follow-up duration shorter than 2 years ($n=324,233$); (4) a prior diagnosis of DM at baseline or an indeterminate DM status during follow-up ($n=13,742$); (5) incomplete baseline data on LDL-C and triglycerides (TG) ($n=93,747$); and (6) sdLDL-C levels below 0 ($n=6$). Following the

application of these exclusion criteria, the study's final cohort included 118,080 participants (Fig. 1).

Data collection

The selection of covariates for this study was guided by prior research and clinical expertise on risk factors associated with DM [14–17]. All data collection and measurement staff underwent rigorous training before the study began to ensure consistency in data collection. The dataset encompassed demographic information (age, gender), anthropometric data (height, weight), lifestyle habits, family history of DM, and systolic/diastolic blood pressure (SBP/DBP) and was collected using a standardized questionnaire. After a minimum of 8 h fasting period, venous blood samples were collected from the participants. Biochemical markers, including TG, serum creatinine (Scr), LDL-C, high-density lipoprotein cholesterol (HDL-C), aspartate aminotransferase (AST), FPG, blood urea nitrogen (BUN), alanine aminotransferase (ALT) and total cholesterol (TC) were measured using an automated biochemical analyzer.

Formula for sdLDL-C

The sdLDL-C formula used in this study is based on established lipid profiling methods and has been widely applied in previous research to estimate sdLDL-C levels in large-scale epidemiological studies [18–21]. This formula leverages the relationships among lipid parameters, including LDL-C, and TG, to provide an indirect but reliable estimation of sdLDL-C concentrations. The estimated sdLDL-C (mg/dL) was determined using the following equation: $\text{sdLDL-C (mg/dL)} = \text{LDL-C} - (1.43 \times \text{LDL-C} - (0.14 \times \ln(\text{TG}) \times \text{LDL-C}) - 8.99)$.

Outcome measures

DM was characterized by the American Diabetes Association's standards as either a FPG measurement ≥ 7.0 mmol/L or self-reported DM during the follow-up [22].

Data analysis

Statistical analyses were carried out using Empower Stats (X&Y Solutions Inc., Boston, MA) version 4.1 and R software version 4.2.2. Normally distributed variables are presented as means \pm standard deviations, with group comparisons performed using One-Way ANOVA. Skewed data are represented by medians and interquartile ranges, and the Kruskal-Wallis H test was applied for group comparisons. The presentation of categorical variables was in the form of frequencies (%), and chi-square tests were used to assess group differences. A two-tailed P value below 0.05 indicated statistical significance.

Among the 118,080 participants, the number and proportion of missing data for each variable were as follows: SBP ($n=18$, 0.02%), DBP ($n=18$, 0.02%), HDL-C

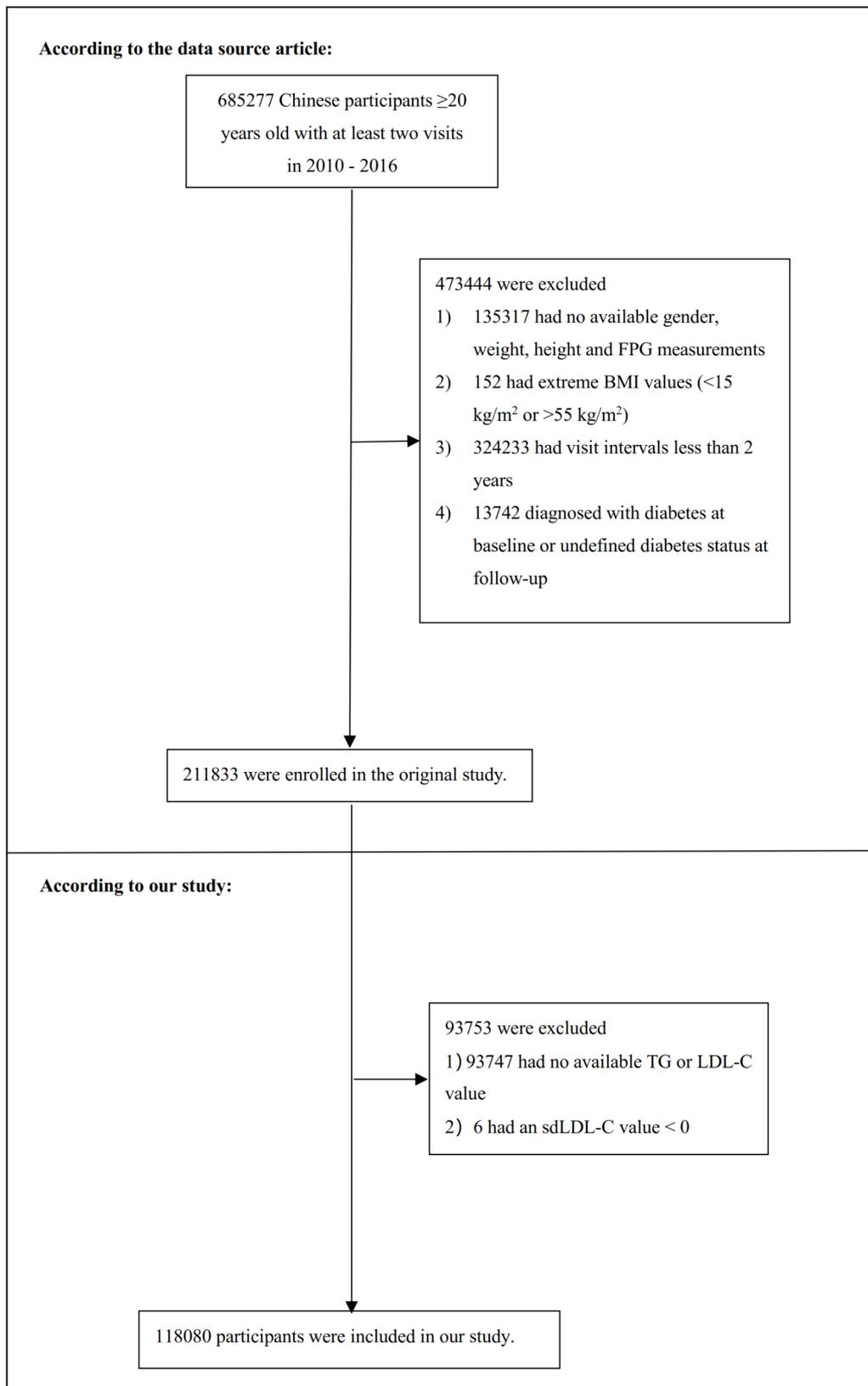


Fig. 1 Study population

($n = 1,333$, 1.13%), AST ($n = 67,529$, 57.19%), ALT ($n = 409$, 0.35%), TC ($n = 1$, <0.01%), Scr ($n = 1,338$, 1.13%), BUN ($n = 3,148$, 2.67%), drinking status ($n = 84,653$, 71.69%), and smoking status ($n = 84,653$, 71.69%). We performed five imputations to ensure the stability and reliability of the imputed datasets. The imputation model included the following variables: gender, BMI, DBP, SBP, drinking status, family history of diabetes, smoking status, age, ALT, HDL-C, BUN, TC, Scr, FPG, and AST. The low missingness rates for most variables (e.g., 0.02% for SBP/DBP, 0.35% for ALT) suggest that missingness is unlikely to depend systematically on unobserved factors. For variables with higher missingness (e.g., AST: 57.19%, smoking/drinking status: 71.69%), we verified that missingness patterns were unrelated to diabetes status or sdLDL-C levels ($P > 0.10$ in logistic regression models predicting missingness). Sensitivity analyses comparing distributions of observed vs. imputed data (e.g., mean/median values, frequencies) showed no significant deviations, supporting the plausibility of missing at random. It was assumed in the analysis that the absent data were missing randomly [23, 24].

In accordance with the STROBE guidelines, three multivariate Cox proportional hazards regression models were constructed. The hazard ratios (HR) and 95% confidence intervals (CI) were calculated to evaluate the connection of sdLDL-C levels with the likelihood of DM development. Model 1 served as the baseline model without any covariate adjustments. Adjustments for variables including age, SBP, DBP, BMI, smoking status, family diabetes history, drinking status, and gender were incorporated in Model 2. Model 3 built upon Model 2 by incorporating additional covariates, including HDL-C, TC, ALT, BUN, FPG, Scr and AST. We assessed the proportional hazards assumption by incorporating an interaction term between sdLDL-C and the logarithm of follow-up time in the Cox regression model. The analysis revealed no significant violations of the proportional hazards assumption. To assess multicollinearity among the variables, the variance inflation factor (VIF) was calculated for all covariates. A VIF threshold of > 5 was used to identify collinear variables, which would be excluded from the multivariate Cox proportional hazards regression models. However, based on this criterion, no covariates were identified as collinear, and thus none were excluded from the analysis (Supplementary Table S1).

Sensitivity analyses were conducted to rigorously evaluate the reliability of our findings. Considering the recognized associations of prediabetes, obesity and hypertension with DM [25–27], individuals with $\text{BMI} \geq 24 \text{ kg/m}^2$, $\text{FPG} \geq 6.1 \text{ mmol/L}$, or those with

elevated blood pressure (DBP of 90 mmHg or more, or SBP of 140 mmHg or more) were excluded from the study. Finally, the E-value was calculated to evaluate the potential influence of unmeasured confounders on the association between sdLDL-C and DM.

In the Cox regression model, the potential nonlinear correlation between sdLDL-C levels and the probability of DM was discovered by applying cubic spline functions and smooth curve fitting. If non-linearity was detected, a recursive algorithm was employed to identify the inflection point. The process began by dividing the dataset into two segments at an arbitrary starting point within the range of sdLDL-C values. For each partition, a two-piece Cox proportional hazards regression model was constructed to evaluate the relationship on either side of the candidate inflection point. The log-likelihood ratio test was then used to determine the most suitable model for explaining the association between the sdLDL-C and the risk of DM.

Results

Baseline features

This research comprised 118,080 participants averaging 44.31 ± 13.08 years in age, of whom 63,709 (53.95%) were male. Baseline features of the participants, categorized by sdLDL-C quartiles, are displayed in Table 1. Individuals in the highest sdLDL-C quartile (Q4 group) tend to be older, male, current drinkers and current smokers in contrast to individuals in the lowest sdLDL-C quartile (Q1 group). Subjects assigned to the Q4 group also exhibit higher SBP, DBP, and BMI, along with elevated levels of ALT, TG, AST, TC, BUN, LDL-C, Scr, and FPG. Conversely, the greatest HDL-C levels are found in the Q1 group.

The incidence rate of DM

As indicated in Tables 2, and 2,737 individuals (2.32%) developed DM over an average follow-up duration of 3.10 years. The incidence rates of DM for the entire cohort and across quartiles of the sdLDL-C were 2.32% (95%CI: 2.23–2.40), 0.67% (95%CI: 0.58–0.76), 1.62% (95%CI: 1.48–1.77), 2.66% (95%CI: 2.48–2.85), and 4.32% (95%CI: 4.08–4.55) for the total population, Q1, Q2, Q3, and Q4, respectively. The overall cohort had a cumulative incidence rate of 7.48 per 1,000 person-years, while the rates for groups Q1, Q2, Q3, and Q4 were 2.14, 5.23, 8.65, and 13.97 per 1,000 person-years, respectively. Participants in higher sdLDL-C quartiles exhibited significantly elevated DM incidence rates. These findings were further corroborated by the Kaplan-Meier curve illustrating cumulative hazard (Fig. 2).

Table 1 The baseline characteristics of participants

sdLDL-C (mg/dL)	Q1(≤ 24.14)	Q2(24.14 to ≤ 31.15)	Q3(31.15 to ≤ 39.76)	Q4(> 39.76)	P-value
Participants	29,518	29,521	29,519	29,522	
Gender					< 0.001
Male	10,309 (34.92%)	14,733 (49.91%)	18,259 (61.86%)	20,408 (69.13%)	
Female	19,209 (65.08%)	14,788 (50.09%)	11,260 (38.14%)	9114 (30.87%)	
Age (years)	39.00 ± 10.91	42.75 ± 12.57	46.28 ± 13.31	49.21 ± 13.11	< 0.001
BMI (kg/m ²)	21.54 ± 2.72	22.77 ± 3.05	24.01 ± 3.17	25.13 ± 3.10	< 0.001
SBP (mmHg)	113.24 ± 14.61	117.37 ± 15.88	121.72 ± 16.57	125.75 ± 16.96	< 0.001
DBP (mmHg)	70.33 ± 9.82	73.07 ± 10.45	75.86 ± 10.80	78.63 ± 11.01	< 0.001
Drinking status					< 0.001
Current-drinker	378 (1.28%)	608 (2.06%)	799 (2.71%)	1041 (3.53%)	
Ex-drinker	3070 (10.40%)	3845 (13.02%)	4684 (15.87%)	4957 (16.79%)	
Never- drinker	26,070 (88.32%)	25,068 (84.92%)	24,036 (81.43%)	23,524 (79.68%)	
Smoking status					< 0.001
Current-smoker	2597 (8.80%)	4170 (14.13%)	5910 (20.02%)	7369 (24.96%)	
Ex-smoker	699 (2.37%)	938 (3.18%)	1199 (4.06%)	1311 (4.44%)	
Never-smoker	26,222 (88.83%)	24,413 (82.70%)	22,410 (75.92%)	20,842 (70.60%)	
Family history of diabetes					0.429
No	28,892 (97.88%)	28,843 (97.70%)	28,845 (97.72%)	28,850 (97.72%)	
Yes	626 (2.12%)	678 (2.30%)	674 (2.28%)	672 (2.28%)	
ALT (U/L)	14.00 (11.00–19.00)	16.60 (12.10–24.00)	20.00 (14.20–29.00)	25.00 (17.60–37.80)	< 0.001
AST (U/L)	20.30 (16.39–25.12)	21.63 (17.30–27.00)	23.00 (18.50–28.83)	25.39 (20.40–32.00)	< 0.001
TG (mg/dL)	53.14 (44.28–63.77)	82.37 (70.86–98.31)	116.03 (97.43–145.25)	183.34 (143.48–243.57)	< 0.001
TC (mg/dL)	157.82 ± 23.77	173.97 ± 24.17	189.56 ± 24.86	219.63 ± 31.13	< 0.001
LDL-C (mg/dL)	86.23 ± 17.13	99.22 ± 17.45	110.98 ± 19.34	132.06 ± 25.66	< 0.001
HDL-C (mg/dL)	55.83 ± 11.43	54.49 ± 11.97	51.95 ± 11.66	50.03 ± 11.56	< 0.001
BUN (mg/dL)	12.70 ± 3.27	12.91 ± 3.32	13.31 ± 3.31	13.60 ± 3.26	< 0.001
Scr (mg/dL)	0.74 ± 0.17	0.78 ± 0.19	0.82 ± 0.18	0.84 ± 0.18	< 0.001
FPG (mg/dL)	86.47 ± 9.70	88.01 ± 10.51	89.87 ± 11.09	92.23 ± 11.68	< 0.001

Values are n (%) or mean ± SD or median (quartile)

sdLDL-C: estimated small dense low-density lipoprotein cholesterol; BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; ALT: alanine aminotransferase; AST: aspartate aminotransferase; TC: total cholesterol; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; TG: triglycerides; BUN: blood urea nitrogen; Scr: serum creatinine; FPG: fasting plasma glucose

Table 2 The rate of incident diabetes

sdLDL-C	Participants (n)	Diabetes events (n)	Cumulative incidence (95% CI) (%)	Per 1000 person-year
Overall	118,080	2737	2.32 (2.23–2.40)	7.48
Q1	29,518	198	0.67 (0.58–0.76)	2.14
Q2	29,521	479	1.62 (1.48–1.77)	5.23
Q3	29,519	786	2.66 (2.48–2.85)	8.65
Q4	29,522	1274	4.32 (4.08–4.55)	13.97
P for trend			< 0.001	< 0.001

sdLDL-C: estimated small dense low-density lipoprotein cholesterol; CI, confidence interval

Univariate analysis

Table 3 summarizes the outcomes of the univariate Cox regression analysis, highlighting factors associated with the risk of DM. Identified risk factors include sdLDL-C, age, BMI, SBP, DBP, alcohol intake, smoking, a family history of DM, and elevated levels of TC, ALT, LDL-C, AST,

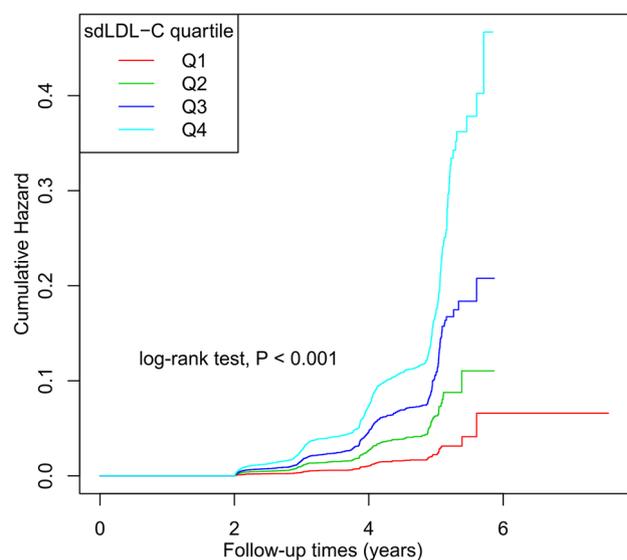


Fig. 2 Kaplan-Meier Cumulative Hazard curve. Kaplan-Meier analysis of incident diabetes based on sdLDL-C levels quartiles (log-rank, $P < 0.001$)

Table 3 The results of the univariate analysis

	Statistics	HR (95%CI)	P value
Gender			
Male	63,709 (53.95%)	Ref	
Female	54,371 (46.05%)	0.50 (0.46, 0.54)	< 0.0001
Age (years)	44.31 ± 13.08	1.06 (1.06, 1.07)	< 0.0001
BMI (kg/m ²)	23.36 ± 3.30	1.22 (1.21, 1.23)	< 0.0001
SBP (mmHg)	119.52 ± 16.70	1.04 (1.04, 1.04)	< 0.0001
DBP (mmHg)	74.47 ± 10.97	1.04 (1.04, 1.05)	< 0.0001
Drinking status			
Current-drinker	2826 (2.39%)	Ref	
Ex-drinker	16,556 (14.02%)	0.55 (0.45, 0.67)	< 0.0001
Never-drinker	98,698 (83.59%)	0.43 (0.36, 0.51)	< 0.0001
Smoking status			
Current-smoker	20,046 (16.98%)	Ref	
Ex-smoker	4147 (3.51%)	0.79 (0.65, 0.96)	0.0167
Never-smoker	93,887 (79.51%)	0.57 (0.52, 0.62)	< 0.0001
Family history of diabetes			
No	115,430 (97.76%)	Ref	
Yes	2650 (2.24%)	1.36 (1.11, 1.67)	0.0026
ALT (U/L)	23.81 ± 21.73	1.00 (1.00, 1.00)	< 0.0001
AST (U/L)	24.16 ± 12.41	1.01 (1.01, 1.01)	< 0.0001
TC (mg/dL)	185.25 ± 34.70	1.01 (1.01, 1.01)	< 0.0001
TG (mg/dL)	122.24 ± 92.49	1.00 (1.00, 1.00)	< 0.0001
LDL-C (mg/dL)	107.12 ± 26.29	1.01 (1.01, 1.01)	< 0.0001
HDL-C (mg/dL)	53.07 ± 11.87	0.99 (0.98, 0.99)	< 0.0001
BUN (mg/dL)	13.13 ± 3.31	1.07 (1.06, 1.08)	< 0.0001
Scr (mg/dL)	0.80 ± 0.18	1.66 (1.53, 1.80)	< 0.0001
FPG (mg/dL)	89.14 ± 10.98	1.14 (1.13, 1.14)	< 0.0001
sdLDL-C (mg/dL)	32.81 ± 11.72	1.04 (1.04, 1.04)	< 0.0001

BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; ALT: alanine aminotransferase; AST: aspartate aminotransferase; TC: total cholesterol; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; TG: triglycerides; BUN: blood urea nitrogen; Scr: serum creatinine; FPG: fasting plasma glucose; sdLDL-C: estimated small dense low-density lipoprotein cholesterol; HR, hazard ratios; CI, confidence interval; Ref, reference

BUN, TG, Scr and FPG. Conversely, HDL-C and female sex were found to be protective factors against DM.

The relevance of sdLDL-C to DM

Table 4 provides a detailed multivariate analysis that evaluated the independent effect of sdLDL-C on the risk of DM development. In Model 1, without adjusting for confounding variables, a significant correlation was found between sdLDL-C and DM risk (HR: 1.04, 95% CI: 1.04–1.04, $P < 0.0001$). This relationship persisted in Model 2 (HR: 1.02, 95% CI: 1.02–1.02, $P < 0.0001$), which adjusted for gender, SBP, age, BMI, DBP, drinking status, smoking status and family history of DM. In Model 3, the association continued to be statistically significant (HR: 1.04, 95% CI: 1.03–1.04, $P < 0.0001$), which included additional adjustments for HDL-C, TC, ALT, AST, FPG, Scr, and BUN based on Model 2. Furthermore, individuals in the highest sdLDL-C quartile experienced a 194% increased

Table 4 Relationship between sdLDL-C and incident diabetes in different models

Exposure	Model 1 (HR, 95% CI, P)	Model 2 (HR, 95% CI, P)	Model 3 (HR, 95% CI, P)
sdLDL-C	1.04 (1.04, 1.04) < 0.0001	1.02 (1.02, 1.02) < 0.0001	1.04 (1.03, 1.04) < 0.0001
sdLDL-C (quartile)			
Q1	Ref	Ref	Ref
Q2	2.56 (2.17, 3.02) < 0.0001	1.58 (1.34, 1.87) < 0.0001	1.69 (1.43, 2.01) < 0.0001
Q3	4.34 (3.71, 5.07) < 0.0001	1.80 (1.54, 2.12) < 0.0001	2.09 (1.76, 2.48) < 0.0001
Q4	7.07 (6.09, 8.21) < 0.0001	2.27 (1.94, 2.65) < 0.0001	2.94 (2.44, 3.55) < 0.0001
P for trend	< 0.0001	< 0.0001	< 0.0001

Model 1: we did not adjust for any covariants

Model 2: we adjusted for gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, and BMI

Model 3: we adjusted for gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

sdLDL-C: estimated small dense low-density lipoprotein cholesterol; HR: hazard ratios; CI: confidence interval; Ref: reference

risk of DM (HR: 2.94, 95% CI: 2.44–3.55, $P < 0.0001$) when contrasted with the lowest quartile in Model 3

The sensitivity analysis findings

As shown in Model 4 of Table 5, the sensitivity analysis on individuals without BMI ≥ 24 kg/m² confirmed a significant positive correlation between sdLDL-C and the risk of DM (HR: 1.04, 95% CI: 1.03–1.05, $P < 0.0001$). Similarly, an additional analysis focusing on individuals with SBP not reaching 140 or DBP not reaching 90 mmHg, similarly demonstrated a significant positive relationship between sdLDL-C and DM incidence (HR: 1.04, 95% CI: 1.03–1.05, $P < 0.0001$), displayed in in Model 5 of Table 5. Furthermore, among individuals with FPG < 6.1 mmol/L, a consistent positive relationship between sdLDL-C and DM was observed (Model 6; HR: 1.04, 95% CI: 1.04–1.05, $P < 0.0001$). In addition, the E-value (1.25) was found to exceed the relative risk associated with sdLDL-C and potential unmeasured confounders (1.24). This indicates that the influence of unknown or unmeasured confounders on the observed relationship between sdLDL-C and DM is likely minimal.

The assessment of the non-linear connection

As demonstrated in Table 6; Fig. 3, a nonlinear connection between sdLDL-C levels and DM risk was discovered using the Cox proportional hazards regression model with cubic spline functions and smooth curve fitting. The inflection point for sdLDL-C was calculated to be 29.49 mg/dL through a recursive algorithm. A segmented Cox regression analysis was then performed to

Table 5 Relationship between sdLDL-C levels and incident diabetes in different sensitivity analyses

Exposure	Model 4 (HR,95%CI, P)	Model 5 (HR,95%CI, P)	Model 6 (HR,95%CI, P)
sdLDL-C	1.04 (1.03, 1.05) <0.0001	1.04 (1.03, 1.05) <0.0001	1.04 (1.04, 1.05) <0.0001
sdLDL-C (quartile)			
Q1	Ref	Ref	Ref
Q2	1.61 (1.24, 2.09) 0.0004	1.53 (1.25, 1.88) <0.0001	1.67 (1.34, 2.08) <0.0001
Q3	2.14 (1.64, 2.81) <0.0001	2.01 (1.64, 2.47) <0.0001	2.24 (1.80, 2.78) <0.0001
Q4	3.02 (2.21, 4.14) <0.0001	2.92 (2.32, 3.68) <0.0001	3.17 (2.47, 4.07) <0.0001
P for trend	<0.0001	<0.0001	<0.0001

Model 4 was sensitivity analysis in participants without BMI ≥ 24 kg/m². We adjusted gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

Model 5 was sensitivity analysis in participants without SBP ≥ 140 or DBP ≥ 90 mmHg. We adjusted gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

Model 6 was sensitivity analysis in participants without FPG ≥ 6.1 mmol/L. We adjusted gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

sdLDL-C: estimated small dense low-density lipoprotein cholesterol; HR, hazard ratios; CI, confidence; Ref: reference

Table 6 The result of the two-piecewise Cox proportional hazards regression model

Incident diabetes	HR (95%CI)	P
Fitting model by two-piecewise Cox proportional hazards regression		
The inflection point of sdLDL-C	29.49	
≤ 29.49	1.08 (1.06, 1.10)	<0.0001
> 29.49	1.03 (1.03, 1.04)	<0.0001
P for the log-likelihood ratio test	<0.001	

We adjusted for gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

sdLDL-C: estimated small dense low-density lipoprotein cholesterol; HR: hazard ratios; CI: confidence

estimate the HR and CI on both sides of this threshold. The HR was measured at 1.08 (95% CI: 1.06–1.10) when

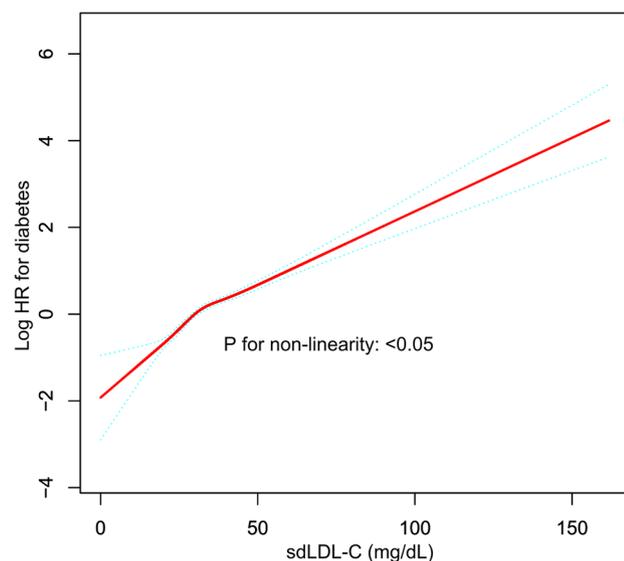


Fig. 3 The nonlinear relationship between sdLDL-C levels and incident diabetes. A nonlinear relationship between sdLDL-C levels and incident diabetes was detected after adjustment for gender, age, SBP, DBP, family history of diabetes, drinking status, smoking status, BMI, HDL-C, TC, ALT, AST, Scr, BUN, and FPG

sdLDL-C was below 29.49 mg/dL, and it decreased to 1.03 (95% CI: 1.03–1.04) when above 29.49 mg/dL.

Discussion

This large-scale retrospective cohort study revealed a non-linear positive relationship between sdLDL-C and the risk of developing DM. Notably, sdLDL-C levels below 29.49 mg/dL were more strongly associated with an elevated risk of DM.

Increasing evidence indicates that sdLDL-C is correlated with atherosclerosis, coronary heart disease and diabetes status [28–30]. In a study of 3,684 T2DM patients undergoing selective coronary angiography, elevated serum sdLDL-C was found to increase coronary heart disease severity and predict future cardiovascular events [31]. Similarly, a cohort study involving 887 community members found that elevated sdLDL-C levels were linked to a higher risk of carotid plaque formation in individuals with normal LDL-C levels, after controlling for variables including age, smoking, BMI, sex, alcohol use, hypertension, diabetes, and follow-up duration [32]. Another study involving 4,388 Japanese participants showed significantly elevated sdLDL-C levels in the early stages of metabolic syndrome and T2DM with impaired glucose tolerance compared to healthy controls [33]. Furthermore, an investigation involving 594 healthy participants between the ages of 35 and 65 without coronary heart disease revealed a meaningful relationship between sdLDL-C quartiles and glucose metabolism markers, even after adjusting for gender, hypertension, BMI, household income, alcohol consumption, smoking and age [34]. In addition, a cross-sectional observational study of newly identified T2DM individuals with ideal or nearly ideal lipid levels also found a greater proportion of sdLDL-C compared to healthy controls [35]. Despite these findings, limited evidence is available on the connection between sdLDL-C levels and the onset of DM. In our study, we identified a nonlinear, positive association

between sdLDL-C and DM risk after adjusting for confounders. Our observations revealed an inflection point at a sdLDL-C concentration of 29.49 mg/dL. Below this threshold, an increase of one unit in sdLDL-C correlated with an 8% greater risk of developing DM (HR = 1.08, 95% CI: 1.06–1.10). Above the inflection point, the risk increase was more modest, with each additional unit of sdLDL-C corresponding to a 3% increased likelihood of developing DM (HR = 1.03, 95% CI: 1.03–1.04). These findings imply that reducing sdLDL-C levels is associated with a lower risk of DM, with a more pronounced risk reduction observed when sdLDL-C levels fall below 29.49 mg/dL. Sensitivity analyses further confirmed the robustness of this association, showing consistent results in individuals without obesity or hypertension. Our findings suggest that sdLDL-C may serve as a valuable biomarker for the risk stratification of DM. Given the observed association between elevated sdLDL-C levels and an increased risk of DM, sdLDL-C could potentially be incorporated into clinical practice to identify individuals at higher risk for developing DM. This would enable earlier interventions, such as lifestyle modifications or targeted therapies, aimed at reducing DM incidence. Furthermore, sdLDL-C could complement existing risk assessment tools, particularly in populations where traditional risk factors, such as obesity or family history, may not fully capture DM risk. Future studies are warranted to validate these findings in diverse populations and to explore the feasibility of integrating sdLDL-C into routine clinical risk assessment frameworks. Such efforts could ultimately enhance the precision of DM prevention strategies and improve patient outcomes.

In our study, the nonlinear relationship between sdLDL-C and DM risk, with an inflection point at 29.49 mg/dL, suggests that the association may vary across different sdLDL-C levels. Participants with sdLDL-C levels above the inflection point were observed to have higher BMI, BUN, FPG, TC, TG, LDL-C, and Scr levels, as well as lower HDL-C levels (Table S2). These factors are well-established risk factors for diabetes and may collectively attenuate the independent contribution of sdLDL-C to DM risk, thereby weakening the association between sdLDL-C and DM above the inflection point. Additionally, at higher sdLDL-C levels, the metabolic pathways contributing to diabetes risk, such as oxidative stress, inflammation, and insulin resistance, may already be maximally activated [29, 34, 36]. This saturation effect could result in a plateau or reduction in the relative contribution of sdLDL-C to DM risk. In contrast, below the inflection point, the relative contribution of sdLDL-C to DM risk may be more pronounced due to the absence or lower levels of other competing risk factors. This could explain the steeper increase in DM risk at lower sdLDL-C levels. Further studies are needed to validate these

hypotheses and elucidate the precise biological mechanisms underlying this nonlinear relationship.

The exact role of sdLDL-C in the development of DM remains inadequately understood. Increased concentrations of sdLDL-C were related to insulin resistance, a fundamental metabolic abnormality [37, 38]. Evidence from prior researches demonstrates that triglyceride metabolism plays a significant role in regulating sdLDL-C production [39]. In states of insulin resistance, an elevated availability of free fatty acids (FFAs) drives liver release of very low-density lipoprotein (VLDL) while simultaneously inhibiting the degradation of apolipoprotein B [40]. The excessive production of VLDL1 facilitates the generation of sdLDL-C through metabolic pathways involving cholesteryl ester transfer protein and hepatic lipase [39, 40]. Furthermore, elevated sdLDL-C levels have been implicated in worsening insulin resistance. It was proposed that the predominance of sdLDL-C may act as a central element in the progress of insulin resistance in individuals without diabetes [41]. This bidirectional interaction suggests that while insulin resistance contributes to increased sdLDL-C levels, the latter may, aggravate insulin resistance in turn. Based on these observations, we propose that a combination of mechanisms, including both insulin resistance and insulin deficiency, may contribute to the development of diabetes in individuals with elevated sdLDL-C levels.

This study presents several significant strengths. First, it is a large-scale cohort study conducted in China, providing an extensive and reliable dataset that improves the applicability of our findings to broader populations. Additionally, we confirmed the consistent association between sdLDL-C and diabetes risk, even after considering numerous traditional confounders, which underscores the robustness of this link. Importantly, our analysis identified a non-linear, positive correlation between sdLDL-C and DM, offering new perspectives to the existing research. Finally, the sensitivity analyses were applied to ensure the validity and reliability of our findings, further strengthening the methodological rigor of the study.

It is necessary to recognize the limitations of this study. First, the definition of DM applied in this research excluded oral glucose tolerance testing, which could have led to an underestimation of DM prevalence. Second, the absence of repeated measurements of sdLDL-C restricted our ability to evaluate the role of temporal alterations in sdLDL-C levels in influencing the risk of DM. Third, since the study was conducted exclusively within a Chinese population, the findings may have restricted applicability to geographically or ethnically diverse populations. Fourth, sdLDL-C levels were estimated using a validated formula based on LDL-C, and TG levels in this study. While this formula provides a

practical and cost-effective alternative to direct measurement methods, such as ultracentrifugation, it is important to acknowledge its potential limitations. The use of an estimated value may introduce some degree of measurement error, which could affect the precision of the observed associations. However, the formula has been validated in large population-based studies and has demonstrated strong correlations with directly measured sdLDL-C levels [18–21]. Although we were unable to validate the formula specifically in our study population due to the lack of directly measured sdLDL-C data, validation studies conducted in similar populations support its applicability in this context. Fifth, despite our efforts to adjust for known confounders, unmeasured variables such as antihyperlipidemic drug, weight loss, dietary habits, physical activity, socioeconomic status and genetic predispositions may have influenced the observed association between sdLDL-C and incident diabetes. However, our calculation of the E-value (1.25) exceeds the estimated relative risk of the association between sdLDL-C and potential unmeasured confounders (1.24). This finding suggests that the observed relationship between sdLDL-C and incident diabetes is likely to be independent of these unmeasured factors. In future prospective studies, we will endeavour to minimise missing data and collect and integrate data on antihyperlipidemic drug, weight loss, dietary habits, physical activity, socioeconomic status and genetic predispositions to ensure a wider representation of the population in order to further validate and strengthen our findings. Sixth, the reliance on FPG and self-reported DM may result in a misclassification bias, potentially underestimating the true prevalence of DM in the study population. This could attenuate the observed association between sdLDL-C and DM, as some cases of DM may have been missed. In the future, we will design studies to further validate the stability of our results by diagnosing diabetes through oral glucose tolerance test and HbA1c. Seventh, although our findings suggest a significant association between elevated sdLDL-C levels and an increased risk of DM, the potential for reverse causality cannot be excluded. Individuals with higher sdLDL-C levels may have underlying metabolic abnormalities, such as insulin resistance or sub-clinical inflammation, that contribute both to elevated sdLDL-C levels and an increased risk of DM. While we adjusted for known confounders in our analysis, residual confounding from unmeasured metabolic factors may still exist. Future longitudinal studies with repeated measurements of sdLDL-C and metabolic markers are needed to better elucidate the temporal relationship and causality between sdLDL-C and DM. Finally, the original study excluded participants with visit intervals of less

than two years and did not provide information on participants with less than 2 years of follow-up. In the present study, 4,636 participants (3.93%) had FPG levels ≥ 6.1 mmol/L, making the development of diabetes within two years inevitable. Therefore, we preferred to exclude participants with less than two years of follow-up, which could have led to selection bias. For future research, we plan to design a new study that includes participants with follow-up durations of less than two years. This approach will enable us to evaluate the effects of short-term follow-up on changes in glucose metabolism and provide a more comprehensive understanding of the study population.

Conclusion

This investigation reveals a nonlinear positive connection between sdLDL-C levels and the risk of developing DM in Chinese adults. Notably, sdLDL-C levels lower than 29.49 mg/dL were strongly associated with a greater risk of DM. It was emphasized that the critical role of sdLDL-C as a predictive biomarker for DM, highlighting its potential application in clinical settings for risk stratification and the development of preventive interventions.

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1186/s12902-025-01880-w>.

Supplementary Material 1

Acknowledgements

We greatly appreciate the raw data provided by Chen, Ying et al.

Author contributions

Yingkai Gao designed the study, performed the statistical analysis and drafted the manuscript.

Funding

This study did not receive any funding.

Data availability

The raw data utilized in this study are available for download from 'DATADRYAD' database (<https://datadryad.org/stash/dataset/doi:10.5061/dryad.ft8750v>).

Declarations

Ethics approval and participant consent

The original study received approval from the Rich Healthcare Group Review Board and waived the requirement for informed consent. The research has been performed in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 6 January 2025 / Accepted: 13 February 2025

Published online: 11 March 2025

References

1. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2011;34(Suppl 1):S62–9.
2. Williams R, Karuranga S, Malanda B, Saeedi P, Basit A, Besancon S, Bommer C, Esteghamati A, Ogurtsova K, Zhang P, et al. Global and regional estimates and projections of diabetes-related health expenditure: results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pr*. 2020;162:108072.
3. Balakumar P, Maung-U K, Jagadeesh G. Prevalence and prevention of cardiovascular disease and diabetes mellitus. *Pharmacol Res*. 2016;113(Pt A):600–9.
4. Thomas MC, Cooper ME, Zimmet P. Changing epidemiology of type 2 diabetes mellitus and associated chronic kidney disease. *Nat Rev Nephrol*. 2016;12(2):73–81.
5. Himasa FI, Singhal M, Ojha A, Kumar B. Prospective for diagnosis and treatment of Diabetic Retinopathy. *Curr Pharm Des*. 2022;28(7):560–9.
6. Yun JS, Ko SH. Current trends in epidemiology of cardiovascular disease and cardiovascular risk management in type 2 diabetes. *Metabolism*. 2021;123:154838.
7. Kanonidou C. Small dense low-density lipoprotein: Analytical review. *Clin Chim Acta*. 2021;520:172–8.
8. Higashioka M, Sakata S, Honda T, Hata J, Yoshida D, Hirakawa Y, Shibata M, Goto K, Kitazono T, Osawa H, et al. Small Dense Low-Density Lipoprotein Cholesterol and the risk of Coronary Heart Disease in a Japanese Community. *J Atheroscler Thromb*. 2020;27(7):669–82.
9. Duran EK, Aday AW, Cook NR, Buring JE, Ridker PM, Pradhan AD. Triglyceride-Rich Lipoprotein Cholesterol, small dense LDL cholesterol, and Incident Cardiovascular Disease. *J Am Coll Cardiol*. 2020;75(17):2122–35.
10. Ikezaki H, Lim E, Cupples LA, Liu CT, Szatlos BF, Schaefer EJ. Small dense low-density lipoprotein cholesterol is the most atherogenic lipoprotein parameter in the prospective Framingham offspring study. *J Am Heart Assoc*. 2021;10(5):e19140.
11. Zhu X, Chen Y, Zhu M, Hu J. The relationship between small dense low-density lipoprotein cholesterol and metabolic syndrome. *Diabet Metab Syndr Ob*. 2024;17:1523–32.
12. Younis NN, Soran H, Pemberton P, Charlton-Menys V, Elseweidy MM, Durrington PN. Small dense LDL is more susceptible to glycation than more buoyant LDL in type 2 diabetes. *Clin Sci*. 2013;124(5):343–9.
13. Chen Y, Zhang XP, Yuan J, Cai B, Wang XL, Wu XL, Zhang YH, Zhang XY, Yin T, Zhu XH, et al. Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. *Bmj Open*. 2018;8(9):e21768.
14. Cai X, Gao J, Liu S, Wang M, Hu J, Hong J, Zhu Q, Tuerxun G, Dang Y, Li N. Hepatic Steatosis Index and the Risk of Type 2 Diabetes Mellitus in China: Insights from a General Population-Based Cohort Study. *Dis Markers*. 2022; 2022:3150380.
15. Cai X, Wang M, Liu S, Yuan Y, Hu J, Zhu Q, Hong J, Tuerxun G, Ma H, Li N. Establishment and validation of a nomogram that predicts the risk of type 2 diabetes in obese patients with non-alcoholic fatty liver disease: a longitudinal observational study. *Am J Transl Res*. 2022;14(7):4505–14.
16. Cai XT, Ji LW, Liu SS, Wang MR, Heizhati M, Li NF. Derivation and validation of a Prediction Model for Predicting the 5-Year incidence of type 2 diabetes in non-obese adults: a Population-based Cohort Study. *Diabet Metab Syndr Ob*. 2021;14:2087–101.
17. Cai X, Zhu Q, Cao Y, Liu S, Wang M, Wu T, Hong J, Ahmat A, Aierken X, Li N. A Prediction Model Based on Noninvasive Indicators to Predict the 8-Year Incidence of Type 2 Diabetes in Patients with Nonalcoholic Fatty Liver Disease: A Population-Based Retrospective Cohort Study. *Biomed Res Int*. 2021; 2021:5527460.
18. Sampson M, Wolska A, Warnick R, Lucero D, Remaley AT. A new equation based on the Standard Lipid Panel for calculating small dense low-density lipoprotein-cholesterol and its use as a risk-enhancer test. *Clin Chem*. 2011;67(7):987–97.
19. Endo K, Kobayashi R, Tanaka M, Tanaka M, Akiyama Y, Sato T, Hosaka I, Nakata K, Koyama M, Ohnishi H, et al. Validation of Estimated Small Dense Low-Density Lipoprotein Cholesterol Concentration in a Japanese General Population. *J Atheroscler Thromb*. 2024;31(6):931–52.
20. Huang H, Xie J, Hou L, Miao M, Xu L, Xu C. Estimated small dense low-density lipoprotein cholesterol and nonalcoholic fatty liver disease in nonobese populations. *J Diabetes Invest*. 2024;15(4):491–9.
21. Hirano T, Ito Y. Accuracy of small dense low-density lipoprotein-cholesterol concentration estimated via Sampson's equation in healthy subjects and patients with diabetes. *J Atheroscler Thromb*. 2023;30(8):979–89.
22. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, Collins BS, Hilliard ME, Isaacs D, Johnson EL, et al. 2. Classification and diagnosis of diabetes: standards of Care in Diabetes-2023. *Diabetes Care*. 2023;46(Suppl 1):S19–40.
23. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30(4):377–99.
24. Groenwold RH, White IR, Donders AR, Carpenter JR, Altman DG, Moons KG. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. *Can Med Assoc J*. 2012;184(11):1265–9.
25. Wu Y, Hu H, Cai J, Chen R, Zuo X, Cheng H, Yan D. Association of hypertension and incident diabetes in Chinese adults: a retrospective cohort study using propensity-score matching. *Bmc Endocr Disord*. 2021;21(1):87.
26. Cao C, Hu H, Zheng X, Zhang X, Wang Y, He Y. Association between central obesity and incident diabetes mellitus among Japanese: a retrospective cohort study using propensity score matching. *Sci Rep-Uk*. 2022;12(1):13445.
27. Li T, Cao C, Xuan X, Liu W, Xiao X, Wei C. The association between creatinine to body weight ratio and the risk of progression to diabetes from pre-diabetes: a 5-year cohort study in Chinese adults. *Bmc Endocr Disord*. 2023;23(1):266.
28. Santos HO, Earnest CP, Tinsley GM, Izidoro L, Macedo R. Small dense low-density lipoprotein-cholesterol (sdLDL-C): analysis, effects on cardiovascular endpoints and dietary strategies. *Prog Cardiovasc Dis*. 2020;63(4):503–9.
29. Ivanova EA, Myasoedova VA, Melnichenko AA, Grechko AV, Orekhov AN. Small Dense Low-Density Lipoprotein as Biomarker for Atherosclerotic Diseases. *Oxid Med Cell Longev* 2017, 2017:1273042.
30. Chen Y, Fu Y, Wang S, Chen P, Pei Y, Zhang J, Zhang R, Niu G, Gu F, Li X. Clinical significance of neutrophil gelatinase-associated lipocalin and sdLDL-C for coronary artery disease in patients with type 2 diabetes mellitus aged ≥ 65 years. *Cardiovasc Diabetol*. 2022;21(1):252.
31. Huang J, Gu JX, Bao HZ, Li SS, Yao XQ, Yang M, Li Y, Zhang AM, Yin Y, Zhang N et al. Elevated Serum Small Dense Low-Density Lipoprotein Cholesterol May Increase the Risk and Severity of Coronary Heart Disease and Predict Cardiovascular Events in Patients with Type 2 Diabetes Mellitus. *Dis Markers*. 2021; 2021:5597028.
32. Ma X, Wang Q, Hu X, Wang X, Zhao Y, Liu X, Li J, Du Y, Wang M, Qiu C et al. Association of sdLDL-C With Incident Carotid Plaques With Stable and Vulnerable Morphology: A Prospective Cohort Study *Stroke*. 2024; 55(3):576–585.
33. Kikkawa K, Nakajima K, Shimomura Y, Tokita Y, Machida T, Sumino H, Murakami M. Small dense LDL cholesterol measured by homogeneous assay in Japanese healthy controls, metabolic syndrome and diabetes patients with or without a fatty liver. *Clin Chim Acta*. 2015;438:70–9.
34. Hsu SH, Jang MH, Torng PL, Su TC. Positive Association between Small Dense Low-Density Lipoprotein Cholesterol Concentration and Biomarkers of Inflammation, thrombosis, and Prediabetes in non-diabetic adults. *J Atheroscler Thromb*. 2019;26(7):624–35.
35. Jui A, Jha K, Mondal H. Small dense low-density lipoprotein level in newly diagnosed type 2 diabetes Mellitus patients with normal low-density lipoprotein. *Cureus J Med Sci*. 2023;15(1):e33924.
36. Hirano T. [Small-Dense LDL, HDL2, 3]. *Rinsho Byori*. 2016;64(6):636–42.
37. Ban Y, Koba S, Tsunoda F, Yokota Y, Ezumi H, Kondo T, Suzuki H, Katagiri T. Predominance of small dense low-density lipoproteins and abnormal glucose regulation in patients with acute coronary syndrome. *Circ J*. 2006;70(4):393–401.
38. Goff DJ, D'Agostino RJ, Haffner SM, Otvos JD. Insulin resistance and adiposity influence lipoprotein size and subclass concentrations. Results from the Insulin Resistance Atherosclerosis Study. *Metabolism*. 2005; 54(2):264–270.
39. Wu L, Parhofer KG. Diabetic dyslipidemia. *Metabolism* 2014, 63(12):1469–1479.
40. Hirano T. Pathophysiology of Diabetic Dyslipidemia. *J Atheroscler Thromb*. 2018;25(9):771–82.
41. Haffner SM, Mykkanen L, Robbins D, Valdez R, Miettinen H, Howard BV, Stern MP, Bowsher R. A preponderance of small dense LDL is associated with specific insulin, proinsulin and the components of the insulin resistance syndrome in non-diabetic subjects. *Diabetologia*. 1995;38(11):1328–36.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.