scientific reports

Check for updates

OPEN

Association between excessive alcohol consumption and incident diabetes mellitus among Japanese based on propensity score matching

Changchun Cao[1,5], Cuimei Wei[2,5], Yong Han[3,5], Jiao Luo[1], Ping Xi[2], Jingying Chen[2], Xiaohua Xiao[2]✉, Haofei Hu[4]✉ & Dongli Qi[4]✉

The available evidence on the connection between excessive alcohol consumption and diabetes is controversial. Therefore, the primary objective of this investigation was to examine the connection between excessive alcohol consumption and incident diabetes in a Japanese population through the utilization of propensity score matching (PSM) analysis. Our retrospective cohort study encompassed a sample of 15,464 Japanese individuals who were initially free of diabetes between the years 2004 and 2015. The study utilized comprehensive medical records of individuals who underwent a physical examination. Employing a one:one PSM analysis, the current research included 2298 individuals with and without excessive alcohol consumption. Furthermore, a doubly robust estimation method was employed to ascertain the connection between excessive alcohol consumption and diabetes. The findings revealed that individuals with excessive alcohol consumption exhibited a 73% higher likelihood of developing diabetes (HR = 1.73, 95% CI 1.08–2.77). Furthermore, upon adjusting for variables, the PSM cohort demonstrated that individuals with excessive alcohol consumption had a 78% increased risk of developing diabetes in comparison to those with non-excessive alcohol consumption (HR = 1.78, 95% CI 1.08–2.93). Individuals with excessive alcohol consumption were found to have a 73% higher risk of developing diabetes compared to those with non-excessive alcohol consumption, even after controlling for propensity score (HR = 1.73, 95% CI 1.08–2.78). Participants in the PSM cohort with excessive alcohol consumption had a 73% higher risk of developing diabetes than those with non-excessive alcohol consumption after controlling for confounding factors. These findings underscore the importance of alcohol consumption guidelines aimed at reducing excessive drinking. Clinicians should be vigilant in screening for alcohol use in patients, particularly those at risk for diabetes, and provide appropriate counseling and resources to support alcohol reduction.

Keywords Excessive alcohol consumption, Diabetes mellitus, Propensity score, Propensity score matching, Inverse probability of treatment weights

Abbreviations

ALT	Alanine aminotransferase
AST	Aspartate aminotransferase;
BMI	Body mass index

[1]Department of Rehabilitation, Shenzhen Dapeng New District Nan'ao People's Hospital, Shenzhen 518000, Guangdong Province, China. [2]Department of Geriatrics, Futian District, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University, No.3002, Sungang West Road, Shenzhen 518000, Guangdong Province, China. [3]Department of Emergency, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University, Shenzhen 518000, Guangdong Province, China. [4]Department of Nephrology, Futian District, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University, No.3002, Sungang West Road, Shenzhen 518000, Guangdong Province, China. [5]These authors contributed equally: Changchun Cao, Cuimei Wei, and Yong Han. ✉email: tu_xi8888@163.com; huhaofei0319@126.com; 386514653@qq.com

CI	Confidence intervals
DBP	Diastolic blood pressure
DM	Diabetes mellitus
FPG	Fasting plasma glucose
GGT	Gamma-glutamyl transferase
HbA1c	Glycosylated hemoglobin
HDL-C	High-density lipoprotein cholesterol
HR	Hazard ratios
IPTW	Inverse probability of treatment weights
PS	Propensity score
PSM	Propensity score matching
Ref	Reference
SBP	Systolic blood pressure
SD	Standard deviation
TC	Total cholesterol
TG	Triglyceride

Diabetes mellitus (DM) is a chronic metabolic disorder characterized by hyperglycemia due to insulin resistance and relative insulin deficiency. The global incidence of DM has been rising, driven largely by increasing rates of obesity[1]. Obesity is a major risk factor for DM as it contributes to insulin resistance and beta-cell dysfunction[2]. Insulin resistance, a hallmark of DM, often precedes the onset of the disease and is exacerbated by excess adipose tissue[3]. The interplay between obesity and insulin resistance significantly elevates the risk of developing DM[4]. Global data from the International Diabetes Federation reveals that in 2019, approximately 463 million individuals aged 20–79 were diagnosed with DM, resulting in an incidence rate of 9.3%[5]. The burden of diabetes is particularly significant in Japan, where the prevalence has been steadily increasing. According to the National Health and Nutrition Survey Japan, the prevalence of diabetes among adults aged 20 and older has risen from 13.7% in 1997 to 20.3% in 2016[6]. This alarming trend underscores the need for focused public health strategies to address this escalating issue in the Japanese population. While DM is generally considered irreversible, preventive measures can play a pivotal role in mitigating its impact[7]. Numerous lifestyle-related risk factors, including excessive alcohol consumption, have been found to be significantly associated with the development of DM[8]. Specifically, heavy alcohol consumption has been shown to elevate the risk of hypertension, obesity, and insulin resistance[9–11], all of which are recognized as pivotal contributors to the onset of DM.

In recent years, the relationship between alcohol consumption and the risk of developing DM has garnered significant interest. Several meta-analyses and Mendelian randomized studies have shown that heavy alcohol consumption increases the risk of diabetes[12–14]. However, some studies have discovered inconsistent conclusions. A meta-analysis of 26 studies involving 706,716 participants showed that heavy drinking does not increase the risk of diabetes[15]. In addition, a negative association between alcohol intake and diabetes risk was recently revealed in the Atherosclerosis Risk in Communities Study[16]. Current research on the relationship between alcohol consumption and DM is still controversial, and the number of cohort studies is limited. Therefore, the association between alcohol consumption and DM still needs further study.

The propensity score (PS) quantifies the likelihood of experiencing a particular exposure, such as varying levels of alcohol intake, based on a predefined array of covariates. Employing the propensity score matching (PSM) technique provides significant benefits in studies facing the challenge of outcomes influenced by a plethora of covariates, where confounding factors might skew results, and when the feasibility of conducting randomized clinical trials is limited due to resource constraints and ethical issues. Traditional regression models, when faced with numerous confounders, may inadvertently introduce bias if they fail to account for all relevant variables or if there's lingering confounding. On the other hand, incorporating too many variables can lead to overfitting, obscuring the true relationship between the exposure and the outcome. Therefore, this research utilized PSM analysis to explore the authentic relationship between high levels of alcohol intake and the occurrence of diabetes mellitus among the Japanese populace.

Methods

Study design and data source

This research utilized open-source data from the NAGALA (NAFLD in Gifu Area, Longitudinal Analysis) database as a secondary analysis of a medical examination program. The center where these programs were conducted, established in 1994, performed over 8000 medical examinations annually, with 60% of participants undergoing one to two exams per year. Due to the high frequency of repeated examinations, the original study cohort comprised all participants who underwent repeated examinations between 2004 and 2015. Researchers can access the original study data from the Dryad Digital Repository (https://datadryad.org/). The dataset (https://doi.org/10.5061/dryad.8q0p192) comprises data for 15,464 participants free of DM at baseline. These data were employed for a secondary analysis in full compliance with Dryad's terms of service. We utilized data from this cohort, which was established to investigate the relationship between ectopic obesity and diabetes and has been previously published[17]. We have used their data to explore the relationship between excessive alcohol consumption and DM. The principal objective of this investigation was to assess the influence of excessive alcohol consumption as an independent variable on the incidence of diabetes mellitus, which was considered the dependent variable.

Study participants

In the initial study, informed written consent was obtained from all participants after receiving approval from the Clinical Research Ethics Committee at Murakami Memorial Hospital[17]. Additionally, the ethical endorsement of this study was obtained from the Ethics Committee of the Shenzhen Dapeng New District Nan'ao People's Hospital (2,022,082,201). Moreover, the study adhered to the principles outlined in the Declaration of Helsinki, and all procedures detailed in the Declarations section were conducted in accordance with applicable regulations and guidelines.

The initial sample for this study comprised 20,944 Japanese individuals who underwent physical examinations and participated in at least two examinations between 2004 and 2015[17]. Subjects meeting any of the following criteria were excluded from the study at baseline: (1) missing data for variables; (2) pre-existing liver diseases, such as hepatitis C or hepatitis B; (3) ethanol consumption over 60 g/day for men and 40 g/day for women; (4) medication usage; (5) diabetes; (6) fasting plasma glucose (FPG) levels ≥ 6.1 mmol/L. After excluding 5480 individuals, a total of 15,464 participants (7034 females and 8430 males) were included in the data analysis (Fig. 1).

Definition of excessive alcohol consumption

The evaluation of alcohol consumption included inquiries to participants about their alcohol consumption in the preceding month, followed by the calculation of their average weekly alcohol consumption. Excessive alcohol consumption was defined as consuming more than 30 g per day for males and more than 20 g per day for females[18–23].

Diagnosis of incident diabetes

DM was defined as fasting plasma glucose levels of ≥ 7 mmol/L, glycosylated hemoglobin levels of ≥ 6.5%[24], or self-reported during the follow-up period.

Covariates

This analysis, including covariates, was guided by clinical insights and findings from prior studies[25–29]. Based on the principles mentioned earlier, the following variables were used as covariates: gender, age, BMI, smoking status, regular exerciser, systolic blood pressure (SBP), diastolic blood pressure (DBP), high-density lipoprotein cholesterol (HDL-C), triglycerides (TG), total cholesterol (TC), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), glycosylated hemoglobin (HbA1c), and fasting plasma glucose (FPG). Data collection on lifestyle and medical history was facilitated through a standardized self-management questionnaire, while measurements of weight, height, and blood pressure were conducted by trained personnel. Laboratory tests were performed uniformly under controlled conditions to gather additional data.

Statistical analyses

Continuous variables adhering to a normal distribution were presented as mean ± standard deviation (SD), while those with a skewed distribution were depicted using the median (interquartile range). Frequencies and percentages were used to report categorical variables. Group differences were assessed using the Wilcoxon rank-sum test (skewed distribution), two-sample t-tests (normal distribution), and the $\chi 2$ test (categorical variables). Incidence rates were expressed in terms of person-years and cumulative incidence.

The present study employed propensity score analysis to ensure comparability between the excessive alcohol consumption and non-excessive alcohol consumption groups at baseline. This was achieved by creating a single group of subjects with similar baseline characteristics, as detailed in the flowchart in Fig. 1. As previously described, employing a comprehensive multivariate logistic regression model, wherein excessive alcohol consumption was designated as the independent variable and fifteen baseline variables served as covariates, the propensity score was meticulously derived[30]. A greedy-matching algorithm with a caliper width of 0.01 was used to build a one:one matching protocol without replacement. Standardized differences were used to assess group equilibrium, with a standardized difference below 10.0% indicating a relatively minor imbalance in a covariate[31]. Furthermore, the Kaplan–Meier method was utilized to calculate the likelihood of diabetes-free survival in each group, and statistical significance was assessed using the log-rank test. The Cox proportional hazards regression model was employed to investigate the relationship between excessive alcohol consumption and the incidence of diabetes within the propensity score-matched cohort. This study applied a doubly robust estimation method that combined the propensity score model and multivariate regression model to examine the association between excessive alcohol consumption and the risk of developing diabetes[32].

Propensity score matching may result in a reduced sample size, leading to unstable estimates and biased results, so we assessed the stability of our results by conducting sensitivity analyses. Inverse probability of treatment weighting (IPTW) helps to mitigate this issue by utilizing all available data, thus preserving the cohort size and enhancing the robustness of our findings. We used the estimated propensity score from the original cohort to compute the IPTW. Specifically, IPTW was calculated as one divided by the PS for individuals with excessive alcohol consumption and as one divided by (1 minus the PS) for individuals with non-excessive alcohol consumption. This approach allowed us to establish a weighted cohort[33]. Our study employed a range of sensitivity analyses to evaluate the robustness of the results and the impact of employing various models for associative inference on the conclusions. Two connection inference models were applied to the weighted and original populations in the sensitivity analysis. The resulting P-values and effect sizes were documented and compared across all models.

Prespecified subgroup analyses were conducted based on BMI and age. Continuous variables were converted to categorical variables using clinical cut-off points. Each stratification accounted for all factors, excluding the stratification factor itself. To maintain the balance of baseline characteristics between the excessive alcohol consumption and non-excessive alcohol consumption groups, only the corresponding matched pairs within

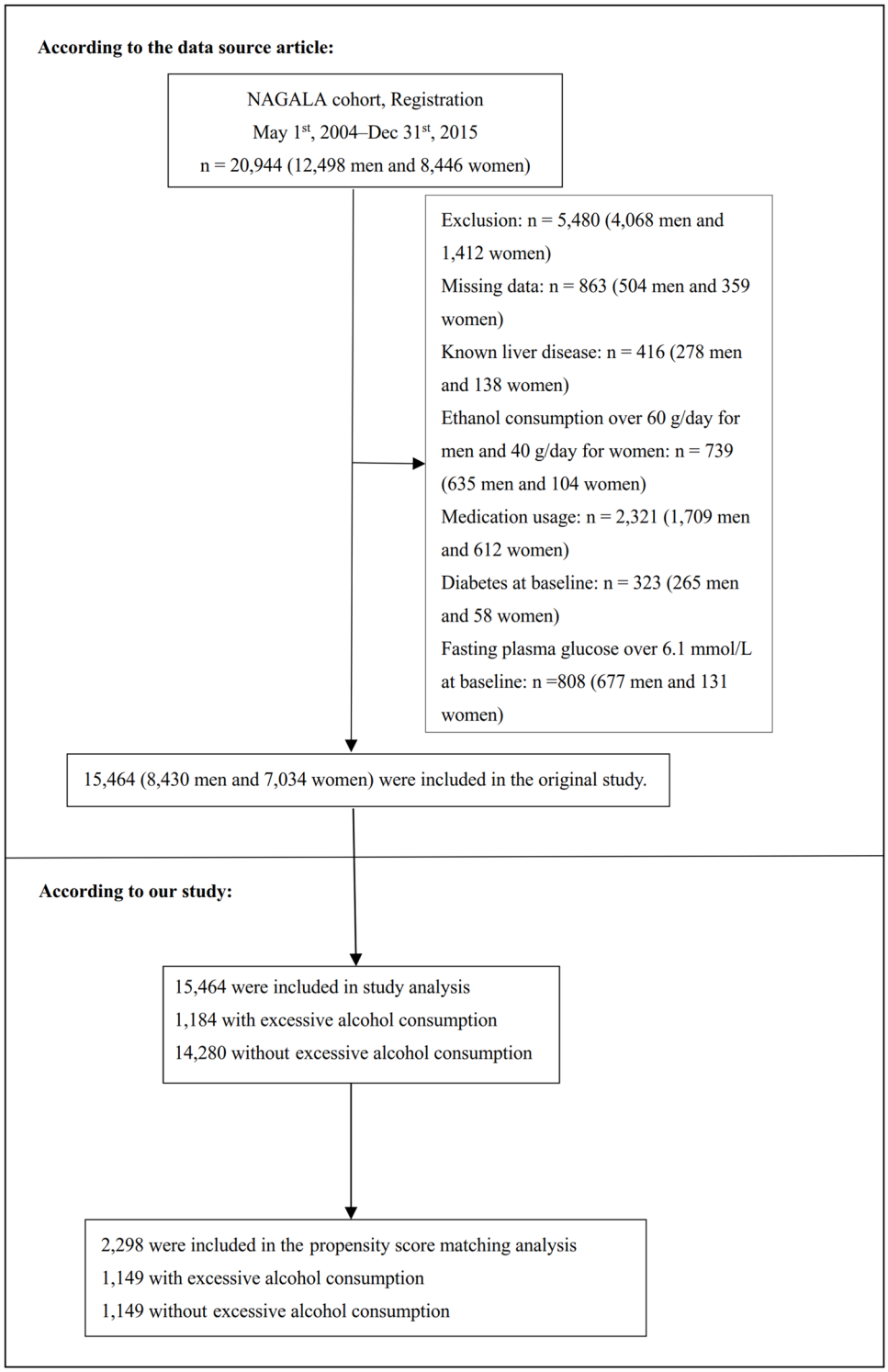

Figure 1. Study population.

the same subgroup were selected for the subgroup analyses. For instance, in the subgroup of participants with a BMI < 24 kg/m^2, only matched pairs from both the excessive alcohol consumption and non-excessive alcohol consumption groups within the BMI < 24 kg/m^2 category were included in the subgroup analysis. Likelihood ratio tests were employed to evaluate modifications and interactions within the subgroups. Our findings were presented according to the STROBE statement[34].

Statistical computations were carried out with Empower Stats (R) version 2.2 (http://www.empowerstats.com, X&Y Solutions, Inc., Boston, MA) and R software 3.6.1 (http://www.R-project.org, The R Foundation). A threshold of $P < 0.05$ in bidirectional tests established the benchmark for statistical significance.

Ethics approval and consent to participate

In the initial study, informed written consent was obtained from all participants after receiving approval from the Clinical Research Ethics Committee at Murakami Memorial Hospital. In addition, the study has also been approved by the Ethics Committee of the Shenzhen Dapeng New District Nan'ao People's Hospital (2022082201), and was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Results

Characteristics of participants

In our study, a total of 15,464 individuals were included, with 54.51% being male. In the original cohort, 4735 participants had alcohol-free consumption, and the mean alcohol consumption was 4.10 g per day in the non-excessive alcohol consumption group and 39.69 g per day in the excessive alcohol consumption group. After matching, 233 participants had no alcohol consumption, and the mean alcohol consumption was 8.23 g per day in the non-excessive alcohol consumption group and 32.47 g per day in the excessive drinking group. Among these individuals, 1184 (7.66%) reported excessive alcohol consumption. The study's participants had an average age of 43.71 ± 8.90 years. With the exception of TC, all baseline traits indicated in Table 1 that were compared between the excessive and non-excessive alcohol consumption groups showed statistically significant differences. Table 1 indicated that individuals engaging in excessive alcohol consumption typically presented elevated blood pressure, increased BMI, older age, and heightened levels of GGT, ALT, HDL-C, FPG, TG, and AST. Additionally, subjects with excessive alcohol consumption had lower HbA1c levels, a higher prevalence of males, current smokers, and individuals with regular exerciser. Through a one:one propensity score matching technique, we successfully matched 1149 individuals with excessive alcohol consumption to 1149 individuals who did not engage in excessive alcohol consumption. Following the implementation of PSM, the standardized differences for nearly all baseline characteristics were less than 10.0%, indicating a good match. Thus, the baseline characteristics of the non-excessive and excessive drinking groups differed only slightly.

Characteristic	Before matching				After matching			
	Non-excessive alcohol consumption	Excessive alcohol consumption	Standardized difference (100%)	*P*	Non-excessive alcohol consumption	Excessive alcohol consumption	Standardized difference (100%)	*P*
Participants	14,280	1184			1149	1149		
Age (years)	43.53 ± 8.89	45.83 ± 8.70	26.1	< 0.001	45.81 ± 9.19	45.76 ± 8.71	0.6	0.885
BMI (kg/m^2)	22.07 ± 3.14	22.71 ± 2.94	21.1	< 0.001	22.78 ± 3.08	22.72 ± 2.93	2.2	0.601
Gender			71.7	< 0.001			4.8	0.252
Male	7440 (52.10%)	990 (83.61%)			934 (81.3)	955 (83.1)		
Female	6840 (47.90%)	194 (16.39%)			215 (18.7)	194 (16.9)		
SBP (mmHg)	113.96 ± 14.83	120.97 ± 15.15	46.8	< 0.001	119.80 ± 16.41	120.78 ± 15.20	6.2	0.140
DBP (mmHg)	71.14 ± 10.39	76.88 ± 10.42	55.2	< 0.001	76.10 ± 11.52	76.72 ± 10.44	5.6	0.183
FPG (mg/dL)	92.74 ± 7.42	95.70 ± 7.12	40.8	< 0.001	95.46 ± 7.04	95.68 ± 7.14	3.1	0.458
HbA1c (%)	5.18 ± 0.32	5.10 ± 0.32	24.6	< 0.001	5.11 ± 0.31	5.10 ± 0.32	2.9	0.482
ALT (U/L)	16 (12–23)	20 (15–26)	21.1	< 0.001	19 (14–26)	20 (15–26)	2.7	0.521
AST (U/L)	17 (14–21)	19 (16–23)	27.2	< 0.001	18 (15–23)	19 (16–23)	1.3	0.753
GGT (U/L)	15 (11–21)	25 (17–40)	61.8	< 0.001	20 (14–33)	25 (17–38)	6.5	0.120
TC (mmol/L)	5.12 ± 0.87	5.15 ± 0.81	3.0	0.341	5.17 ± 0.85	5.14 ± 0.81	3.2	0.437
TG (mmol/L)	0.50 (0.30–0.89)	0.64 (0.37–1.13)	22.7	< 0.001	0.573 (0.346–0.965)	0.632 (0.374–1.120)	6.6	0.112
HDL-C (mmol/L)	1.46 ± 0.40	1.49 ± 0.43	7.2	0.014	1.49 ± 0.45	1.48 ± 0.42	1.0	0.822
Smoking status			82.6	< 0.001			11.8	0.019
Never smoker	8751 (61.28%)	280 (23.65%)			334 (29.069%)	279 (24.282%)		
Ever smoker	2572 (18.01%)	380 (32.09%)			325 (28.285%)	370 (32.202%)		
Current smoker	2957 (20.71%)	524 (44.26%)			490 (42.646%)	500 (43.516%)		
Regular exerciser			6.0	0.042			4.2	0.313
No	11,804 (82.661%)	951 (80.321%)			940 (81.810%)	921 (80.157%)		
Yes	2476 (17.339%)	233 (19.679%)			209 (18.190%)	228 (19.843%)		

Table 1. Baseline characteristics before and after propensity score matching. Values were n (%) or mean ± SDor median (25th to 75th percentiles). SD, standard deviation; BMI, body mass index; systolic blood pressure;DBP, diastolic blood pressure; FPG, fasting plasma glucose; HbA1c, glycosylated hemoglobin; ALT, alanineaminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; TC, total cholesterol;TG, triglyceride; HDL-C, high-density lipoprotein cholesterol.

Incidence rate of diabetes

Table 2 displayed diabetes mellitus incidence rates pre- and post-PSM. Following a median monitoring duration of 5.39 years, diabetes was diagnosed in 373 subjects. Before applying PSM, the prevalence rates for the overall population, non-excessive alcohol consumption group, and excessive alcohol consumption group were 2.41%, 2.27%, and 4.14%, respectively. The cumulative rates per 1000 person-years were 3.99, 3.75, and 6.18 for the total population, non-excessive alcohol consumption group, and excessive alcohol consumption group, respectively. Following propensity score matching, the prevalence rates of diabetes were as follows: 3.26% in the overall group, 2.35% in the non-excessive alcohol consumption group, and 4.18% in the excessive alcohol consumption group. Additionally, the cumulative rates of diabetes were found to be 5.42 per 1000 person-years in the overall group, 3.93 per 1000 person-years in the non-excessive alcohol consumption group, and 6.89 per 1000 person-years in the excessive alcohol consumption group. Kaplan–Meier analysis revealed a significant difference in the likelihood of diabetes-free survival between individuals with excessive alcohol consumption and those with non-excessive alcohol consumption before propensity score matching ($P = 0.001$; Fig. 2a). Following propensity score matching, this difference remained significant ($P = 0.021$; Fig. 2b).

The results of the relationship between excessive alcohol consumption and diabetes risk before and after propensity score matching.

Table 3 showcased the connection between excessive alcohol consumption and diabetes mellitus occurrence before and after propensity score matching. The crude model that did not adjust for any covariates revealed a notable link between excessive alcohol consumption and diabetes risk (HR = 1.79, 95% CI 1.33–2.42, $P = 0.0001$ before matching; HR = 1.73, 95% CI 1.08–2.77, $P = 0.0227$ after matching). This trend persisted in Model 1 with HR of 1.33 (95% CI 0.96–1.83) before matching and 2.02 (95% CI 1.25–3.29) after matching when we adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, and HDL-C. Further refinement (Model 2) incorporated additional variables, including age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, HDL-C, HbA1c, and FPG, maintaining the significant link (HR = 1.41, 95% CI 1.01–1.97, $P = 0.0433$ before matching; HR = 1.78, 95% CI 1.08–2.93, $P = 0.0242$ after matching). This association between excessive alcohol consumption and diabetes risk persisted, even after propensity score adjustments (HR = 1.60, 95% CI 1.14–2.24, $P = 0.0064$ before matching; HR = 1.73, 95% CI 1.08–2.78, $P = 0.0223$ after matching).

Sensitivity analysis

A sensitivity test was conducted to ascertain the stability of the link between excessive alcohol consumption and diabetes risk, as detailed in Table 4. These analyses encompassed weighted cohorts. Our findings uniformly demonstrated a significant association between excessive alcohol consumption and diabetes onset. When accounting for confounding factors, we observed an increased diabetes risk of a 77% elevation (HR = 1.77, 95% CI 1.53–2.04, $P < 0.001$) in the weighted group.

The results of the subgroup analysis

We used a subgroup analysis to detect the effect of potential confounders, which might affect the relationship between excessive alcohol consumption and the risk of diabetes. We treated age and BMI as the stratification variables to evaluate the trend of effect sizes in these variables. Table 5 showed that the association between excessive alcohol consumption and diabetes was notably stronger in individuals with BMI < 24 kg/m^2.

Discussion

Findings from the PSM cohort study reveal an independent correlation between heavy alcohol intake and diabetes occurrence. Post-adjustment for the PS, individuals engaging in excessive alcohol use demonstrated a 73% increased likelihood of diabetes onset. Sensitivity analyses across both weighted and original cohorts uniformly underscored a strong link between heavy drinking and diabetes, thereby bolstering the credibility of our results. Furthermore, in subgroup analysis, we observed that the association between excessive alcohol consumption and diabetes was notably stronger in individuals with BMI < 24 kg/m^2.

Variable	Participants (n)	DM events (n)	Cumulative incidence (95%CI)	Per 1000 person-year
Before Matching				
Total	15,464	373	2.41 (2.17–2.66)	3.99
Non-excessive alcohol consumption	14,280	324	2.27 (2.02–2.51)	3.75
Excessive alcohol consumption	1184	49	4.14 (3.00–5.27)	6.18
After matching				
Total	2298	75	3.26 (2.54–3.99)	5.42
Non-excessive alcohol consumption	1149	27	2.35 (1.47–3.23)	3.93
Excessive alcohol consumption	1149	48	4.18 (3.02–5.34)	6.89

Table 2. Incidence rate of incident diabetes before and after propensity score matching. CI, confidence interval; DM, diabetes mellitus.

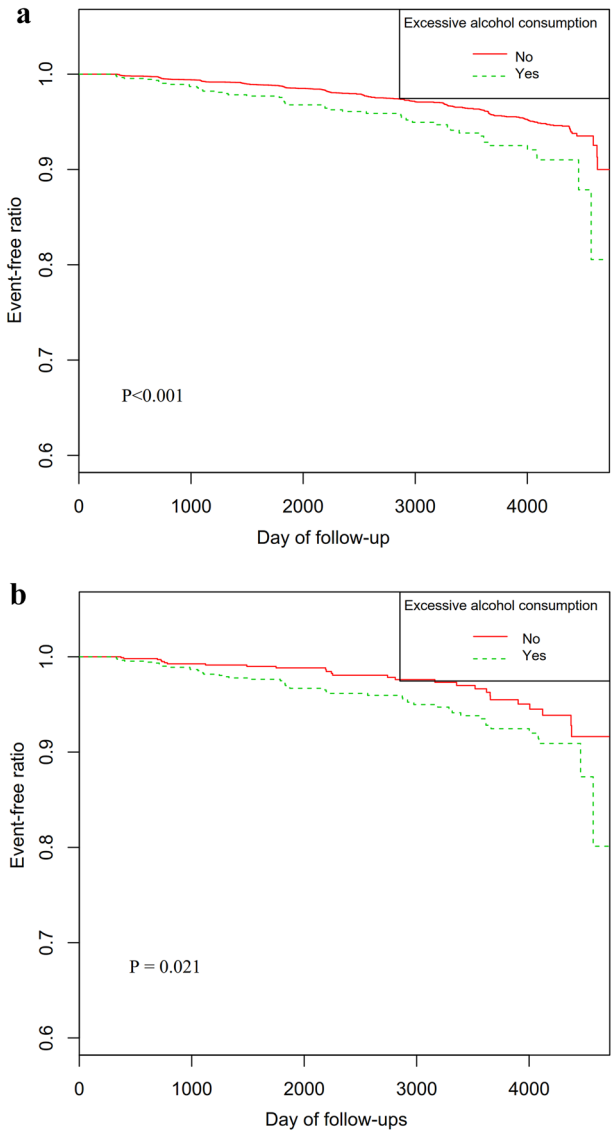

Figure 2. Kaplan–Meier event-free survival curve based on Excessive alcohol consumption and Non-Excessive alcohol consumption. (**a**) Kaplan–Meier event-free survival curve based on Excessive alcohol consumption and Non-Excessive alcohol consumption in the original cohort. (**b**) Kaplan–Meier event-free survival curve based on Excessive alcohol consumption and Non-Excessive alcohol consumption in the propensity score matching cohort.

Variable	Crude model (HR, 95% CI, P)	Model 1 (HR, 95% CI, P)	Model 2 (HR, 95% CI, P)	Model 3 (HR, 95% CI, P)
Before matching				
Non-excessive alcohol consumption	Ref	Ref	Ref	Ref
Excessive alcohol consumption	1.79 (1.33, 2.42) 0.0001	1.33 (0.96, 1.83) 0.0882	1.41 (1.01, 1.97) 0.0433	1.60 (1.14, 2.24) 0.0064
After matching				
Non-excessive alcohol consumption	Ref	Ref	Ref	Ref
Excessive alcohol consumption	1.73 (1.08, 2.77) 0.0227	2.02 (1.25, 3.29) 0.0043	1.78 (1.08, 2.93) 0.0242	1.73 (1.08, 2.78) 0.0223

Table 3. Association between excessive alcohol consumption and incident diabetes before and after propensity score matching. Crude model: we did not adjust for other covariates. Model 1: we adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, and HDL-C. Model 2: we adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, HDL-C, HbA1c, and FPG. Model 3: we adjusted for propensity score. HR, Hazard ratios; CI, Confidence interval; Ref, Reference.

Variable	Crude model (HR, 95% CI, P)	Model 4 (HR,95%CI, P)	Model 5 (HR,95%CI, P)
Non-excessive alcohol consumption	Ref	Ref	Ref
Excessive alcohol consumption	1.85 (1.62, 2.11) < 0.0001	2.02 (1.76, 2.32) < 0.0001	1.77 (1.53, 2.04) < 0.0001

Table 4. Association between excessive alcohol consumption and incident diabetes in a weighted cohort model. Crude model: we did not adjust for other covariates. Model 4: we adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, and HDL-C. Model 5: we adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, HDL-C, HbA1c, and FPG. HR, Hazard ratios; CI, Confidence interval; Ref, Reference.

Characteristic	No of participants	HR (95%CI)	*P* value	*P* for interaction
BMI				0.0022
< 24kg/m^2	1090	6.74 (1.94, 23.45)	0.0027	
≥ 24kg/m^2	232	0.50 (0.11, 2.30)	0.3764	
Age				0.3613
< 45 years	534	0.49 (0.07, 3.55)	0.4840	
≥ 45 years	670	3.14 (1.16, 8.53)	0.0249	

Table 5. Effect size of excessive alcohol consumption on incident diabetes in prespecified and exploratory subgroups. Note 1: The above model has been adjusted for age and BMI. Note 2: In each case, the model was adjusted for age, BMI, gender, smoking status, regular exerciser, SBP, DBP, ALT, AST, GGT, TC, TG, HDL-C, HbA1c, and FPG, but not adjusted for the stratification variable.

Alcohol consumption is one of the key risk factors contributing to the global burden of disease through its harmful effects on health, including cardiovascular disease, cirrhosis of the liver, neuropsychiatric disorders, and cancer[35–37]. Numerous studies have demonstrated that heavy alcohol consumption significantly increases the incidence of DM[12–14]. For instance, a meta-analysis of 38 observational studies, encompassing over 1.9 million participants, indicated an elevated diabetes risk associated with heavy drinking[12]. Nevertheless, any potential risk reduction in moderate drinkers may be limited to specific populations, such as women and non-Asians[12]. A Mendelian randomized study including 72,299 middle-aged and elderly volunteers found that heavy alcohol consumption was strongly associated with the risk of diabetes mellitus (HR = 1.295, 95% CI 1.059–1.583), after adjusting for age, residence areas, gender, BMI, hypertension, hyperlipidemia, education, daily activity, energy intake, and smoking status[13]. Furthermore, a Mendelian randomization study involving 408,540 European participants revealed a 1.10-fold increased odds of DM (95% CI 1.06–1.13) among those consuming more than 14 drinks per week[14]. Notably, these findings align with our own results. In our analysis, employing the doubly robust estimation method in the PSM cohort, we observed a significant 73% increase in the risk of incident diabetes associated with excessive alcohol consumption. Importantly, this risk remained unchanged even after adjusting for propensity scores.

Our research offers an in-depth analysis of the link between excessive alcohol consumption and diabetes risk within real-world contexts. We conducted meticulous adjustments for various covariates, including sociodemographic factors, lifestyle habits, and clinical parameters. Furthermore, our study benefits from a substantial sample size. The findings highlight a significant association between excessive alcohol consumption and the incidence of diabetes. Our results suggest that people with excessive alcohol consumption have a high risk of getting diabetes in the future. Additionally, our study contributes to the advancement of propensity score methods in correlation studies.

However, it is important to note that some researchers have reported contrasting findings regarding the association between excessive alcohol consumption and the risk of diabetes. For example, in a meta-analysis incorporating 706,716 participants (275,711 males and 431,005 females) from 26 studies, the findings indicated that, in comparison to the minimal category of alcohol consumption, the relative risk for diabetes in the heavy alcohol consumption category was 0.95 (95% CI 0.83, 1.09; $P = 0.480$)[15]. Similarly, another prospective analysis, spanning a median follow-up period of 21 years and involving 12,042 study participants, revealed intriguing results[16]. Over the course of this extended observation, 3795 participants ultimately developed diabetes[16]. Surprisingly, this prospective study identified a negative correlation between alcohol intake and the risk of diabetes, suggesting that moderate to high alcohol consumption could act as a protective factor against the development of diabetes[16]. Interestingly, among participants categorized as heavy alcohol consumers, an interesting observation was made: when they reduced their alcohol intake during follow-up, their risk of developing diabetes appeared to increase[16].

The inconsistent findings in the existing literature regarding the relationship between excessive alcohol consumption and diabetes can be attributed to several factors. First, study participants exhibit diversity in terms of race, gender, nationality, age, and other demographic characteristics. Second, different studies may choose to include or exclude certain covariates, leading to variations in the results. Third, variations in follow-up time across studies could also contribute to inconsistencies. Shorter follow-up periods may not capture the long-term effects of alcohol consumption on insulin resistance and impaired beta-cell function, whereas longer follow-up

periods may introduce additional confounding factors and lead to changes in confounders. Fourth, different approaches have been employed to address confounding factors. Our study aligns with the prevailing evidence in the literature, supporting the hypothesis that excessive alcohol consumption is associated with an increased risk of diabetes. This underscores the significance of interventions aimed at reducing alcohol consumption as a preventive measure for diabetes.

The exact pathway through which alcohol influences the onset of diabetes remains to be fully elucidated. It is hypothesized that the association between alcohol-induced diabetes risk involves mechanisms such as insulin resistance and compromised β-cell functionality. Direct pancreatic damage from excessive alcohol intake has been posited[38]. Evidence from animal research indicates that high levels of ethanol consumption impair islets and β-cells, disrupting insulin pathways, diminishing β-cell mass, and thus reducing insulin output while elevating fasting glucose concentrations[39,40]. Ethanol exposure is also thought to induce β-cell apoptosis via mitochondrial dysfunction, marked by an uptick in reactive oxygen species and a drop in adenosine triphosphate production[41]. Additionally, ethanol might impede the insulin-stimulated activation of phosphatidylinositol 3-kinase and the expression of glucose transporter four in skeletal muscle, further leading to insulin resistance[39]. Our research corroborates the theory that excessive alcohol intake significantly elevates the risk of diabetes development.

Study strengths and limitations

Our study exhibits several notable strengths. Firstly, it represents the first investigation to employ Propensity Score Matching in assessing the relationship between excessive alcohol consumption and the risk of diabetes. PSM is a robust methodology that addresses diverse data requirements by minimizing inter-group disparities, balancing confounding variables, and achieving an effect comparable to randomization. Secondly, our study reinforces the robustness of its findings through a range of sensitivity analyses. By using Inverse Probability of Treatment Weights, we established a weighted cohort to further evaluate the association between excessive alcohol consumption and diabetes incidence. This approach provides an additional means to assess the performance of traditional regression models.

However, it's important to acknowledge the limitations of this research. Firstly, the study exclusively involved individuals of Japanese ancestry, which may restrict the generalizability of our findings to other ethnic groups. Further research in diverse populations is necessary to validate the broader applicability of our results. Secondly, the raw data lacked information from the 2-h oral glucose tolerance test, and participants with FPG levels ≥ 6.1 mmol/L were not included. As a result, our study may underestimate the actual relationship between excessive alcohol consumption and diabetes risk. In future research, we intend to design a study that includes participants with FPG levels ≥ 6.1 mmol/L and incorporates a 2-h oral glucose tolerance test to explore the relationship between excessive alcohol consumption and diabetes more accurately. Third, in the non-excessive alcohol consumption group, 233 participants had no alcohol consumption. In the non-excessive alcohol consumption group, participants may have had a sick-quitter effect bias, which may have affected our results' stability. In future studies, we will exclude participants who did not consume alcohol to minimize this bias further. Lastly, while the PSM method helps ensure the balance of measured covariates, it cannot address potential imbalances in unmeasured confounding variables, such as family history of diabetes or dietary habits. To minimize the potential impact of these unmeasured variables on our results, we adopted a caliper width of 0.01.

Conclusions

Based on the findings of this retrospective cohort study employing PSM, a substantial correlation exists between excessive alcohol consumption and diabetes. Consequently, this study offers clinical evidence that can be referenced for managing the risk of diabetes by controlling excessive alcohol consumption.

Data availability

The raw data can be downloaded from the 'DATADRYAD' database (www.Datadryad.org). Dryad Digital Repository. https://datadryad.org/stash/dataset/doi:10.5061%2Fdryad.8q0p192.

Received: 7 February 2024; Accepted: 22 July 2024
Published online: 27 July 2024

References

1. Cao, C. *et al.* Association between central obesity and incident diabetes mellitus among Japanese: A retrospective cohort study using propensity score matching. *Sci. Rep.* **12**(1), 13445 (2022).
2. Kahn, S. E., Hull, R. L. & Utzschneider, K. M. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* **444**(7121), 840–846 (2006).
3. Shulman, G. I. Cellular mechanisms of insulin resistance. *J Clin. Invest.* **106**(2), 171–176 (2000).
4. Reaven, G. M. Insulin resistance: The link between obesity and cardiovascular disease. *Med. Clin. North Am.* **95**(5), 875–892 (2011).
5. 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.* **157**, 107843 (2019).
6. Ministry of Health, Labour and Welfare, Japan. (2017). *National Health and Nutrition Survey Japan. 2020.*
7. Uusitupa, M. *et al.* Prevention of Type 2 diabetes by lifestyle changes: A systematic review and meta-analysis. *Nutrients* **11**(11), 2611 (2019).
8. Zhang, Y. *et al.* Combined lifestyle factors and risk of incident type 2 diabetes and prognosis among individuals with type 2 diabetes: A systematic review and meta-analysis of prospective cohort studies. *Diabetologia* **63**(1), 21–33 (2020).
9. Mukamal, K. J. *et al.* Alcohol and risk for ischemic stroke in men: the role of drinking patterns and usual beverage. *Ann. Intern. Med.* **142**(1), 11–19 (2005).
10. Hwang, S. & Choi, J. W. Association between excessive alcohol consumption and hypertension control in hypertensive patients. *Chronic Illn.* **19**(3), 625–634 (2023).

11. Tatsumi, Y. *et al.* Association between alcohol consumption and incidence of impaired insulin secretion and insulin resistance in Japanese: The Saku study. *Diabetes Res. Clin. Pract.* **135**, 11–17 (2018).
12. Knott, C., Bell, S. & Britton, A. Alcohol consumption and the risk of Type 2 diabetes: A systematic review and dose-response meta-analysis of more than 1.9 million individuals from 38 observational studies. *Diabetes Care* **38**(9), 1804–1812 (2015).
13. Liu, M. & Park, S. A causal relationship between alcohol intake and type 2 diabetes mellitus: A two-sample Mendelian randomization study. *Nutr. Metab. Cardiovasc. Dis.* **32**(12), 2865–2876 (2022).
14. Lu T, Nakanishi T, Yoshiji S, Butler-Laporte G, Greenwood C, Richards JB: Dose-dependent association of alcohol consumption with obesity and type 2 diabetes: mendelian randomization analyses. *J. Clin. Endocrinol. Metab.* 2023.
15. Li, X. H., Yu, F. F., Zhou, Y. H. & He, J. Association between alcohol consumption and the risk of incident type 2 diabetes: a systematic review and dose-response meta-analysis. *Am. J. Clin. Nutr.* **103**(3), 818–829 (2016).
16. He, X., Rebholz, C. M., Daya, N., Lazo, M. & Selvin, E. Alcohol consumption and incident diabetes: The Atherosclerosis Risk in Communities (ARIC) study. *Diabetologia* **62**(5), 770–778 (2019).
17. 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.)* **43**(1), 139–148 (2019).
18. Schmidt, M. I. *et al.* Cohort profile: Longitudinal study of adult health (ELSA-Brasil). *Int. J. Epidemiol.* **44**(1), 68–75 (2015).
19. Nath, P. *et al.* Effect of nonalcoholic fatty liver disease (NAFLD) on COVID-19: A single-center study of 3983 patients with review of literature. *Cureus* **14**(7), e26683 (2022).
20. Zheng, X. *et al.* Association between nonalcoholic fatty liver disease and incident diabetes mellitus among Japanese: A retrospective cohort study using propensity score matching. *Lipids Health Dis.* **20**(1), 59 (2021).
21. Karnikowski, M., Cordova, C., Oliveira, R. J., Karnikowski, M. G. & Nobrega, O. T. Non-alcoholic fatty liver disease and metabolic syndrome in Brazilian middle-aged and older adults. *Sao Paulo Med. J.* **125**(6), 333–337 (2007).
22. Yoon, Y. S., Oh, S. W., Baik, H. W., Park, H. S. & Kim, W. Y. Alcohol consumption and the metabolic syndrome in Korean adults: the 1998 Korean National Health and Nutrition Examination Survey. *Am. J. Clin. Nutr.* **80**(1), 217–224 (2004).
23. Alkerwi, A. *et al.* Alcohol consumption and the prevalence of metabolic syndrome: A meta-analysis of observational studies. *Atherosclerosis* **204**(2), 624–635 (2009).
24. Classification and Diagnosis of Diabetes. Standards of medical care in diabetes-2021. *Diabetes Care* **44**(Suppl 1), S15–S33 (2021).
25. Polsky, S. & Akturk, H. K. Alcohol consumption, diabetes risk, and cardiovascular disease within diabetes. *Curr. Diab. Rep.* **17**(12), 136 (2017).
26. Llamosas-Falcon, L. *et al.* The relationship between alcohol consumption, BMI, and Type 2 diabetes: A systematic review and dose-response meta-analysis. *Diabetes Care* **46**(11), 2076–2083 (2023).
27. Cao, C., Han, Y., Hu, H., He, Y. & Luo, J. Non-linear relationship between pulse pressure and the risk of pre-diabetes: A secondary retrospective Chinese cohort study. *BMJ Open* **14**(3), e80018 (2024).
28. Cao, C. *et al.* The nonlinear correlation between alanine aminotransferase to high-density lipoprotein cholesterol ratio and the risk of diabetes: A historical Japanese cohort study. *BMC Endocr. Disord.* **23**(1), 124 (2023).
29. Cao, C. *et al.* Non-linear relationship between high-density lipoprotein cholesterol and incident diabetes mellitus: A secondary retrospective analysis based on a Japanese cohort study. *BMC Endocr. Disord.* **22**(1), 163 (2022).
30. Ahmed, A. *et al.* Outcomes in ambulatory chronic systolic and diastolic heart failure: A propensity score analysis. *Am Heart. J.* **152**(5), 956–966 (2006).
31. Austin, P. C. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat. Med.* **28**(25), 3083–3107 (2009).
32. Koch, B., Vock, D. M. & Wolfson, J. Covariate selection with group lasso and doubly robust estimation of causal effects. *Biometrics* **74**(1), 8–17 (2018).
33. Austin, P. C. & Stuart, E. A. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* **34**(28), 3661–3679 (2015).
34. von Elm, E. *et al.* The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Int. J. Surg.* **12**(12), 1495–1499 (2014).
35. Forouzanfar, M. H. *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386**(10010), 2287–2323 (2015).
36. Shield, K. D., Parry, C. & Rehm, J. Chronic diseases and conditions related to alcohol use. *Alcohol. Res.* **35**(2), 155–173 (2013).
37. Connor, J. Alcohol consumption as a cause of cancer. *Addiction* **112**(2), 222–228 (2017).
38. Federico, A. *et al.* The effects of alcohol on gastrointestinal tract, liver and pancreas: Evidence-based suggestions for clinical management. *Eur. Rev. Med. Pharmacol. Sci.* **19**(10), 1922–1940 (2015).
39. Zhao, L. N. *et al.* The diabetogenic effects of excessive ethanol: Reducing beta-cell mass, decreasing phosphatidylinositol 3-kinase activity and GLUT-4 expression in rats. *Br. J. Nutr.* **101**(10), 1467–1473 (2009).
40. Wang, S. *et al.* Ethanol induced impairment of glucose metabolism involves alterations of GABAergic signaling in pancreatic beta-cells. *Toxicology* **326**, 44–52 (2014).
41. Dembele, K., Nguyen, K. H., Hernandez, T. A. & Nyomba, B. L. Effects of ethanol on pancreatic beta-cell death: interaction with glucose and fatty acids. *Cell Biol Toxicol* **25**(2), 141–152 (2009).

Author contributions

C.C., C.W., and Y.H. contributed to the study concept and design, researched and interpreted the data, and drafted the manuscript. X.X., H.H., D.Q., J.L., P.X., and J.C. analyzed the data and reviewed the manuscript. C.C., C.W., and Y.H. oversaw the project's progress, contributed to the discussion, and reviewed the manuscript. X.X., H.H., and D.Q. are the guarantors of this work and, as such, had full access to all the data in the study and took responsibility for the data's integrity and the data analysis's accuracy. All authors read and approved the final manuscript.

Funding

This study was supported by Shenzhen Municipal Science and Technology Innovation Commission for Sustainable Development (KCXFZ20201221173213036), Shenzhen High-level Hospital Construction Fund, Sanming Project of Medicine in Shenzhen (No. SZSN202311007).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.X., H.H. or D.Q.