

To Explore the Causal Link between Genetic Predisposition to Alcohol Consumption and Type 2 Diabetes Mellitus

Qingxiang Zeng, Minghua Dong*

Key Laboratory of Prevention and Treatment of Cardiovascular and Cerebrovascular Diseases, Ministry of Education, Gannan Medical University, Ganzhou, China
Email: *dmhua2000@qq.com

How to cite this paper: Zeng, Q.X. and Dong, M.H. (2024) To Explore the Causal Link between Genetic Predisposition to Alcohol Consumption and Type 2 Diabetes Mellitus. *Journal of Biosciences and Medicines*, 12, 538-549.

<https://doi.org/10.4236/jbm.2024.1212040>

Received: November 21, 2024

Accepted: December 28, 2024

Published: December 31, 2024

Abstract

Type 2 diabetes is a chronic metabolic disorder, and the behavior and quantity of alcohol consumption may be influenced by genetic factors. Moderate alcohol consumption has been associated with a reduced risk of type 2 diabetes, whereas excessive drinking may increase the risk. Alcohol consumption might indirectly contribute to the development of type 2 diabetes through factors such as body mass index (BMI). This paper explores the complex relationship between alcohol consumption and type 2 diabetes, investigating the causal association using Mendelian randomization (MR) methodology. Individual genetic background significantly impacts drinking habits, with certain genetic variants being linked to higher alcohol consumption tendencies, which also influence the risk of type 2 diabetes. The study employed MR methods, utilizing genetic variants as instrumental variables to assess the causal relationship between alcohol consumption and type 2 diabetes. Data from the European population samples obtained from the IEU Open GWAS Project database were used, including summary data from genome-wide association studies (GWAS) related to alcohol consumption and type 2 diabetes. The inverse variance weighted (IVW) method was applied to estimate the causal effect. The IVW results indicated no statistically significant association between alcohol consumption and increased risk of type 2 diabetes (OR = 1.0046, 95% CI: [0.8722, 1.1571], P = 0.9495), suggesting no significant causal relationship between genetically predicted alcohol consumption and type 2 diabetes risk. These findings contribute to future research and development in the prevention and management of type 2 diabetes.

Keywords

Type 2 Diabetes, Capacity for Alcohol, Mendelian Randomization

1. Introduction

Diabetes has become a serious global public health issue that significantly affects human health [1]. Diabetes is a group of chronic metabolic diseases caused by insufficient insulin secretion or defects in its action [2]. According to the 10th edition of the International Diabetes Federation's Global Diabetes Map published in 2021, 537 million adults (aged 20 - 79 years) worldwide had diabetes in 2021, with one in ten people affected. It is projected that this number will rise to 643 million by 2030 and 783 million by 2045 [3]. Increasing evidence suggests a complex association between alcohol consumption and type 2 diabetes. While some studies propose that harmful drinking may be associated with an increased risk of type 2 diabetes, this association is not linear and may be influenced by individual genetic background, lifestyle, and other environmental factors [4]. This study aims to understand the relationship between alcohol consumption and type 2 diabetes using Mendelian Randomization (MR).

1.1. Alcohol Consumption and Type 2 Diabetes

The relationship between alcohol consumption and type 2 diabetes has received considerable attention. Wang Yanhuan *et al.* found in their prospective study "Alcohol Consumption and Risk of Type 2 Diabetes Among Adult Males in Deqing County, Zhejiang Province" that individuals who consume large amounts of alcohol in a single session and drink frequently have an increased risk of developing type 2 diabetes. In contrast, those who consume smaller amounts of alcohol in multiple sessions may have a lower risk [5]. Xu Li *et al.* further confirmed these findings in their review "Research Progress and Prospects on Moderate Alcohol Consumption and Diabetes Prevention," which also showed that moderate alcohol consumption can reduce the incidence of diabetes, while excessive drinking increases the risk [6]. Zhu Yali *et al.* mentioned in their cross-sectional study "Association Between Alcohol Consumption and Pre-diabetes and Diabetes in Bao'an District, Yan'an City" that light drinking ($0.1 \text{ g/day} < \text{ethanol intake} < 19.9 \text{ g/day}$) is a protective factor against pre-diabetes and diabetes, whereas heavy drinking ($\text{ethanol intake} \geq 20 \text{ g/day}$) is a risk factor [7]. These findings suggest that alcohol consumption is an important factor in the development of type 2 diabetes.

1.2. Genetic Predisposition and Drinking Behavior

Genetic studies have shown that an individual's genetic makeup can influence their drinking behavior, including frequency and quantity. Foreign research has indicated that the expression of alcohol-related genes, such as initial reactions to alcohol and maximum daily drinking episodes, is genetically influenced. For example, rs1229984 in ADH1B (Alcohol Dehydrogenase 1B) and rs671 in ALDH2 (Aldehyde Dehydrogenase 2) can prevent heavy drinking (Kimura and Higuchi 2011) [8]. SNPs such as rs1229984 in ADH1B, rs13130794 in KLB, rs144198753 in BTF3P13, rs1260326 in GCKR, rs13107325 in SLC39A8, and rs11214609 in DRD2 are significantly associated with drinking behavior [9]. Therefore, genetic

predispositions related to drinking can influence drinking behavior.

1.3. Mendelian Randomization Studies

To better understand the causal relationship between alcohol consumption and type 2 diabetes, we adopted the statistical method of Mendelian Randomization (MR). MR uses genetic variants as instrumental variables to evaluate the causal relationship between exposure factors (such as alcohol consumption) and health outcomes (such as type 2 diabetes). The core idea of MR is to utilize the random allocation characteristics of genetic variants, similar to randomization in randomized controlled trials, to overcome potential confounding and reverse causality [10].

1.4. Research Objectives and Significance

Given the above background, this study aims to further explore the causal link between genetic predispositions for alcohol consumption and type 2 diabetes using Mendelian Randomization. Specifically, we will assess how different levels of genetic predisposition to alcohol consumption affect the risk of developing type 2 diabetes and explore potential biological mechanisms to provide new insights into the prevention and management of type 2 diabetes. Understanding the causal relationship between genetic predispositions for alcohol consumption and type 2 diabetes not only helps reveal disease mechanisms but also provides scientific evidence for targeted prevention strategies. Additionally, this study will guide clinical practice, particularly in designing personalized medical plans, to reduce the incidence of type 2 diabetes and improve patients' quality of life.

2. Methods

2.1. Data Sources and Basic Information

Genome-wide association studies (GWAS) summary data related to alcohol consumption were obtained from the IEUOpenGWASProject database for European populations. The IEUOpenGWASProject, developed by the Medical Research Council Integrative Epidemiology Unit (MRCIEU) at the University of Bristol, is a comprehensive resource of manually curated GWAS summary datasets available for download or query. By searching for alcohol-related traits, we obtained a sample from 2017 with 112,117 participants and 12,935,395 SNPs. GWAS summary data related to type 2 diabetes were also obtained from the IEUOpenGWASProject database for European populations. By searching for type 2 diabetes traits, we obtained a sample from 2021 with 490,089 participants and 24,167,560 SNPs. Since all the data were extracted from publicly available summary statistics, ethical approval was not required.

2.2. Data Extraction and Preprocessing

In this study, to establish causality, the Two-Sample Mendelian Randomization (TSMR) method was employed utilizing standardized Genome-Wide Association

Studies (GWAS) summary statistics. The GWAS datasets used encompass essential information on exposure and outcome variables, specifically including several key elements: Single Nucleotide Polymorphisms (SNPs), the chromosome number where the SNP resides (Chromosome, chr), the precise location on the chromosome (Position, pos), the effect size (Effect Size, β), which denotes the impact of each allele on the outcome variable, expressed in terms of standard error. Notably, in this study, the effect sizes are presented as Odds Ratios (OR) and converted to their natural logarithmic form ($\log(\text{OR})$) following conventional practices for subsequent analyses. Furthermore, the dataset includes important parameters such as the standard error of the effect size (Standard Error, SE), the effect allele (Effect Allele, EA) associated with the SNP, and the significance level (P-value) of the risk allele in the exposure factor. Collectively, these data serve as the foundation for causal inference in this study.

2.3. Selection of Instrumental Variables

To ensure the validity and applicability of the selected instrumental variables in this study, single nucleotide polymorphisms (SNPs) that meet the fundamental criteria for instrumental variables were initially screened from the dataset of the exposure factor. The screening process followed the steps outlined below:

1) Initial Screening of Significant SNPs: Candidate SNPs were preliminarily identified from the exposure factor based on a threshold of $P < 5 \times 10^{-8}$. Considering the relatively smaller sample size for alcohol consumption, which limits the number of available instrumental variables, an additional criterion of $P < 5 \times 10^{-5}$ was adopted to supplement the pool of candidate SNPs, thereby increasing the number of potential instruments.

2) Removal of Linkage Disequilibrium (LD) SNPs: Linkage disequilibrium (LD) refers to the non-random association of alleles at different loci on a chromosome, indicating that these alleles tend to be inherited together more often than would be expected by chance. To avoid potential biases introduced by linkage disequilibrium, the PLINK clumping algorithm was applied to the GWAS summary data of the exposure factor. The clumping process involved setting the linkage disequilibrium parameter r^2 threshold to 0.01 and the genetic distance to 5000 kilobases (kb), further refining the list of SNPs to ensure they met the independence requirement for instrumental variables.

2.4. Validation of Instrumental Variables

In Mendelian Randomization (MR) studies, the reliability of causal inference conclusions heavily relies on the appropriate selection of instrumental variables. To ensure this, the present study strictly adheres to one of the three core assumptions in MR analysis concerning the choice of instrumental variables—the relevance assumption. This assumption stipulates that the instrumental variables must maintain a stable and strong association with the exposure factor. Specifically, SNPs serving as instrumental variables should exhibit a significant correlation with the

exposure factor under investigation, such as the relationship between alcohol intake and the risk of Parkinson's disease. The three main hypotheses of Mendelian randomization studies are shown in **Figure 1**.

To further ensure that the selected instrumental variables comply with this assumption, the study implemented additional measures, including the use of F-statistics to evaluate potential weak instrument bias. In this context, a weak instrument refers to genetic variants that poorly explain variations in the exposure factor, characterized by a weak link with the exposure. It is crucial to note that weak instruments are not synonymous with invalid instruments; there is a fundamental distinction between the two. In practice, the risk of weak instrument bias is particularly pronounced when the sample size is insufficient to provide adequate statistical power. By employing the aforementioned methods, the aim of this study is to minimize such biases to the greatest extent possible, thereby enhancing the accuracy and credibility of the causal inference outcomes.

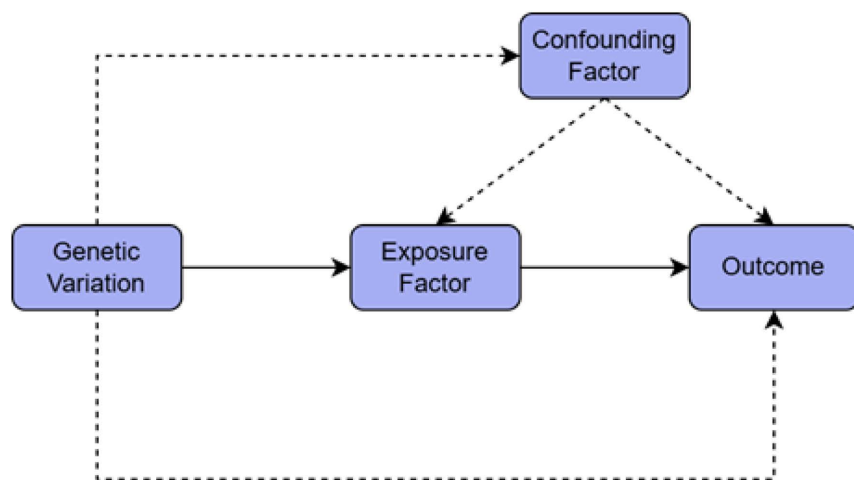


Figure 1. Three main assumptions of MR studies.

2.5. Estimation of Causal Effects

In this study, to accurately assess the causal effects of exposure factors on outcome variables, multiple statistical methods were primarily employed, with a focus on the Inverse Variance Weighted (IVW) method.

1) Inverse Variance Weighted (IVW) Method: The IVW method is the most commonly used approach for estimating causal effects [11]. Its fundamental principle assumes that all selected genetic variants are valid instrumental variables (IVs) and that these IVs affect the outcome variable only indirectly through the exposure factor. Based on this assumption, the IVW method expects that the effect of each IV on the outcome variable is proportional to its effect on the exposure factor. Therefore, the causal effect of the exposure factor on the outcome variable can be estimated by calculating the weighted average of the ratios of the outcome variable effects to the exposure variable effects for each IV, with weights typically being the inverse variances of the IV effect estimates. This method provides a

robust estimate of the causal effect, equivalent to the slope obtained from a linear regression model regressing SNP-outcome associations on SNP-exposure associations.

However, a critical limitation of the IVW method lies in its assumption that the intercept of the regression model is fixed at zero. This implies that if the effect of an IV on the exposure factor is zero, then its effect on the outcome variable should also be zero. This assumption can be challenged in practice due to pleiotropy, where some IVs may directly or indirectly influence the outcome variable, not solely through the exposure factor. In such cases, the causal effect estimate derived from the IVW method may be biased.

2) Causal Effect Estimation: In this study, the IVW method was utilized to calculate the causal effect of the exposure factor on the outcome variable. Specifically, by implementing the aforementioned statistical procedures, we obtained the overall effect size, standard error (which was subsequently transformed into the final Odds Ratio (OR) and its 95% Confidence Interval (CI)), and the significance test value. In this study, a two-tailed test was used by default, and a causal effect was considered statistically significant when the P-value was less than 0.05.

2.6. Statistical Analysis Software and Workflow

The statistical analyses for this study were conducted using R software (version 4.4.1). For the Two-Sample Mendelian Randomization (TSMR) analysis, the primary tool used was the TwoSampleMR package, developed by Gibran Hemani *et al.* This R package is publicly available on GitHub.

For visualization purposes, the study utilized the forest and TwoSampleMR packages within R to generate graphs relevant to the Mendelian Randomization analysis. These included Forest Plots, Scatter Plots, and Funnel Plots, which are essential for visualizing the results and assessing the robustness of the causal inferences made.

3. Results

3.1. Selection of SNPs Related to Alcohol Consumption and Instrumental Variables

By screening the GWAS summary data for alcohol consumption, we initially obtained SNPs with $P < 5 \times 10^{-5}$, LD parameter $r^2 < 0.01$, and genetic distance > 5000 kb. Calculations of r^2 and F statistics for each SNP revealed that the F statistics for the 247 candidate instrumental variables were all greater than 10, excluding weak instrument bias due to insufficient sample size and ensuring the reliability of the MR analysis. The detailed information of each SNP in the drinking-related GWAS data is shown in **Table 1**.

3.2. Validation of Instrumental Variables

The 247 selected instrumental variables all met the criterion of $P < 5 \times 10^{-5}$, indicating that these SNPs are strongly associated with alcohol consumption and

satisfy the relevance assumption for instrumental variables. After integrating the exposure (alcohol consumption) and outcome (type 2 diabetes) data, the corresponding P values, beta values, and standard errors (SE) for these SNPs in both the alcohol consumption and type 2 diabetes databases are summarized in **Table 2**.

Table 1. Detailed information of SNPs in the GWAS summary data for alcohol consumption.

SNP	chr	EA	OA	eaf	P-value	beta	se
rs2216808	1	A	G	0.4077	2.39E-05	-0.01263	0.002988
rs1599180	1	A	G	0.4767	5.51E-06	-0.01369	0.003012
rs785119	1	G	A	0.2244	2.29E-05	0.01266	0.002989
rs72679815	1	T	C	0.01211	4.43E-05	0.01217	0.002979
rs138653631	1	A	G	0.09695	1.13E-05	-0.01307	0.002977
rs12137803	1	T	C	0.0187	1.61E-05	-0.01289	0.002988
rs56170575	1	C	A	0.2796	1.09E-05	0.01325	0.003012
rs77654900	1	T	C	0.3031	1.73E-05	-0.01319	0.00307
rs116209269	1	T	C	0.02164	4.67E-05	-0.01211	0.002974
rs884118	1	T	A	0.1348	4.84E-05	0.01223	0.00301
rs55989809	1	A	C	0.06889	1.68E-05	0.01305	0.003034
rs12404604	1	A	G	0.7811	4.01E-06	0.01374	0.00298
rs74525179	1	C	T	0.02379	2.67E-05	-0.01255	0.002988
rs3009872	1	C	T	0.4336	7.65E-07	-0.01478	0.00299
rs10890472	1	T	G	0.6092	1.23E-05	0.0132	0.00302
rs114544721	1	G	T	0.01485	3.72E-05	0.01236	0.002997
rs11800009	1	T	C	0.05505	1.10E-05	-0.01327	0.003018
rs116748149	1	C	T	0.02423	2.87E-05	0.01248	0.002983
rs10798977	1	G	A	0.6065	1.88E-05	0.01287	0.003007
rs41313906	1	A	G	0.0541	3.92E-05	0.01244	0.003025

Table 2. P-values, beta values, and SE values for SNPs in alcohol consumption and type 2 diabetes datasets.

SNP	Alcohol Consumption			Type 2 Diabetes Datasets		
	P-value	beta	se	P-value	beta	se
rs1000368	1.39E-05	0.01313	0.003021	0.2057	0.0085	0.0068
rs10018066	4.13E-06	-0.01372	0.00298	0.6753	0.0034	0.008
rs10168008	4.04E-05	0.01226	0.002987	0.8215	-0.004	0.0176
rs10739501	3.69E-05	0.01231	0.002983	0.9039	-9.00E-04	0.0075
rs10757838	2.86E-05	0.01271	0.003037	0.507399	-0.0044	0.0066
rs10778818	2.01E-05	0.01277	0.002996	0.9178	-7.00E-04	0.0064
rs10786331	1.75E-05	0.01284	0.002989	0.4824	-0.0046	0.0066
rs10798977	1.88E-05	0.01287	0.003007	0.0465297	0.014	0.007
rs10817503	6.81E-06	-0.01344	0.002988	0.2507	0.0086	0.0075
rs10890472	1.23E-05	0.0132	0.00302	0.0545306	-0.0162	0.0084
rs10892761	5.85E-06	0.01348	0.002974	0.6303	-0.003	0.0063
rs10942435	1.86E-05	0.01315	0.003072	0.6092	-0.0057	0.0111
rs10964652	2.60E-05	0.01255	0.002983	0.8262	-0.0018	0.0084
rs1104608	9.04E-06	-0.01352	0.003046	0.0891107	0.0114	0.0067
rs11074422	8.43E-06	0.01355	0.003041	0.6574	0.0029	0.0066

3.3. Estimation of Causal Effects

The estimation of causal effects using different methods is shown in **Table 3**. Overall, there is no causal relationship between genetically predicted alcohol consumption and the risk of type 2 diabetes (T2D). The IVW method results indicate that there is no statistically significant association between alcohol consumption and an increased risk of T2D (OR = 1.0046, 95% CI: [0.8722, 1.1571], P = 0.9495). Similar results were obtained using the Weighted Median (WM) method (OR = 0.9013, 95% CI: [0.7777, 1.0445], P = 0.1671), Simple Model (OR = 0.8635, 95% CI: [0.6201, 1.2023], P = 0.8635), and Weighted Mode (OR = 0.8306, 95% CI: [0.6258, 1.2023], P = 0.2003) (all Ps > 0.05), which supports the stability of the IVW method results. The results of scatter plot, funnel plot, forest plot and leave one method sensitivity analysis of MR Analysis are shown in **Figures 2-5**.

Table 3. MR estimates of the causal effect of alcohol consumption on type 2 diabetes using different methods.

Method	OR	P-value	LCI	UCI
MREgger	19.20769496	1.32E-11	8.511096121	43.34759476
Weightedmedian	0.901256731	0.16705797	0.77767855	1.044472289
Inversevarianceweighted	1.004577564	0.949489294	0.872196812	1.157050871
Simplemode	0.86349136	0.385753144	0.62012885	1.202358718
Weightedmode	0.83063458	0.200318015	0.625766548	1.102573808

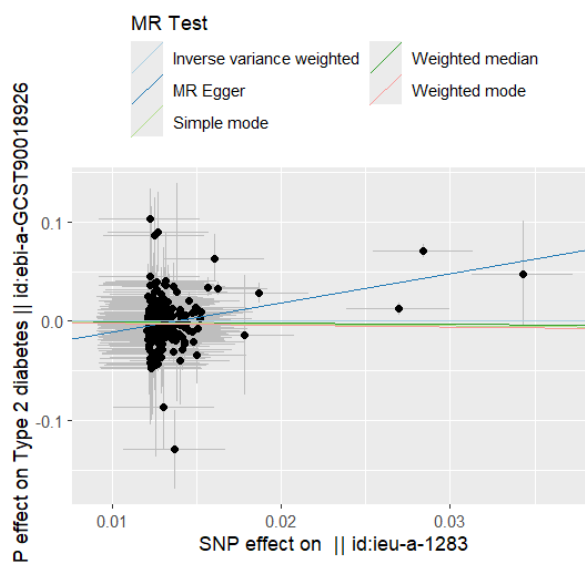


Figure 2. Scatter plot of MR analysis.

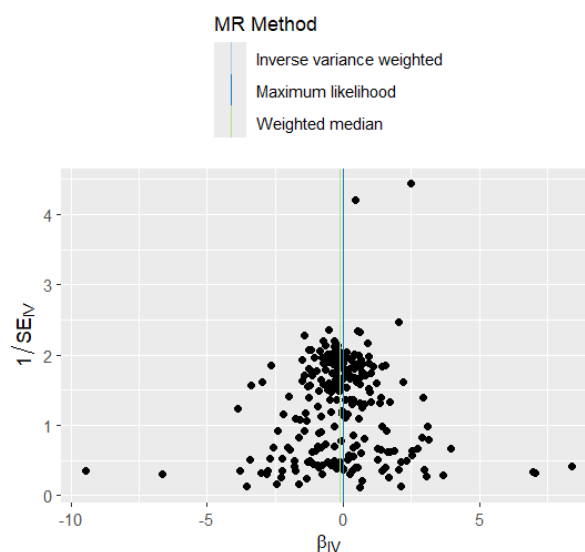


Figure 3. Funnel plot of MR analysis.

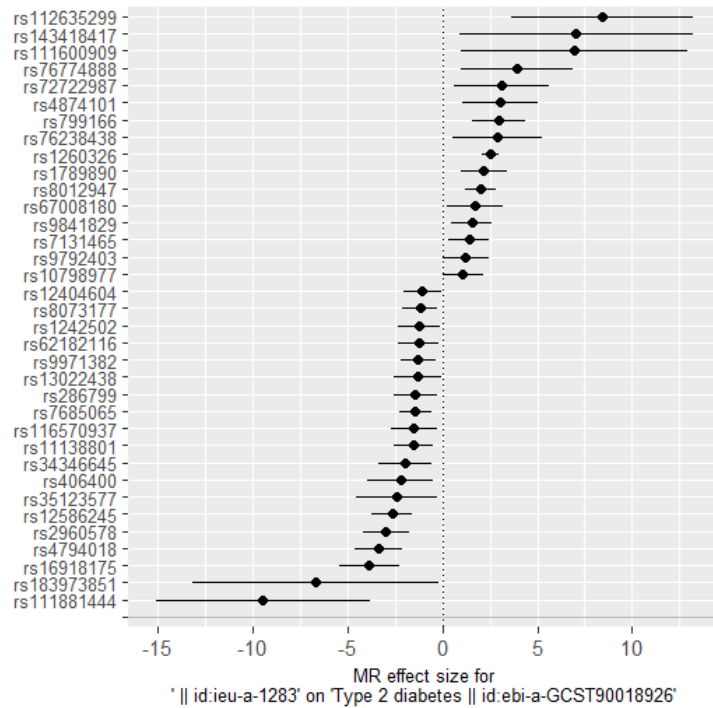


Figure 4. Forest plot of MR analysis.

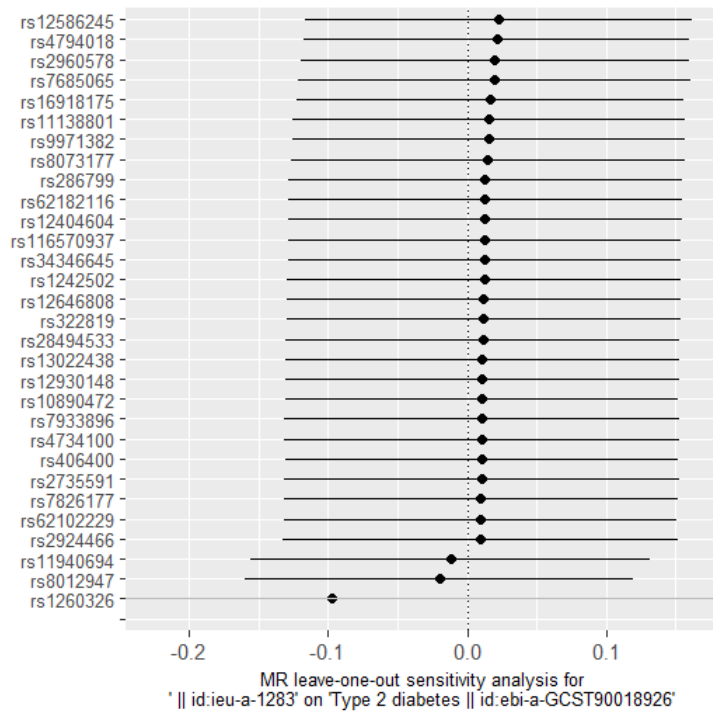


Figure 5. Leave-one-out sensitivity analysis results.

4. Discussion

4.1. Summary

This paper reviews the research progress on the complex associations between

alcohol consumption and T2D and explores their causal relationship using Mendelian Randomization (MR) methods. The relationship between alcohol consumption and T2D is not linear; moderate alcohol consumption may improve insulin sensitivity and reduce inflammation, thereby providing some protective effect against T2D. However, excessive alcohol consumption increases the risk of the disease, particularly through weight gain and exacerbation of insulin resistance. An individual's genetic background may influence drinking habits, with certain genetic variants being associated with higher drinking tendencies, and these variants may also affect the risk of T2D.

The aim of this study was to further investigate the causal link between genetic predisposition to alcohol consumption and the risk of T2D using MR methods. The study utilized European population samples from the IEU Open GWAS Project database, including GWAS summary data related to alcohol consumption and T2D. The experimental methods included the selection, validation of instrumental variables, and estimation of causal effects. During the selection process, SNPs significantly associated with alcohol consumption were chosen, and SNPs in linkage disequilibrium were removed. Furthermore, the causal effect was estimated using the Inverse Variance Weighting (IVW) method.

The results show that, based on the selected 247 instrumental variables, there is no significant causal relationship between genetically predicted alcohol consumption and the risk of T2D. The IVW method results indicate that there is no statistically significant association between alcohol consumption and an increased risk of T2D (OR = 1.0046, 95% CI: [0.8722, 1.1571], P = 0.9495).

4.2. Limitations of This Study

The conclusions drawn from this study contradict those of previous studies, possibly due to the weak influence of available genetic variants on exposure factors. For example, while genes associated with alcohol consumption can influence drinking behavior to some extent, they are also affected by postnatal factors such as lifestyle, social environment, family, social interactions, and psychological factors.

Regarding the impact of drinking behavior on T2D, the focus is primarily on the dose of alcohol consumed per occasion. Although genes associated with alcohol consumption can affect the quantity and activity of alcohol-metabolizing enzymes in the body, they only relate to the upper limit of alcohol consumption per occasion and do not have a direct relationship with the actual amount of alcohol consumed.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Qian, X., Wu, F., Pu, Z., Yan, Q., Cheng, M., Guo, F. and Li, Y. (2019) Dietary Patterns

- among Residents Aged 35 Years and Older in Shanghai and Their Impact on Type 2 Diabetes Mellitus. *Shanghai Journal of Preventive Medicine*, 2, 97-103.
<https://doi.org/10.19428/j.cnki.sjpm.2019.19152>
- [2] Yi, S., Peng, J. and Pang, R. (2018) Efficacy of Cognitive Behavioral Therapy for Type 2 Diabetes Patients with Comorbid Depression and Anxiety. *Chinese Journal of Clinical Psychology*, 2, 414-416, 420.
<https://doi.org/10.16128/j.cnki.1005-3611.2018.02.044>
- [3] International Diabetes Federation (IDF) (2021) IDF Diabetes Atlas 10th Edition
<http://www.diabetesatlas.org>
- [4] Zheng, Y., Jiang, S., Guo, Y., Sun, S., Wu, F. and Shi, Y. (2016) Analysis of Major Chronic Diseases and Risk Factors among Middle-Aged and Elderly Populations in China. *Shanghai Journal of Preventive Medicine*, 6, 353-357.
<https://doi.org/10.19428/j.cnki.sjpm.2016.06.001>
- [5] Wang, Y., Dong, X., Zhu, J., Wang, N., Chen, Y., Jiang, Q. and Fu, C. (2021) Prospective Study on Alcohol Consumption and the Risk of Type 2 Diabetes among Adult Males in Rural Deqing County, Zhejiang Province. *Shanghai Journal of Preventive Medicine*, 7, 577-582.
- [6] Xu, L. and Zhang, H. (2021) Research Progress and Prospects on the Relationship between Moderate Alcohol Consumption and Diabetes Prevention and Control. *Modern Preventive Medicine*, 10, 1849-1853.
- [7] Zhu, Y. (2019) Cross-Sectional Study on the Relationship between Alcohol Consumption and the Prevalence of Diabetes and Prediabetes among Residents of Baota District, Yan'an City. Master's Thesis, Yan'an University.
<https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201902&filename=1019856807.nh>
- [8] Rose, R.J. and Kaprio, J. (2010) Genetics, Environment, and Adolescent Smoking and Drinking Behavior: Retrospective and Prospective Findings from the Finnish Twin Study. *Acta Psychologica Sinica*, 10, 1062-1072.
- [9] Kimura, M. and Higuchi, S. (2011) Genetics of Alcohol Dependence. *Psychiatry and Clinical Neurosciences*, 65, 213-225.
<https://doi.org/10.1111/j.1440-1819.2011.02190.x>
- [10] Yu, T., Xu, W., Su, Y. and Li, J. (2021) Basic Principles, Methods, and Limitations of Mendelian Randomization Studies. *Chinese Journal of Evidence-Based Medicine*, 21, 1227-1234.
- [11] Bowden, J. and Holmes, M.V. (2019) Meta-Analysis and Mendelian Randomization: A Review. *Research Synthesis Methods*, 10, 486-496.
<https://doi.org/10.1002/jrsm.1346>